

Durham E-Theses

*Gene expression during late embryogenesis in pea
(Pisum sativum L)*

Melinda Jane Mayer

How to cite:

Mayer, Melinda Jane (1993) Gene expression during late embryogenesis in pea (*Pisum sativum* L).
Doctoral thesis, Durham University.

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a <https://etheses.durham.ac.uk/id/eprint/5722/> is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

GENE EXPRESSION DURING LATE EMBRYOGENESIS IN PEA
(*Pisum sativum* L.)

The copyright of this thesis rests with the author.
No quotation from it should be published without
his prior written consent and information derived
from it should be acknowledged.

A thesis submitted by Melinda Jane Mayer, B.Sc (Bristol) in
accordance with the requirements of the University of Durham
for the degree of Doctor of Philosophy.

Department of Biological Sciences
August 1993



14 JAN 1994

TO MY PARENTS

ABSTRACT

GENE EXPRESSION DURING LATE EMBRYOGENESIS IN PEA (*Pisum sativum* L.).

A thesis submitted by Melinda Jane Mayer, B.Sc.(Bristol) in accordance with the requirements of the University of Durham for the degree of Doctor of Philosophy. Department of Biological Sciences, August 1993.

Two cDNA libraries were constructed from desiccating pea cotyledons. Differential screening of the libraries with cDNA from an earlier developmental stage (physiological maturity) demonstrated that the abundant message population during dehydration shows some noticeable differences to the message populations present before desiccation.

Clones hybridising to a polyubiquitin probe were isolated from a cDNA library. These clones were identified as messages for the two types of ubiquitin extension proteins (with 52 and 79 residue tails), already characterised in other species as being involved in ribosome biogenesis. The pea ubiquitin extension tail amino acid sequences showed considerable homology to tails from other plants, animals, yeast and protozoa, including a nuclear localisation site and a putative zinc-binding nucleic acid binding domain, the positions of which are conserved within the tail sequences. Sequencing of a second polyubiquitin cDNA from pea leaf demonstrated that pea contains a ubiquitin multigene family of at least four members.

The expression of several genes associated with plant response to stress and two abundant seed messages (*Leg A* and *J*) was examined in developing and dehydrating cotyledons and axes. This confirmed conspicuous variations in the message levels of the genes examined as the cotyledons aged, with different members of the ubiquitin and legumin multigene families showing differential expression with age. It was also demonstrated that the expression pattern of certain messages in the cotyledons was different to that in the axes and other seed tissues. This was confirmed by an analysis of total and albumin protein fractions in cotyledons and axes.

The effect on specific message and protein levels of premature desiccation treatments indicated that the temporal expression of several seed genes is related to the state of hydration of the seed, artificial desiccation leading to premature maturation. Seed storage protein message and protein levels were especially increased by premature desiccation. Legumin seed storage protein messages were also shown to be responsive to exogenous ABA applied to immature cotyledons during the seed filling stage. However, the other stress-related messages examined in pea (ubiquitin and a pea putative metallothionein) were not responsive to exogenous ABA at this developmental stage.

CONTENTS

CONTENTS	i
ABBREVIATIONS	xiv
FIGURES AND TABLES	xvii
CHAPTER 1 - INTRODUCTION	1
1.1 PLANT RESPONSES TO DESICCATION	1
1.1.1 The Effect of Water Stress	1
1.1.1.1 Effects of a Loss of Turgor	1
1.1.1.2 Effects of Cell Flaccidity	2
1.1.1.3 Effects of Prolonged Dehydration	3
1.1.2 Adaptation Strategies	5
1.2 SEED DEVELOPMENT	7
1.2.1 The Seed as a Biological System	7
1.2.2 The Events from Fertilisation to Growth	9
1.2.3 Water Relations of the Seed	12
1.2.3.1 Moisture Content of the Seed	13
1.2.3.2 Desiccation Tolerance in the Seed	14
1.2.3.3 Precocious Germination and Water Loss	17
1.2.3.4 The Effects of Premature Desiccation and Embryo Isolation on Seed Metabolism	20
1.2.4 Changes in Hormone Levels in the Seed	25
1.2.4.1 The Involvement of ABA in Plant Systems	25

1.2.4.2	Changes in ABA Levels and Sensitivity During Natural and Premature Maturation	27
1.2.4.3	The Effects of ABA on Message and Protein Levels	30
1.2.5	Changes in mRNA and Protein Levels in the Developing and Germinating Seed	33
1.2.5.1	Changes in mRNA Mass and Complexity during Development	33
1.2.5.2	Patterns of mRNA and Protein Accumulation	37
1.2.5.3	The Effect of Desiccation on mRNA and Protein Populations	43
1.2.5.4	The Regulation of mRNA and Protein Levels during Seed Development and Germination	44
1.3	LEA PROTEINS AND DEHYDRINS	50
1.3.1	Group 1 LEAs	52
1.3.2	Group 2 LEAs	53
1.3.3	Group 3 LEAs	54
1.3.4	Other Unclassified LEAs	55
1.3.5	Induction and Regulation of LEAs	57
1.3.6	Proposed Roles of LEAs	59
1.4	GENES AND PROTEINS USED IN EXPRESSION STUDIES	61
1.4.1	Legumin Storage Protein	61
1.4.1.1	The <i>Leg A</i> Subfamily	63
1.4.1.2	The <i>Leg J</i> Subfamily	64
1.4.1.3	Expression of Legumin Genes and Proteins	66
1.4.1.4	Control of Legumin Gene Expression	71

1.4.2	The Wheat Em Late Embryogenesis Abundant Protein	74
1.4.3	Ubiquitin	79
1.4.3.1	Ubiquitin-protein Conjugation and Protein Degradation	80
1.4.3.2	Ubiquitin Multigene Families	83
1.4.3.3	Differential Expression of Ubiquitin Genes	88
1.4.4	Chaperonin 60 - A Molecular Chaperone	94
1.4.5	<i>PsMT_A</i> - A Putative Metallothionein Gene	96
1.4.6	CpTI - Cowpea Trypsin Inhibitor Protein	99
1.5	AIMS OF THE PROJECT	102
	CHAPTER 2 - MATERIALS	105
2.1	MATERIALS FOR PLANT GROWTH	105
2.2	GLASSWARE AND PLASTICWARE	105
2.3	KITS FOR MOLECULAR BIOLOGY	105
2.4	BIOCHEMICALS, CHEMICALS AND CONSUMABLES	106
2.5	LINKERS, ADAPTORS AND PRIMERS	107
2.6	PLASMID DNA FOR HYBRIDISATION AND SUBCLONING	108
2.6.1	PJC5.2	108
2.6.2	PRC3.1	108
2.6.3	PCU1 and PCU2	109
2.6.4	PMuge Kla	109
2.6.5	PASD1	109

2.6.6 <i>PsMT_A</i>	109
2.7 ANTISERUM	110
CHAPTER 3 - METHODS	111
3.1 GROWTH AND TREATMENT OF PLANT MATERIAL	111
3.1.1 Growth and Harvesting of Peas	111
3.1.2 Premature Desiccation Treatments	111
3.1.2.1 Air Drying	111
3.1.2.2 Silica Drying	112
3.1.3 Treatment with Exogenous ABA	112
3.2 RNA EXTRACTION AND ISOLATION PROCEDURES	112
3.2.1 Extraction of Total RNA from Cotyledons	112
3.2.2 Extraction of Total RNA from Embryonic Axes	114
3.2.3 Isolation of Poly(A) ⁺ RNA	115
3.2.3.1 Oligo(dT)-Cellulose Spun Column Chromatography	115
3.2.3.2 Isolation of mRNA using Magnetic Particles	116
3.3 ANALYSIS OF NUCLEIC ACIDS BY GEL ELECTROPHORESIS	117
3.3.1 Analysis of RNA by Formaldehyde Gel Electrophoresis	117
3.3.2 Analysis of DNA by Agarose Gel Electrophoresis	118
3.3.2.1 Restriction and Purification of DNA	118
3.3.2.2 Size Analysis of DNA	119
3.3.2.3 Isolation of DNA Fragments by Electroelution	119

3.3.2.4 Isolation of DNA Fragments using Silica Fines	119
3.3.2.5 Alkaline Agarose Gels	120
3.4 ANALYSIS OF NUCLEIC ACIDS BY BLOTTING AND HYBRIDISATION	121
3.4.1 Northern Blotting of RNA Formaldehyde Gels	121
3.4.2 RNA Dot Blots	121
3.4.3 Hybridisation of DNA Probes to Immobilised RNA	122
3.4.3.1 Preparation of Radioactive DNA Probes	122
3.4.3.2 Treatment of Blots and Hybridisation of Probes	123
3.4.4 Southern Blotting of DNA Agarose Gels	124
3.4.5 Hybridisation of DNA Probes to Immobilised DNA	124
3.4.6 Colony Hybridisation	125
3.5 CONSTRUCTION OF A cDNA LIBRARY FROM DESICCATING PEA COTYLEDONS	126
3.5.1 Construction of a cDNA Library using PUC18 as a Cloning Vector	126
3.5.1.1 cDNA Synthesis	126
3.5.1.2 Preparation of Vector	128
3.5.1.3 Ligation of cDNA to Vector using EcoRI/ NotI Adaptors	128
3.5.1.4 Ligation of cDNA to Vector using EcoRI Phosphorylated Linkers	129
3.5.1.5 Ligation of cDNA to Vector using BamHI	

12mer Linkers	130
3.5.1.6 Assessment of Linker Activity	130
3.5.1.7 Transformation of Bacteria with Plasmid DNA	131
3.5.1.8 Preparation of Competent Cells	132
3.5.1.9 Preparation of Plasmid DNA for Restriction Analysis and Subcloning	133
3.5.1.10 Preparation of Plasmid DNA for Direct Sequencing	134
3.5.2 Construction of a cDNA Library using Lambda-gt10 as a Cloning Vector	135
3.5.2.1 Growth of Host Cells and Lambda	136
3.5.2.2 Ligation of cDNA to Lambda-gt10 Arms using EcoRI/NotI Adaptors	136
3.5.2.3 Titre, Amplification and Storage of Lambda cDNA Libraries	138
3.6 SCREENING OF LAMBDA cDNA LIBRARY WITH DNA PROBES	138
3.6.1 Immobilisation of Plated Lambda Library DNA to Nitrocellulose Filters	138
3.6.2 Hybridisation of DNA Probe to Immobilised Lambda cDNA Library	139
3.7 PREPARATION OF A SUBTRACTED cDNA PROBE	140
3.7.1 Synthesis of Poly(A) ⁺ RNA	141
3.7.2 Synthesis and Treatment of cDNA	141
3.7.3 Hybridisation of 18DAF Poly(A) ⁺ RNA to 24,26,28DAF	

Radiolabelled cDNA	142
3.7.4 Hydroxyapatite Chromatography	143
3.7.5 Radiolabelling of Subtracted cDNA to High Specific Activity	144
3.7.6 Hybridisation of Subtracted cDNA Probe to Immobilised Lambda cDNA Library	145
3.8 SIMULTANEOUS SCREENING OF ISOLATED LAMBDA cDNA CLONES	145
3.9 PREPARATION AND ANALYSIS OF CLONED cDNA INSERTS	146
3.9.1 Preparation of Recombinant Lambda DNA for Analysis	146
3.9.1.1 Quigen Midiprep	146
3.9.1.2 Small-scale Isolation of Lambda DNA	148
3.9.1.3 Restriction of Recombinant Lambda DNA	149
3.9.2 Amplification and Isolation of cDNA Clones by PCR Technology	150
3.9.2.1 Amplification of Inserts from Recombinant Lambda-cDNA Clones by PCR	150
3.9.2.2 Amplification of Lambda-cDNA Library for use as a Probe	151
3.9.3 Subcloning of PCR products	151
3.9.3.1 Sequencing of Subclones	152
3.10 PROTEIN EXTRACTION AND ANALYSIS	152
3.10.1 Extraction of Total Protein from Pea Cotyledons and Embryonic Axes	152
3.10.1.1 Extraction in PBS	152
3.10.1.2 Extraction in SDS Sample Buffer	153

3.10.2 Pulse Labelling and Extraction of Total Protein from Pea Cotyledons	153
3.10.3 Extraction of Albumins from Pea Cotyledons	154
3.10.4 Analysis of Proteins by Polyacrylamide Gel Electrophoresis	154
3.10.4.1 One-dimensional Gel Electrophoresis	154
3.10.4.2 Fluorography of ¹⁴ C-labelled Protein separated by PAGE	155
3.10.5 Analysis of Proteins by Western Blotting	156
3.10.5.1 Immobilisation of Proteins on Nitrocellulose	156
3.10.5.2 Immunological Detection of Labelled Proteins	156
 CHAPTER 4 - CONSTRUCTION AND SCREENING OF A cDNA LIBRARY FROM DESICCATING PEA COTYLEDONS	 158
4.1 RESULTS	158
4.1.1 The Construction of a cDNA Library from Desiccating Pea Cotyledons using PUC18 as a Cloning Vector	158
4.1.2 The Construction of a cDNA Library from Desiccating Pea Cotyledons using Lambda-gt10 as a Cloning Vector	164
4.1.3 Screening of the Lambda-cDNA Libraries with a Probe for <i>Leg J</i>	170
4.1.4 Screening of Lambda-cDNA Libraries with Probes for Plant Stress Protein Messages	173

4.1.4.1	Cowpea Trypsin Inhibitor Protein	174
4.1.4.2	<i>PsMT_A</i> - A Pea Putative Metallothionein	175
4.1.4.3	Wheat Em Late Embryogenesis Abundant Polypeptide	175
4.1.4.4	Pea Leaf Polyubiquitin	178
4.1.5	An Investigation of Pea Leaf and Cotyledon Ubiquitin Messages	179
4.1.5.1	Sequence of a Second Polyubiquitin cDNA from Pea Leaves	179
4.1.5.2	Sequences of Ubiquitin cDNA Clones Isolated from Desiccating Pea Cotyledons	184
4.1.6	Differential Screening of 24,26,28DAF cDNA Libraries	192
4.1.6.1	24,26,28DAF cDNA Library Screened with 18DAF cDNA	192
4.1.6.2	Screening of the 24,26,28DAF cDNA Library with a Subtracted Probe	193
4.1.6.3	Screening of Selected Clones from the First 24,26,28DAF cDNA Library with 18DAF cDNA	193
4.1.6.4	Screening of Selected Clones from the Second 24,26,28DAF cDNA Library with 18DAF cDNA	199
4.1.6.5	Analysis and Subcloning of Differentially Expressed Clones Isolated from the Second 24,26,28DAF Lambda-cDNA Library	203
4.2	DISCUSSION	208

4.2.1 Construction of a cDNA Library from Desiccating Pea Cotyledons	208
4.2.2 The Occurrence of Stress Protein Messages in Desiccating Pea Cotyledons	213
4.2.2.1 Cowpea Trypsin Inhibitor Protein	213
4.2.2.2 Pea <i>PsMT_A</i>	213
4.2.2.3 Wheat Em Polypeptide	214
4.2.2.4 Pea Polyubiquitin	215
4.2.3 An Investigation of cDNAs Coding for Pea Ubiquitin	215
4.2.3.1 Attempted Isolation of Ubiquitin Clones from the Second 24,26,28DAF cDNA Library	215
4.2.3.2 Isolation of Two Pea Ubiquitin Extension Protein cDNAs from the First 24,26,28 DAF cDNA Library	217
4.2.3.3 Polyubiquitin cDNAs in Pea	219
4.2.4 Differential Screening of 24,26,28DAF cDNA Libraries	221
4.2.4.1 24,26,28DAF cDNA Library Screened with 18DAF cDNA	221
4.2.4.2 Screening of the 24,26,28DAF cDNA Library with a Subtracted Probe	223
4.2.4.3 Screening of Selected Clones from the First 24,26,28DAF cDNA Library with 18DAF cDNA	224
4.2.4.4 Screening of Selected Clones from the	

Second 24,26,28DAF cDNA Library with 18DAF cDNA	226
CHAPTER 5 - CHANGES IN THE LEVELS OF SPECIFIC MESSAGES AND PROTEINS IN DEVELOPING AND TREATED PEA SEEDS	232
5.1 RESULTS	235
5.1.1 The Effect of Cotyledon Age on Specific Message Levels	235
5.1.1.1 <i>PsMT_A</i>	236
5.1.1.2 Pea Ubiquitin	236
5.1.1.3 <i>Leg J</i>	246
5.1.1.4 <i>Leg A</i>	247
5.1.2 The Effect of Plant Tissue - Specific Message Levels in the Developing Embryonic Axis and in Non-seed Tissues	248
5.1.2.1 <i>PsMT_A</i>	249
5.1.2.2 Pea Ubiquitin	249
5.1.2.3 <i>Leg J</i>	251
5.1.2.4 <i>Leg A</i>	253
5.1.3 The Effect of Premature Desiccation on Specific Message Levels in Immature Cotyledons	254
5.1.3.1 <i>PsMT_A</i>	255
5.1.3.2 Pea Ubiquitin	261
5.1.3.3 <i>Leg J</i>	261
5.1.3.4 <i>Leg A</i>	264
5.1.4 The Effect of Exogenous ABA on Specific Message	

Levels in Immature Cotyledons	265
5.1.4.1 <i>PsMT_A</i>	265
5.1.4.2 Pea Ubiquitin	266
5.1.4.3 <i>Leg J</i>	266
5.1.4.4 <i>Leg A</i>	269
5.1.5 Em Message Levels in Pea	270
5.1.6 Cowpea Trypsin Inhibitor Levels in Pea	271
5.1.7 Hybridisation of the PM46 cDNA Insert to Pea RNA	271
5.1.8 The Effect of Seed Age and Tissue on the Total and Albumin Protein Fractions in Developing Pea Seeds	274
5.1.9 The Effect of Seed Age and Premature Desiccation on Levels of the Molecular Chaperone Chaperonin 60	277
5.1.10 The Effect of Premature Desiccation on Cotyledonary Total Protein	280
5.2 DISCUSSION	285
5.2.1 The Analysis of Northern and Dot Blot Results	285
5.2.2 The Effect of Cotyledon Age on Specific Message Levels in Developing Pea Seeds	288
5.2.3 The Effect of Plant Tissue - Specific Message Levels in the Developing Axis and in Non-seed Tissues	298
5.2.4 The Effect of Premature Desiccation on Specific Message Levels in Immature Cotyledons	303
5.2.5 The Effect of Exogenous ABA on Specific Message Levels in Immature Cotyledons	310

5.2.6	Em Message Levels in Pea	318
5.2.7	Cowpea Trypsin Inhibitor Levels in Pea	319
5.2.8	Hybridisation of the PM46 cDNA Insert to Pea Total RNA	320
5.2.9	The Effect of Seed Age and Tissue on the Total and Albumin Protein Fractions in Developing Pea Seeds	321
5.2.10	The Effect of Cotyledon Age and Premature Desiccation on Levels of Chaperonin 60 Protein	324
5.2.11	The Effect of Premature Desiccation on Cotyledonary Total Protein	326
CHAPTER 6 - CONCLUSIONS		330
6.1	CHANGES IN MESSAGE AND PROTEIN LEVELS DURING PEA SEED DEVELOPMENT	330
6.2	THE EFFECT OF PREMATURE DESICCATION ON MESSAGE AND PROTEIN LEVELS IN DEVELOPING PEA COTYLEDONS	335
6.3	THE EFFECT OF EXOGENOUS ABA ON MESSAGE LEVELS IN IMMATURE PEA COTYLEDONS	340
REFERENCES		
ACKNOWLEDGEMENTS		

ABBREVIATIONS

ABA	abscisic acid
ATP	adenosine-5'-triphosphate
b	bases
bp	base pair
BSA	bovine serum albumin
cdNA	complementary deoxyribonucleic acid
cv	cultivar
DAF	days after flowering
DAI	days after imbibition
DAP	days after pollination
dATP	deoxyadenosine-5'-triphosphate
dCTP	deoxycytosine-5'-triphosphate
DEPC	diethyl pyrocarbonate
dpa	days post-anthesis
DTT	dithiothreitol
dGTP	deoxyguanosine-5'-triphosphate
DMSO	dimethylsulphoxide
DNase	deoxyribonuclease
dTTP	deoxythymidine-5'-triphosphate
<i>E. coli</i>	<i>Escherichia coli</i>
Em	early methionine-labelled polypeptide
EDTA	ethylene diamine tetraacetic acid
EGTA	ethylene glycol tetraacetic acid
HAI	hours after imbibition
HAP	hydroxyapatite

HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulphonic acid
HGT	high gelling temperature (agarose)
hsp	heat shock protein
iaa	isoamyl alcohol
kb	kilobase
kDa	kiloDalton
LEA	late embryogenesis-abundant
leuNAase	L-leucyl-beta-naphthylamidase
MES	4-Morpholineethansulphonic acid
MOPS	3-(N-Morpholino) propane-sulphonic acid
mRNA	messenger ribonucleic acid
MW	relative molecular weight
PEG	polyethylene glycol
OD	optical density
oligo(dt)	oligodeoxythymidylic acid
OLB	oligonucleotide labelling buffer
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
pI	isoelectric point
pfu	plaque forming unit
PMSF	phenylmethanesulphonyl fluoride
poly(A) ⁺ RNA	polyadenylated RNA (=mRNA)
PPO	2,5 diphenyloxazole
PVP	polyvinylpyrrolidone
RNase	ribonuclease

rRNA	ribosomal RNA
Rubisco	ribulose biphosphate carboxylase
SA-PMP	streptavidin-paramagnetic particle (Promega)
SDS	sodium dodecyl sulphate
SSC	saline sodium citrate
TCA	trichloroacetic acid
TEMED	NNN'N'-tetramethylethylene-diamine
uv	ultraviolet

FIGURES AND TABLES

- Figure 1.** Changes in mean seed fresh weights (A) and dry weights (B) during development of pea cv. Feltham First, taken from Boulter *et al* (1990). 11
- Figure 2.** cDNA synthesis from 18DAF and 24,26,28DAF RNA isolated from pea cotyledons for use in the construction of a PUC18 cDNA library. 160
- Figure 3.** cDNA synthesis from 24,26,28DAF RNA isolated from pea cotyledons for use in the construction of the first lambda cDNA library 165
- Figure 4.** First strand cDNA synthesis from 18DAF and 24,26,28DAF RNA from pea cotyledons for use in the differential screening of cDNA libraries and for the synthesis of double stranded cDNA for the construction of the second 24,26,28DAF lambda cDNA library. 168
- Figure 5.** PCR amplification product of a clone isolated from a screen of the first 24,26,28DAF cDNA library with a probe for *Leg J* (PJC5.2). 172
- Figure 6.** Restriction map of lambda-gt10. 176
- Figure 7.** Restriction and hybridisation analysis of recombinant lambda DNA from clones isolated from a screen of the first 24,26,28DAF cDNA library with a probe for pea polyubiquitin (PCU1). 180

Figure 8.	A comparison of the sequences of two polyubiquitin cDNAs isolated from pea leaves.	182
Figure 9.	A comparison of the sequences of two ubiquitin extension protein cDNAs isolated from desiccating pea cotyledons.	185
Figure 10.	A comparison of the deduced amino acid sequence of the pea 52 residue ubiquitin extension protein tail with 52 and 53 residue tails from a range of other species.	187
Table 1.	Percentage amino acid sequence homologies between deduced pea 52 amino acid ubiquitin extension protein tail and 52 and 53 residue tail sequences from other plants, animals, yeast and micro-organisms.	189
Figure 11.	A comparison of the deduced amino acid sequence of the pea 79 residue ubiquitin extension protein tail with 80-81 residue tails from other plants and a 76 residue tail from yeast	190
Table 2.	Percentage amino acid sequence homologies between the deduced pea 79 amino acid ubiquitin extension protein tail sequence and 76-81 residue tail sequences from other plants and yeast.	191
Figure 12.	PCR amplification of stored phage particle solutions.	195
Figure 13.	Differential screening by Southern blot	196

hybridisation of 50 selected clones from the first 24,26,28DAF cDNA library amplified by PCR.

- Figure 14.** Artifactual sequence obtained from the subcloning of the PCR product of clone 35, isolated from the first 24,26,28DAF cDNA library, into EcoRI-restricted PUC18. 198
- Figure 15.** Differential screen of 49 selected clones isolated from the second 24,26,28DAF cDNA library. 201
- Figure 16.** Restriction analysis of recombinant lambda DNA from three clones isolated from a differential screen of the second 24,26,28DAF cDNA library. 204
- Figure 17.** Sequence of the partial cDNA insert from clone PM46, isolated and subcloned from the second 24,26,28DAF cDNA library. 205
- Figure 18.** Sequences of the cloning sites of two clones isolated from the second 24,26,28DAF cDNA library, showing a loss of cDNA inserts. 207
- Figure 19.** Northern blot analysis of total RNA from developing pea cotyledons, probed with a) *PsMT_A*; b) PCU1 (pea polyubiquitin); c) PJC5.2 (*Leg J*); d) PRC3.1 (*Leg A*). 237
- Figure 20.** Dot blot analysis of total RNA from developing pea cotyledons, embryonic axes and non-seed tissues, probed with a) *PsMT_A*; b) PCU1 (pea polyubiquitin); c) PJC5.2 240

- (*Leg J*).
- Figure 21.** Densitometric analysis of dot blots of total RNA from developing pea cotyledons, embryonic axes and non-seed tissues probed with a) *PsMT_A*; b) PCU1 (pea polyubiquitin); c) PJC5.2 (*Leg J*). 242
- Figure 22.** Northern blot analysis of total RNA extracted from developing pea embryonic axes probed with PJC5.2 (*Leg J*). 252
- Figure 23.** Dot blot analysis of total RNA extracted from pea cotyledons from untreated developing pods and from pods isolated from the plant and exposed to premature desiccation. 256
- Figure 24.** Densitometric analysis of dot blots of total RNA extracted from pea cotyledons from untreated developing pods and from pods isolated from the plant and exposed to premature desiccation probed with a) *PsMT_A*; b) PCU1 (pea polyubiquitin); c) PJC5.2 (*Leg J*); d) PRC3.1 (*Leg A*). 258
- Figure 25.** Northern blot analysis of total RNA extracted from immature pea cotyledons from pods which were either untreated, prematurely dried by air or silica drying, or incubated in water. 262
- Figure 26.** Northern blot analysis of total RNA extracted from pea cotyledons from untreated pods and from pods isolated from the plant and 267

incubated with varying concentrations of ABA and water.

- Figure 27.** Northern blot analysis of total RNA from 272
developing pea cotyledons, probed with PM46.
- Figure 28.** Polyacrylamide gel electrophoresis of albumin 275
and total protein fractions from developing
and desiccated cotyledons and embryonic axes.
- Figure 29.** Western blot analysis of total protein 278
extracted from developing cotyledons and leaf
tissues reacted with antiserum for chaperonin
60.
- Figure 30.** Western blot analysis of total protein from 281
pea cotyledons from immature and prematurely
desiccated pods reacted with antiserum for
chaperonin 60.
- Figure 31.** Fluorograph of total protein extracted from 284
pea cotyledons from immature and prematurely
desiccated pods, pulse labelled for four hours
with a ^{14}C -amino acid mixture.
- Table 3.** Radioactivity (cpm) of total protein extracts 285
from immature and prematurely desiccated
cotyledons pulse labelled with a ^{14}C -amino acid
mixture.

CHAPTER 1 - INTRODUCTION

1.1 PLANT RESPONSES TO DESICCATION

1.1.1 The Effect of Water Stress

Water is essential in plant cells, both for the maintenance of cellular and macromolecular structures and for the biological activity of the cell. Higher plants have the ability to maintain a water potential that is much higher than that of their external environment, as long as a water supply is available. Apart from the free water in the protoplasm and associated with the cell membranes, plant cells also contain a certain amount of 'bound' water, within macromolecular structures such as proteins, that can only be removed by extreme temperatures. If the water supply fails, or plants are exposed to other external desiccating conditions such as high winds or intense heat, the resulting loss of cellular water can have several effects on the plant cell, varying in severity and reversibility (reviewed by Levitt, 1980 and by Bewley and Black, 1985 and references therein).

1.1.1.1 Effects of a Loss of Turgor

A relatively small loss of water from the plant will lead to a fall in cell turgor. In response to this, plant growth, which requires full turgor, is inhibited both rapidly and severely, preventing cell enlargement and



division. The second response to turgor loss is the closure of stomata, which is controlled by the turgor of the surrounding guard and epidermal cells. This in turn leads to metabolic disturbance, as the decrease or cessation of gas exchange causes a rapid drop in photosynthesis and a slower fall in respiration. Dehydration has also been shown to inhibit the activity of several metabolic enzymes including the chlorophyll a/b protein in corn, PSII in sunflower and ribulosebisphosphate carboxylase in bean, cotton, barley and wheat (reviewed in Levitt, 1980). All of these effects are reversible upon the supply of water being returned. However, an increase in dehydration will lead to a greater loss of cellular water and the production of flaccidity or wilting.

1.1.1.2 Effects of Cell Flaccidity

Dehydration to the point of cell flaccidity has a complex range of effects upon plant development, depending upon the point in the plant life cycle at which it is exposed to water stress, the severity of and length of exposure to stress and the type of plant. Stress during early reproduction can inhibit development and decrease yields in some plants (e.g. soybean, cowpea) but increases yields in others (e.g. potato, maize, cotton), while dehydration during the later stages of development often accelerates maturity (reviewed by Levitt, 1980 and references therein). Desiccation to cell flaccidity also produced the metabolic effects mentioned above due to stomatal closure and enzyme

inhibition. As with the loss of turgor these metabolic effects are reversible, and although the developmental inhibition may give a loss in yield or reproductive potential the plants suffer no lasting damage. But if the water stress continues, plant cells experience more serious injuries.

1.1.1.3 Effects of Prolonged Dehydration

Continued dehydration has a variety of structural and metabolic effects on plant cells (reviewed by Levitt, 1980 and by Bewley and Black, 1985). The loss of water and subsequent rehydration both cause extensive damage to membranes and proteins, affecting cell and plastid structure and functions, and inactivating enzymes. The disruption of plasma membranes leads not only to an increase in permeability, allowing the leakage of ions and solutes, but also to the rearrangement or loss of protein and ion transport systems which, together with the increased concentration of the cytoplasm, inhibits the uptake of essential ions such as phosphorus. Apart from the direct consequences of water loss, there are a variety of metabolic effects. The cessation of photosynthesis and respiration leads to a lack of respirable material and the inability to translocate photosynthate, leading to eventual starvation injury, despite the presence of reserves. Dehydration also has several effects on cellular protein. Desiccation stress causes a rapid breakdown of proteins to amino acids, together with an equally rapid fall in protein synthesis (both of

which are reversible) due in part to a loss of energy and enzyme inactivation. Certain amino acids, notably proline, have been shown to accumulate in plants under desiccation stress (e.g. wheat, barley, tobacco), although some (e.g. sunflower, soybean) require a very severe desiccation stress before proline accumulation (reviewed by Levitt, 1980 and references therein). Part of the increase in proline is due to *de novo* synthesis and it is assumed to have some as yet unidentified role in drought resistance, possibly by association with hydrophobic side chains in proteins to give overall hydrophilicity and so maintain hydration of macromolecules (Schobert, 1977). As well as unspecific protein breakdown there is a general inactivation of enzymes due to protein loss and conformational changes; despite this, certain hydrolytic enzymes and oxidases may increase in activity upon desiccation, notably RNase.

Nucleic acids are also broken down upon prolonged dehydration; the RNA content (and thus protein synthesis) and the polyribosome levels are both shown to decrease, presumably due to the increased levels of RNase, and a fall in the synthesis of nucleic acids has also been demonstrated. This is in contrast to the nucleic acid response to slight dehydration, which has been reported to lead to an increase in the RNA content. A further factor that is influenced by desiccation is the hormone balance. Generally, wilting leads to a decrease in growth promoters and an increase in growth retardants, such as ABA, the latter response being rapid and

due in part to *de novo* synthesis (see section 1.2.4).

Many of these responses to dehydration are interrelated; some suggest an attempt by the plant to protect itself from the stress: the synthesis of ABA closing stomata and so reducing transpiration loss; the decrease in growth promoters reinforcing the cessation of cell growth; the breakdown of proteins to avoid damaged forms and produce free amino acids, some of which may have a protective role. However, on prolonged exposure to desiccation the sum of these injuries may lead to death of the plant tissues (although some younger, less damaged areas of the plant may survive). Accordingly, those plants whose habitats have a high incidence of water loss or drought have evolved a number of strategies to minimise the dangers of desiccation damage.

1.1.2 Adaptation Strategies

Plants can avoid serious damage from desiccation by drought avoidance, drought tolerance, or a combination of the two (reviewed by Levitt, 1980). Avoidance is a more complete adaptation to water stress as it allows growth and possibly development with a very low water supply; tolerance merely allows the plant to survive during the period of desiccation stress and resume growth upon rehydration. Drought avoiders maintain high internal water levels by a variety of morphological adaptations, including cuticle and root adaptations, decreased exposed leaf area, water storage organs and the use of alternative pathways, such as

crassulacean acid metabolism and C₄ photosynthesis. Drought tolerators survive decreased cellular water potentials; the mechanisms are not fully understood but are thought to include solute and reserve accumulation, an increased proportion of bound water and aggregation of protein molecules.

Such strategies are usually employed by plants which are likely to be exposed to desiccation stress or drought in their usual habitats (xerophytes) - most higher plants cannot survive a loss of 40-90% of their normal water content (mesophytes). However, the cellular consequences of prolonged dehydration, and more recent research into the expression of new proteins or messages on exposure to desiccation stress or ABA treatment, have indicated that most higher plants do contain some in-built resistance to water loss (discussed further in section 1.3). One of the obvious approaches to examine how plant tissue survives dehydration is to look at a part of the plant which is 'programmed' to survive extensive water loss - the seed, which first acts as actively dividing, then expanding and synthesising tissue, before undergoing severe desiccation stress while retaining its viability, then withstands a rapid influx of water to germinate and grow. It is reasonable to assume that any systems for protecting the plant cell from dehydration injury will be expressed in the seed at at least one stage of its development; examination of the regulation and function of such systems may aid understanding of dehydration resistance

in the whole plant and, eventually, have some economical significance in the improvement of crop plants.

1.2 SEED DEVELOPMENT

1.2.1 The Seed as a Biological System

Seeds of higher plants undergo a variety of metabolic stages during the transition from fertilised ovule to seedling. The seed has the ability to respond either to some external or internal stimulus, or as part of a developmental programme, to 'switch' metabolic and synthetic modes (Bewley, Kermode and Misra, 1989; Goldberg, Barker and Perez-Grau, 1989). These changes can occur at several stages throughout the life of the seed: from rapid cell division to the anabolic stage, where the cells expand and synthesise and accumulate protein, carbohydrate and lipid reserves for storage; through the maturation of the embryo, including synthesis of new polypeptides and the production of long-lived mRNA; to a period of developmental arrest (and in some cases dormancy), where the seed remains viable in a desiccated state; and then the synthesis of germination- and post-germination-specific messages and proteins and the catabolism of stored reserves, to allow rapid germination and growth of the young seedling (reviewed by Harada *et al*, 1988). These phases involve very different metabolic processes, requiring the ability to redirect enzymatic action and the control of gene expression to allow the seed to

develop, germinate and grow normally when under suitable conditions. Further, the seed, or its maternal environment, needs some mechanism of preventing germination and/or growth in conditions which are not conducive to seedling survival. Thus the different stages of seed development necessitate a range of synthetic and degradative activities combined not only with efficient control systems for changing the emphasis of seed metabolism but also with some method(s) of perception and response to environmental conditions. Seed development also encompasses some unique features, such as the faculty to withstand extreme desiccation, a capacity not usually found in other plant organs of mesophytes; this allows the seed to remain viable for long periods under conditions which would, in other tissues, cause severe and irreversible damage. Another specialised ability is the production and storage of long-lived mRNA (reviewed by Payne, 1976). This, when in conjunction with the rapid re-initiation of transcription and translation from a very low level (Misra and Bewley, 1985; Lalonde and Bewley, 1986) immediately after imbibition, allows germination to commence rapidly in response to water uptake (Thompson and Lane, 1980; reviewed by Harada *et al*, 1988).

It is clear that seeds provide a unique biological system with which to study many aspects of temporal and spatial gene control, protein synthesis, and the responses of plant tissues to stress and environmental change. To understand how the plant effects this complex development it

is necessary to examine not only the changes in nucleic acid and protein synthesis, but also their activity within the seed and their relation to the age of the seed, the environmental climate, and internal conditions, such as the levels of reserves, the hormone balance and the seed moisture content.

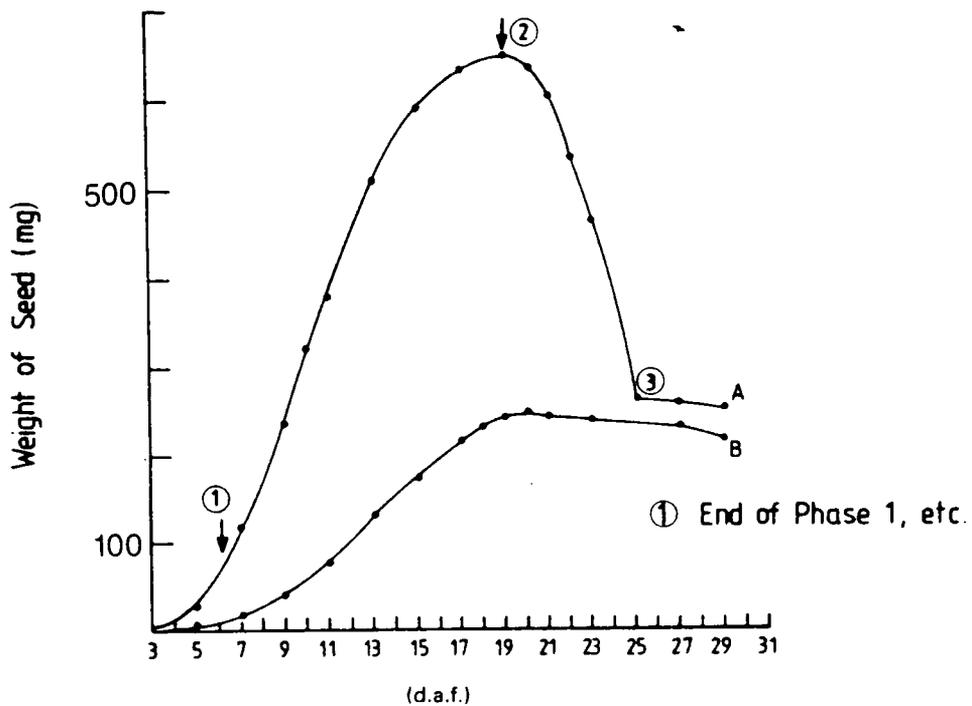
1.2.2 The Events from Fertilisation to Growth

The phases of seed development and the structure of seeds vary widely among higher plants. As well as the period of development and the size of the seed, the primary reserve material and the tissues used for storage also differ, some storing reserves in the endosperm, others with reduced endosperm and either one or two cotyledons as the major storage tissue. However, the general pattern of seed formation and growth is similar in a large number of species, although the actual timing of development will vary with the species or cultivar used. Development is usually measured as days after pollination (DAP) or days post anthesis (dpa); in plants such as pea, where fertilisation occurs within 24 hours of the full opening of the flower, development is measured in days after flowering (DAF).

After fertilisation of the ovule, the embryonic axis and surrounding cellular structures are formed by cellular division and differentiation until a suitable cell number is reached (reviewed by Boulter *et al*, 1990); during this period of morphogenesis, which in pea comprises less

than one third of the developmental time scale, the seed is small with a high water content. The seed then begins a filling stage, involving cell expansion. Here it increases in weight and seed diameter as the embryo develops and the storage tissues synthesise and accumulate nutrients (Le Deunff and Rachidian, 1988; reviewed by Boulter *et al*, 1990). In plants such as *Brassica napus* and the legumes the cotyledons are the primary reserve tissue, whereas the endosperm fills this role in others (e.g. cereals, castor bean, tobacco). The fresh and dry weights of the seed continue to increase until it reaches physiological maturity (see Fig.1), the end of development of the seed organs. The expansion and accumulation period takes approximately half of the term of development and at its end the seed is physiologically mature, having accumulated sufficient carbohydrates, storage proteins and (in some seeds) lipids, to sustain germination and growth until the new seedling can synthesise its own reserves. By this stage the seed has reached a decreased moisture content of between 55 and 60% (reviewed by Bewley, 1979 and Boulter *et al*, 1990). After maturation, the seed becomes dehydrated on the plant as the vascular connections between the pod and the mother plant break down (Le Deunff and Rachidian, 1988); the moisture content of the seed falls to a final level (between 5 and 15% in most species), accompanied by a decrease in metabolic activities as the seed reaches a state of quiescence (Boulter *et al*, 1990). Further growth occurs only after rehydration,

Figure 1. Changes in mean seed fresh weights (A) and dry weights (B) during development of pea cv. Feltham First, taken from Boulter *et al* (1990). 1, 2 and 3 indicate the end of the developmental phases of cell division and differentiation, seed filling and seed desiccation respectively, with 2 indicating the point of physiological maturity



where the seed imbibes water and undergoes cell expansion, division and elongation to produce a seedling.

1.2.3 Water Relations of the Seed

To function efficiently, the seed needs to be able to respond quickly to stimuli, both internal and external. This perception mechanism must be sensitive enough to allow a rapid response to environmental change and yet be able to withstand weaker stimuli caused by minor environmental fluctuations. In the latter respect, the mother plant and surrounding seed structures (pod, testa, pericarp etc.) may act as a buffer. As imbibition of water is the impetus for germination, and desiccation of the mother plant has been shown to accelerate maturation of developing seeds in some species (reviewed by Levitt, 1980), it is obvious that levels of water within and around the seed are an important consideration in the regulation of seed development. It has been demonstrated that most seeds undergo dehydration towards the end of the maturation period (reviewed by Bewley, 1979). It is therefore reasonable to assume that the loss or redistribution of water may play an important part in initiating or enhancing the production of nucleic acids and proteins to prepare the seed for extreme desiccation and, later, germination. Thus the loss of water could effect the transition from an assimilative phase to a mature, dry seed, capable of germination and seedling growth (Saab and Obendorf, 1989). Several groups have studied varying water

levels in seeds and how they relate to the ability of the seed or isolated embryo to germinate, both *in vivo* and *in vitro*.

1.2.3.1 Moisture Content of the Seed

Young immature seeds usually have a fairly high water content. The percentage of water in relation to seed weight gradually falls as the seed grows and assimilates nutrients; this is followed by a sharper decrease as the mature seed dehydrates. In soybean (*Glycine max*) the initially high water levels decrease continuously during growth, with a sharp decline at the point where the seed approaches its maximum dry weight (Saab and Obendorf, 1989). Finkelstein and Crouch (1986) found similar results with *Brassica napus*, the young seeds (before assimilation) containing from 80 to 90% water, which declines to 5 to 15% at maturity. An investigation of the seed moisture content of pea by Le Deunff and Rachidian (1988) divided the variation in water content into three phases. In phase one (P1), after fertilisation, the moisture content is constant at 85% (relative to fresh weight). This decreases during P2, as the seed weight increases, to 55% at physiological maturity, the same level as that found in soybean (Rosenberg and Rinne, 1986). The subsequent desiccation of the seed leads to a fall in both seed weight and moisture content, the dry seed containing 14 to 18% water. Le Deunff and Rachidian (1988) also carried out an examination of the relationship

between the seeds, the pod and the mother plant, using leakage conductivity to approximate the flow of sap from parent to seed. This demonstrated that the arrival of sap from the mother plant to the testa and apoplast decreases slowly during P2, and very rapidly after maturity. Around this stage, the pod (having reached its maximum weight) begins to senesce, as the vascular connections between pod and plant are disrupted. This engenders a water deficit, instigating the rapid dehydration seen in P3.

Thus, the seed first undergoes slow relative moisture loss as it accumulates macromolecules without taking in large amounts of water, so that the percentage water decreases by comparison; then loses water with increasing speed as the connections with the mother plant, the seed's source of water, break down and then sever. Consequently, the seed has a period of gradual water loss in which to prepare itself for more severe dehydration. It is conceivable that this loss could be used as a stimulus, either direct or indirect, for the initiation of protective measures.

1.2.3.2 Desiccation Tolerance in the Seed

It is vital that the seed protects itself against substantial loss of water, the presence of which is so necessary for normal cellular activity. In many seeds, the removal of pods, seeds or embryos for culture or germination *in vitro* has shown that there is a transition from

desiccation intolerance to desiccation tolerance during development. In most seeds this occurs midway through development - for example between 20 and 25DAP in castor bean (Kermode and Bewley, 1985); 20 and 26DAP in *Phaseolus vulgaris* (Misra and Bewley, 1985); 38 and 44DAF in soybean (Blackman *et al*, 1991); and around 35DAP in maize (Oishi and Bewley, 1992) - while pea seeds achieve desiccation tolerance gradually during the P2 seed filling phase (Le Deunff and Rachidian, 1988). The acquisition of desiccation tolerance in barley (*Hordeum vulgare*) occurs somewhat earlier, between 12 and 16DAP, and allows the embryo to remain viable after extreme water loss - 16DAP excised embryos retained the ability to germinate precociously upon rehydration after a desiccation treatment which removed 96-98% of the embryo's water (Bartels, Singh and Salamini, 1988). These workers also found that concomitant with the acquisition of desiccation tolerance was the appearance of a group of 25-30 proteins whose levels were noticeably enhanced, or newly expressed, between 10 and 16DAP. The mRNA population also altered between 12 and 16DAP, indicating some switch in expression. As this occurs fairly early in the development of the barley grain - between the stages of cell division and tissue differentiation - this change is not related to the synthesis of storage proteins, or to a response to desiccation, as the maturation drying occurs much later in the developmental period. Bartels and co-workers concluded that these proteins may be connected with desiccation

tolerance and are apparently developmentally programmed, as no desiccation cue occurs at this stage of development. Blackman *et al* (1991) also identified a set of proteins in soybean axes, expressed later in seed development during late embryogenesis, but whose timing and level of expression corresponded to the acquisition and loss of desiccation tolerance during seed development and germination. Their experiments showed that these 'maturation proteins' were induced by slow drying of immature excised soybean seeds, and also by the incubation of immature isolated seeds at high relative humidity; however, desiccation tolerance was only induced by the former treatment and Blackman *et al* concluded that the proteins were not sufficient for the induction of desiccation tolerance, although they may contribute to tolerance.

Desiccation tolerance is lost as the seed germinates and is not found in seedlings or plants of any of the species mentioned - in soybean, tolerance is lost only 18 hours after imbibition (Blackman *et al*, 1991), while other plants such as pea become intolerant to desiccation after the emergence of the radicle (Lalonde and Bewley, 1986). Thus it appears that desiccation tolerance is developmentally programmed and confined to the seed in mesophytes (Bartels, Singh and Salamini, 1988; Blackman *et al*, 1991). However, in xerophytes which exhibit desiccation tolerance in the whole plant the ability to withstand protoplasmic dehydration is retained. When Bartels *et al* (1990) examined several ABA-

inducible genes isolated from leaves and embryos of the resurrection plant *Craterostigma plantagineum* they identified several areas of sequence homology to ABA-responsive genes found in the embryos of mesophytes. They suggested that some of the same factors associated with desiccation tolerance in seeds could be expressed in the whole plant in xerophytes, with such specialised plants being able to override the developmental 'shut-off' or down-regulation seen in mesophytes to extend the seed's invulnerability to the whole plant.

1.2.3.3 Precocious Germination and Water Loss

It has been demonstrated in many species that, during *in planta* seed development, the seed is unable to germinate before maturation; however, once the immature seeds are removed from the maternal environment and placed under suitable conditions a small percentage of germination can be observed, the germination frequency increasing with seed age {for example pea (Le Deunff and Rachidian, 1988) and soybean (Rosenberg and Rinne, 1986)}. In other species, isolation of the embryo from the whole seed or kernel is necessary to give germination, as in castor bean (Kermode and Bewley, 1985), barley (Evans, Black and Chapman, 1975) and maize (Oishi and Bewley, 1992). In pea the germination of detached seeds during the P2 filling stage gave fairly weak seedling growth (Le Deunff and Rachidian, 1988); this was also noted by Rosenberg and Rinne (1986) with soybean seeds isolated at

35DAF (68-79% moisture content). However, most groups noted a significant increase in germination frequency and the attainment of normal seedling growth when immature seeds were exposed to a slow drying treatment before imbibition. This dehydration/rehydration treatment was only effective in inducing germination after desiccation tolerance had been achieved. Looking at the seedling growth of precociously germinated soybeans, Rosenberg and Rinne (1986) concluded that although the ability of the seed to germinate relied on the seed being a certain developmental age and being placed under the correct conditions, the initiation of cell division for seedling growth and the synthesis of certain polypeptides required the drying treatment, bringing the seed moisture content to below 60% and giving seed germination and growth equivalent to that seen in naturally matured seeds.

Kermode and Bewley (1985) showed that 100% germination of castor bean seeds removed from the capsule and testa and placed in water did not occur until the seeds were mature (60DAP); however, a dehydration/rehydration treatment applied after the acquisition of desiccation tolerance (20-25DAP) gave germination and normal seedling growth (Kermode and Bewley, 1988). Precocious germination of whole immature seeds after premature dehydration has also been seen in *Phaseolus vulgaris* (Dasgupta and Bewley, 1982) and maize (Oishi and Bewley, 1992). Kermode and Bewley (1988) concluded that a drying treatment applied after desiccation tolerance was achieved was an absolute requirement for

successful germination and seedling growth in castor bean and possibly other species. Bewley, Kermode and Misra (1989) also showed that a partial drying treatment giving only limited water loss in immature castor bean and *P.vulgaris* seeds gave the same high frequency of germination and seedling viability, indicating that the dehydration treatment did not have to be severe in these species. In contrast, Oishi and Bewley (1992) found that minimal drying of maize kernels detached from the plant and kept in high relative humidity did not lead to germination.

In the whole excised seed it would appear that desiccation is the cue for efficient germination and seedling growth. However, somewhat different results are seen when the immature embryos are removed from the whole seed environment. Precocious germination of immature embryos, non-dried and placed in water, has been shown in several species, including *P.vulgaris* (Dasgupta and Bewley, 1982); rape (Finkelstein and Crouch, 1986); castor bean (Kermode and Bewley, 1988); maize (Oishi and Bewley, 1992); wheat, where functional maturity of the embryo has been demonstrated at 15dpa (natural maturity at 60dpa) when removed from the seed environment (Morris *et al*, 1991); and barley, whose detached embryos can germinate precociously at only 8DAP, before the acquisition of desiccation tolerance (Bartels, Singh and Salamini, 1988). This last example demonstrates that desiccation tolerance is not vital for the ability to germinate, only for the resistance to dehydration treatments;

from the literature it appears that desiccation tolerance appears soon after the excised embryo is able to germinate. Cook *et al* (1988) showed that precocious germination of pea could be induced in isolated cultured embryos using a medium with low osmotic pressure, although the percentage of germination and seedling vigour were not as high as in naturally matured embryos. They also found that culture on a high osmoticum prevented germination and encouraged the same developmental processes seen *in vivo*, as had previously been observed in rape (Finkelstein and Crouch, 1986). It appears that isolation from certain seed tissues and placement in either water or low osmoticum allows precocious germination and seedling growth without any desiccation treatment. The less vigorous growth seen after the germination of isolated embryos could be ascribed to the lack of storage reserves; however, further investigations have suggested that desiccation and embryo isolation have different effects on the metabolism of the seed and the capability for vigorous seedling growth.

1.2.3.4 The Effects of Premature Desiccation and Embryo Isolation on the Seed Metabolism

Rosenberg and Rinne (1986) found that although immature soybean seeds isolated from the plant had the ability to germinate after 35DAP, the germination did not lead to appreciable seedling growth until the seeds were physiologically mature between 50 and 60DAP, when the

moisture content was below 60%. It was only at this later stage that seeds were able to synthesise malate synthase and isocitrate lyase (required for lipid utilisation) upon germination, and Rosenberg and Rinne detected a correlation between the production of these enzymes and the ability of the germinated seed to commence successful seedling growth. They further demonstrated that premature desiccation of 33DAP seeds not only gave levels of malate synthase and isocitrate lyase upon germination equivalent to those seen in germinated mature seeds and 100% normal seedling growth and vigour, but also gave a rapid loss of chlorophyll from the developing seed tissues during premature dehydration due to degradation, as seen in natural maturation. Later experiments (Rosenberg and Rinne, 1987) showed that the changing levels of starch, soluble sugars, proteins and oils during germination and seedling growth of prematurely matured seeds were comparable to those seen in naturally matured seeds; they were able to conclude that premature desiccation applied to desiccation-tolerant seeds induces the same metabolism seen after natural maturation drying and so could be used to examine the effects of desiccation on seed metabolism. Kermode and Bewley (1988) also investigated the production of post-germinative enzymes, when comparing the effects on normal castor bean seed metabolism of premature desiccation and embryo isolation. They found that upon germination both prematurely desiccated seeds and isolated embryos exhibited the same pattern of reserve mobilisation as that seen in naturally matured seeds,

but the synthesis of post-germinative enzymes LeuNAase (L-leucyl-beta-naphthylamidase) and isocitrate lyase was lower in isolated embryos than the usual high levels seen in naturally and artificially matured germinated seeds. A similar difference between the metabolism of isolated embryos, prematurely dried and naturally dried kernels of maize was seen by Oishi and Bewley (1992). They showed that although after premature desiccation 35DAP maize kernels exhibited a completely germinative and post-germinative pattern of protein synthesis upon rehydration, comparable to that seen in naturally matured germinated seeds, isolated embryos from 35DAP kernels continued to synthesise some developmental proteins as well as some of those associated with germination. From their results, Bewley and co-workers concluded that although embryo isolation gives the same immediate morphological response as premature and natural drying (i.e. germination) and allows the instigation of some post-germinative processes, desiccation is necessary to give a permanent switch between developmental, germinative and growth pathways. Further experiments (Bewley, Kermode and Misra, 1989) showed that this 'switch' in metabolic modes in *P. vulgaris* and castor bean occurs at the level of transcription, with desiccation leading to a halt in the synthesis of developmental protein mRNAs and the induction of germination- and growth-associated messages upon imbibition, accompanied by some post-transcriptional control, as the remaining developmental messages are specifically degraded

upon rehydration. Apart from these qualitative changes, Kermode, Pramanik and Bewley (1989) also found that the changes in quantitative levels of total RNA were comparable between naturally matured and prematurely dried castor bean seeds, both showing a fall in total RNA content upon desiccation and an increase upon imbibition, again by transcriptional and post-transcriptional controls; however, the final level of total RNA reached in prematurely desiccated germinated seeds was less than that found in naturally matured germinated seeds.

Rosenberg and Rinne (1986, 1988, 1989) also examined the synthesis of polypeptides which were accumulated in soybean seeds during both natural and premature maturation. They isolated several soluble, non-storage 'maturation polypeptides' from seeds which were capable of seedling growth, the polypeptides being absent from younger or undried seeds. They demonstrated *de novo* synthesis of both the polypeptides and their messages upon desiccation and found that the three most abundant polypeptides - 21, 31 and 128K - which accumulated gradually in maturing seeds were also synthesised during early germination (5-30HAI), their levels falling at the onset of seedling growth (30-72HAI). Rosenberg and Rinne (1989) suggested that the presence of these polypeptides could be associated with the ability of the germinated seed to initiate seedling growth; alternatively, such maturation polypeptides could be associated with the protection of the seed against

desiccation, as suggested by Blackman *et al* (1991).

These results supply several theories on the relationships between the seed's ability to germinate and grow, its internal water levels, the presence of the mother plant and the presence of surrounding seed tissues. Firstly, it seems that in some seeds (such as soybean and pea) the maternal plant tissues prevent precocious germination until development has produced a mature seed. In others (e.g. castor bean, barley, maize) the tissues surrounding the embryo fulfil the same role. This inhibition can be removed either by isolation of the seed/embryo or premature desiccation of the seed; however, only desiccation will effect the full transition from seed development through successful germination to seedling growth, with full expression of post-germinative enzymes and no further production of developmental proteins. This irreversible change can be induced by dehydration at any point after the acquisition of desiccation tolerance; further, it has been shown that premature desiccation of the seed can induce the synthesis of non-storage proteins normally associated with late embryogenesis which may have a role in the control of the seed's growth pattern, or in protection of the seed against the effects of desiccation.

1.2.4 Changes in Hormone Levels in the Seed

1.2.4.1 The Involvement of ABA in Plant Systems

It is generally accepted that phytohormones play a role in the mediation of gene expression and protein synthesis to allow the plant to adjust to its environment, although in many cases the mechanism of phytohormone action and the directness of its role in plant responses to stress are not fully understood (reviewed by Kuhlemeier, Green and Chua, 1987, Morris and Bowles, 1987 and Skriver and Mundy, 1990). The most widely researched phytohormone associated with seed development is abscisic acid (ABA). In many of the seed and plant systems reviewed ABA has been shown to be involved in plant responses to external stimuli and stresses, such as a lack of water, light, heat and anaerobic stress (reviewed by Kuhlemeier, Green and Chua, 1987). The increase in biosynthesis of ABA upon such stresses is induced by a loss of cell turgor and is thought to be important in adapting the plant tissues to survive water loss and other adverse effects by the inhibition of cell expansion and the expression of new proteins and messages (reviewed by Zeevaart and Creelman, 1988 and Skriver and Mundy, 1990). As well as a possible involvement in adaptation to stress, ABA has a variety of proposed roles within the seed. It has been demonstrated in a number of species (by using isolated immature seeds and embryos incubated with exogenous ABA) that ABA enhances the synthesis of developmental and storage

proteins, while suppressing germination- and growth-related proteins in both dried and non-dried embryos {e.g. barley (Bartels, Singh and Salamini, 1988); rape (Finkelstein and Crouch, 1986); wheat (Williamson and Quatrano, 1988); castor bean (Kermode, Dumbroff and Bewley, 1988); and maize (Oishi and Bewley, 1992); reviewed by Morris and Bowles, 1987}. In castor bean, germination is also suppressed in mature dry seeds incubated with a high concentration of ABA, although the developmental processes are absent (Kermode, Dumbroff and Bewley, 1989). Ackerson (1984) also showed that the endogenous levels of ABA controlled precocious germination, as immature excised soybean seeds did not achieve a high percentage of germination until the ABA content was less than 4ug/g fresh weight, a level usually seen *in vivo* only at physiological maturity at the beginning of desiccation (reviewed by Zeevaart and Creelman, 1988). Further proof of ABA's involvement in the prevention of precocious germination is the vivipary exhibited by ABA-deficient mutants of several plant species (reviewed by Zeevaart and Creelman, 1988 and Skriver and Mundy, 1990). Apart from the general continuance of developmental protein synthesis, ABA has also been implicated in the specific regulation of storage protein synthesis (reviewed by Zeevaart and Creelman, 1988, Morris and Bowles, 1987 and by Skriver and Mundy, 1990) and a range of other seed and plant messages (see section 1.2.4.3). Morris *et al* (1991) found that ABA was essential for the onset of seed dormancy in wheat, and a role in induction and

maintenance of dormancy has also been propounded in other plants (reviewed by Skriver and Mundy, 1990). ABA has also been implicated in the accumulation of proline upon desiccation stress, although the degree and mechanism of its involvement have not yet been elucidated (reviewed by Skriver and Mundy, 1990). A further proposed role for ABA was in the control of seed growth. However, experiments by Wang *et al* (1987) with *Pisum sativum* showed that the increase in ABA levels seemed to follow growth rather than cause it, while the use of pea ABA-deficient mutants by De Bruijn and Vreugdenhil (1992) confirmed that ABA had no significant effect on seed growth rate.

1.2.4.2 Changes in ABA Levels and Sensitivity During Natural and Premature Maturation.

In *Pisum sativum*, Ross and McWha (1990) found that over 90% of the plant's total ABA content was present in the developing seed, suggesting a high requirement for the phytohormone during development. In several species it has been shown that ABA levels rise in the embryo and seed tissues during development, then decrease fairly rapidly upon desiccation to low levels {e.g. pea (Wang *et al*, 1987; Ross and McWha, 1990); castor bean (Kermode, Dumbroff and Bewley, 1989); *Brassica napus* (Harada *et al*, 1988); wheat and barley (Morris *et al*, 1988)}. Ross and McWha (1990) and Wang *et al* (1987) both found that ABA levels in the pea embryo increased in parallel with the fresh weight of the seed; however, the

level in the testa was lower and peaked before that in the embryo, giving a biphasic pattern of ABA accumulation in the whole seed. Apart from the changes in absolute levels, several groups have recorded a fall in sensitivity to exogenous ABA upon maturation, as internal levels fall (reviewed by Zeevaart and Creelman, 1988); for example, Morris, Jewer and Bowles (1991) found that the concentrations of ABA which inhibited precocious germination in immature embryos of wheat and barley were not sufficient to prevent germination of more mature seeds. Kermode, Dumbroff and Bewley (1989) showed that premature desiccation could induce the rapid fall in ABA levels seen in naturally maturing seeds of castor bean, despite the usual response of vegetative tissue of increased levels of ABA on desiccation stress. Such embryos exhibited a high germinability; however, partially dried immature embryos also germinated well, while this treatment gave only a small decrease in the endogenous level of ABA. Kermode, Dumbroff and Bewley concluded that this lack of agreement between germinability and endogenous ABA levels could be explained by sensitivity, with natural, premature and partial drying all leading to a decline in sensitivity of tissues to ABA. This conclusion was supported by the fact that natural, premature and partial drying all decreased the sensitivity of isolated embryos to exogenous ABA by 10-fold. Benech Arnold, Fenner and Edwards (1991) showed that a greater desiccation stress, achieved by growing plants of *Sorghum bicolor* under drought conditions, gave

very high levels of ABA during early development with a sharp decrease at maturation; however, such stressed plants had a much higher incidence of precocious germination during early maturation than the control plants. Again, they were able to explain the lack of correlation between ABA levels and the inhibition of precocious germination by the fact that the sensitivity of isolated embryos grown under drought stress to exogenous ABA was only one tenth of that of the controls. Kermode, Dumbroff and Bewley (1989) suggested that this desiccation-induced fall in embryo sensitivity to ABA may play an important part in the 'switch' from development to germination and growth, removing ABA's inhibition of precocious germination and promotion of developmental protein synthesis. Oishi and Bewley (1992) proposed that, *in vivo*, ABA could be the main factor preventing precocious germination in developing seeds before maturation, while after maturation the negative water potential of the seed will inhibit germination (Finkelstein and Crouch, 1986; Morris, Jewer and Bowles, 1991) as both ABA levels and embryo sensitivity fall during desiccation. It has been suggested that this alteration in sensitivity to hormones may be mediated by drying (Kermode, Dumbroff and Bewley, 1989). A change in sensitivity to hormones upon desiccation has been demonstrated in other plants as a factor of seed development - for example, either natural or premature desiccation is necessary to enable alpha-amylase genes in cereal grains to be responsive to gibberellic acid (GA) (Evans, Black and

Chapman, 1975). As GA induces the expression of a range of genes which function in reserve mobilisation and growth the use of desiccation to induce (GA) or reduce (ABA) sensitivity to hormones may provide a mechanism for the 'switch' between developmental and germinative modes.

1.2.4.3 The Effect of ABA on Message and Protein Levels

A large number of ABA-inducible or *rab* (responsive to ABA) genes have been isolated from both seeds and whole plants (reviewed by Zeevaart and Creelman, 1988 and Skriver and Mundy, 1990). Most of these can be divided into two groups - seed storage proteins and LEA (late embryogenesis abundant) proteins or dehydrins.

The use of exogenous ABA on cultured embryos showed that storage proteins and their messages from rape, soybean and wheat can be induced by ABA, while incubation of excised soybean cotyledons with fluoridone, an inhibitor of ABA biosynthesis, inhibited the accumulation of the storage protein beta-conglycinin (reviewed by Zeevaart and Creelman, 1988). Another phytohormone, jasmonic acid, has also been implicated in the induction of storage proteins; the results of Mason and Mullet (1990) suggested that jasmonic acid may mediate the induction of vegetative storage proteins by water stress in soybean. However, Finkelstein and Crouch (1986) showed that the presence of high levels of endogenous ABA was unnecessary for controlled storage protein synthesis in cultured rape embryos. The use of ABA-deficient mutants

showed that the accumulation of storage proteins in tomato and *Arabidopsis* did not depend on the presence of endogenous ABA (reviewed by Zeevaart and Creelman, 1988). These findings, combined with the fact that the induction of storage proteins such as napin and cruciferin is normally slow and not to very high levels, and that the expression of storage proteins does not necessarily follow the pattern of ABA accumulation, have led reviewers to conclude that ABA is not a major factor in storage protein regulation (Zeevaart and Creelman, 1988; Skriver and Mundy, 1990). Induction of storage protein gene expression by exogenous ABA may merely be a feature of the general mechanism of precocious germination inhibition and the continuation of development. *Rab* genes and proteins have been isolated from a wide variety of plants (reviewed by Skriver and Mundy, 1990 and Espelund *et al*, 1992). Of those that are found in the seed the majority are LEAs (late embryogenesis abundant proteins), whose expression increases at the same time as the peak of endogenous ABA is reached at the end of the seed or grain filling period. Many of these have been shown to be induced in immature seeds by the application of exogenous ABA {e.g. in maize (Gomez *et al*, 1988); cotton (Baker, Steel and Dure, 1988); wheat (Morris *et al*, 1988); rape (Harada *et al*, 1989); and barley (Hong, Barg and Ho, 1992)}. Barratt and Clark (1991) also isolated two ABA-responsive LEA proteins from pea which showed significant homology to a pea disease resistance protein. Several can also be induced in seedlings or the

whole plant by the application of exogenous ABA or desiccation stress and are often classified as 'dehydrins' {e.g. barley and maize dehydrins (Close, Kortt and Chandler, 1989); rice (Mundy and Chua, 1988); tomato (Cohen and Bray, 1992); and pea dehydrins (Robertson and Chandler, 1992)}. However, the ability to induce these normally seed-specific genes in the whole plant by stress is not universal: an aleurone- and embryo-specific LEA gene from barley is inducible in all organs of 3 day old seedlings by ABA or drought stress, but this inducibility decreases rapidly as the seedlings age, suggesting developmental regulation (Hong, Barg and Ho, 1992); while an ABA-responsive LEA isolated from carrot seeds cannot be induced in the whole plant after germination (Goupil *et al*, 1992). In most of these cases ABA induction is linked with osmotic stress and many groups have suggested that it is the stress that is the primary factor in gene expression and regulation, ABA being either a part of the response or a side effect (Finkelstein and Crouch, 1986; Morris *et al*, 1990). However, many LEA and dehydrin genes have been shown to contain ABA-responsive elements (Marcotte, Bayley and Quatrano, 1988; Mundy and Chua, 1988; reviewed by Skriver and Mundy, 1990; see section 1.3). Moreover, treatment of desiccation-intolerant calli of *Craterostigma plantagineum* (which contains several genes which are responsive to ABA and desiccation) with exogenous ABA can induce desiccation tolerance, presumably by the induction of such genes (Bartels *et al*, 1990). Skriver and Mundy (1990)

suggested that some of the *rab* genes form part of a general response to osmotic stress, which is developmentally regulated in maturing seeds but may also be induced in earlier stages of seed development or vegetative tissues by ABA or osmotic stress. Perhaps connected with this is the observation by some groups that some *rab* proteins may bind nucleic acids and act as regulatory proteins (reviewed by Skriver and Mundy, 1990). The fact that the physiological response to ABA can be either rapid (e.g. LEA message induction in stressed seedlings) or slow with long term effects (e.g. storage proteins) indicates that ABA affects gene expression by different mechanisms, and possibly using different receptors (Zeevaart and Creelman, 1988).

1.2.5 Changes in mRNA and Protein Levels in the Developing and Germinating Seed

1.2.5.1 Changes in mRNA Mass and Complexity during Development

In general, most researchers have found that mRNA levels increase to a peak during mid to late embryogenesis, usually corresponding to the peak of storage protein synthesis, then fall to a low level in the mature dry seed (Misra and Bewley, 1985 with *Phaseolus vulgaris*; Ishibashi and Minamikawa, 1989 with cowpea; reviewed by Harada *et al*, 1988). The same pattern was seen with total RNA in castor bean endosperm (Kermode, Pramanik and Bewley, 1989) and

P. vulgaris (on a per seed axis basis, Misra and Bewley, 1985), although Ishibashi and Minamikawa (1989) did not see such a large fall in total RNA content in cowpea and Misra and Bewley (1985) recorded an increase in total RNA on the basis of fresh weight. The translational activity also varies with seed age, increasing to a maximum at mid-development, again corresponding to the peak of major seed protein synthesis, then falling to a lower level in the dehydrated seed (Misra and Bewley, 1985; Ishibashi and Minamikawa, 1989). The decline in mRNA, translational activity and (in some cases) total RNA corresponds to the decreased rates of RNA and protein synthesis and metabolism and a loss of polyribosomes during seed desiccation (Misra and Bewley, 1985; Rosenberg and Rinne, 1986; reviewed by Goldberg, Barker and Perez-Grau, 1989 and Harada *et al*, 1988). Upon imbibition, however, the rates of RNA and protein synthesis and respiration increase rapidly (Kermode, Pramanik and Bewley, 1989; reviewed by Harada *et al*, 1988); the rate of translation also increases after imbibition, although in cowpea there is a lag period of 12HAI (Ishibashi and Minamikawa, 1989). Several groups have attempted to discover if the changes in mRNA content and activity reflect changes in the whole mass of the mRNA or specific regulation of certain genes at different developmental stages.

In an extensive study of the hybridisation kinetics of an excess of seed mRNA from soybean cv. Dare to homologous cDNA, Goldberg *et al* (1981) showed that in terms of diversity

most of the 14,000 to 18,000 different embryonic mRNA species were present throughout development and in the mature dry seed. However, the frequency distributions of different groups of species - classified loosely as rare, abundant and superabundant - changed substantially during seed growth. During cell division (30DAF) the mRNA population consists mainly of abundant (800 molecules/cell/sequence) and rare (17m/c/s) messages, with superabundant sequences forming only 4% of the mass. However, at mid-maturation (75DAF) this class of 7 to 10 superabundant messages (>19,000m/c/s) increased dramatically to form 50% of the mRNA mass, the rest being comprised of fairly abundant (550m/c/s) and rare (3m/c/s) messages, the former present as a greater proportion of the mid-maturation mass than at the earlier stage, and the latter forming only 9% of the mass.

The superabundant messages are embryo specific, encoded by low frequency repetitive DNA sequences (presumed to be small multigene families) and are thought to code for the abundant seed proteins synthesised at this developmental stage, such as seed storage proteins, seed lectins and trypsin inhibitors. Although present in the dry seed they seem to be at lower concentrations than expected, suggesting that the superabundant mRNA mass falls fairly rapidly before or during desiccation. Visualisation of the superabundant mRNAs on gel showed, as expected, no separate bands during cell division but 7-10 bands, forming 50% of the total mRNA mass, at mid-maturation. The pattern changed slightly during

late maturation; several of the mid-maturation bands decreased in intensity or were not visible, while at least one of the remaining 5-6 bands was not seen at the mid-maturation stage. Analysis of mRNA from axes and cotyledons indicated that the majority of superabundant messages were found in the cotyledons, again supporting their relationship to storage proteins. No separate bands were seen in post-germination mRNA, confirming the fall in superabundant message concentration after mid-maturation. Most of the superabundant messages were not found in leaf polysomes; however, the majority of the other diverse embryonic mRNA species were present in leaf polysomes, although at very low concentrations.

Similar results on the complexity and abundance of mRNA species were obtained from cotton and wheat, as reviewed by Goldberg, Barker and Perez-Grau (1989) who concluded that the majority of different mRNA species found in the embryo are present throughout development and in the dry seed and mature plant. However, the results of Morton et al (1983) from *Pisum sativum* cv Feltham First, while confirming the rise in very abundant messages during maturation in parallel with the increase in major seed protein synthesis (storage proteins legumin, vicillin and convicillin; possibly pea lectin and the major albumin), suggested that the number of diverse species fell during development from ca.20,000 to ca.200, due to the loss of rare ($<10^3$ copies/cell) sequences whose levels fell from 9DAF (end of cell division) and were

absent by 19DAF (towards the end of storage protein synthesis). It is hard to correlate these results with the previous findings; it is possible that pea has a different system of regulation to soybean, but this seems unlikely as they are in the same family and the soybean pattern of persistence of diversity is maintained in the less closely related families of cotton and wheat. The hybridisation experiments were not taken to completion so the figures may not be completely accurate; further, the fact that Morton and co-workers used cotyledons rather than the whole seed may bias the results - as Goldberg *et al* (1981) showed that the majority of the superabundant messages were located in the cotyledons, the prevalence of such messages in pea cotyledons may make the presence of rare messages hard to detect.

1.2.5.2 Patterns of mRNA and Protein Accumulation

Several groups have demonstrated that the prevalent mRNA population changes quite substantially during development and that these alterations can be related to the stage of development, the synthesis and accumulation of seed proteins and the preparation of the seed for germination and growth (reviewed by Goldberg, Barker and Perez-Grau, 1989). During cell division the mRNA population consists mainly of messages for proteins associated with early embryogenesis, including factors for growth and differentiation of the seed structures, along with the constitutively expressed 'housekeeping' genes; these include the components of the

energy-generating systems and protein synthesis, some of which form part of the fraction Goldberg *et al* (1981) defined as 'abundant'; for example the ribulose bisphosphate carboxylase small subunit and ribosomal proteins (reviewed by Gatehouse *et al*, 1986). Towards mid-maturation the 'superabundant' messages become prevalent, at a time of high translational activity and the accumulation of large amounts of storage proteins. The message levels decline during late maturation and are stored in the dry seed, some at very low levels (Kermode, Pramanik and Bewley, 1989; reviewed by Goldberg, Barker and Perez-Grau, 1989). Storage protein and seed lectin synthesis may continue after the message levels fall but this also decreases during late maturation (Ishibashi and Minamikawa, 1989). As these superabundant messages become less prevalent a third dominant population of LEA messages appears, accompanied by the synthesis of LEA or 'maturation' proteins, many of which are thought to be necessary for seed survival and seedling growth (Rosenberg and Rinne, 1988; Kermode, Pramanik and Bewley, 1989; reviewed by Goldberg, Barker and Perez-Grau, 1989). The LEA messages persist during drying, some showing an increase in abundance upon dehydration, and are stored in the dry seed at fairly high levels, along with their proteins (see section 1.3).

Underlying these major population changes there is a steady expression of the constitutive messages and of embryo-specific transcripts, including embryonic enzymes such as cotyledonary alpha-amylase. In addition there is a less

prevalent population of messages which appears around mid-maturation, is stored in the seed and continues to be expressed in the germinating seed and growing seedling; the expression of this last group has been of great interest, as information on the timing of induction and the regulation of messages which are involved in germination may aid an understanding of the 'switch' from developmental to germination and growth processes (reviewed by Goldberg, Barker and Perez-Grau, 1989).

Of the messages stored in the dry seed, most are degraded upon imbibition* (Cuming, 1984; Misra and Bewley, 1985; Lalonde and Bewley, 1986; Kermode, Pramanik and Bewley, 1989; reviewed by Goldberg, Barker and Perez-Grau, 1989). However, there is a further subset of messages - 'long-lived mRNA' - the majority of which appear in the seed during mid-

[*FOOTNOTE - The results of Goldberg *et al* (1981) suggest that for the majority of messages, except perhaps the seed-specific superabundant fraction, this degradation and initiation of messages is not absolute; however, the continuation of synthesis is at such a low level; as demonstrated by their results from leaf polysomes, that it can be disregarded with reference to message populations which are important at this stage of development and growth, except to support the theory of gene enhancement and repression for regulation rather than that of an on/off switch mechanism]

maturation and increase in transcript concentration upon germination (reviewed by Payne, 1976 and by Goldberg, Barker and Perez-Grau, 1989; Thompson and Lane, 1980; Lalonde and Bewley, 1986). This stored mRNA supports protein synthesis during early germination, using the stored ribosomal proteins and tRNA which quickly reassemble upon imbibition (reviewed by Payne, 1976), until the germinating seed resumes active transcription. According to Thompson and Lane (1980) this can occur as early as 40 minutes after imbibition, with new mRNA becoming responsible for the majority of translation within 24HAI (Ishibashi and Minamikawa, 1989). The messages transcribed during germination include those necessary for germination that were stored in the seed, whose levels increase early in germination then gradually decline in prevalence as the seed finishes germinating, and a new set of post-germination messages for mobilisation of reserves and growth, giving a very different pattern of seed proteins to that in the seed during development (Thompson and Lane, 1980; Misra and Bewley, 1985; Lalonde and Bewley, 1986; Williamson and Quatrano, 1988). Payne (1976) suggested that long-lived mRNA species were transcribed during grain filling. Ishibashi and Minamikawa (1989) showed that in cowpea all of the requirements for germination and successful seedling growth were present in the seed at the point of physiological maturity at 17 to 19DAF (cowpeas, unlike many seeds, do not require a dehydration period to give successful germination and growth, although maturation drying does occur as part of

the developmental plan); the mRNA population at this developmental stage showed no difference to the mature dry seed, but a considerable difference to the prevalent mRNA population of mid- to late maturation (13-15DAF), indicating that some change in expression occurred to reduce the messages for seed storage proteins and increase those for germination and post-germinative processes - the increase need not be great, as the mass of poly(A)⁺RNA decreases noticeably between 13-15 and 17-19DAF. Harada, Baden and Comai (1988) and Harada *et al* (1988) conducted a detailed survey of the mRNAs found during germination and post-germination in rape. They found that although most of the post-germination abundant mRNAs were not detectable in developing or dry seeds there was a set of messages which was also found in immature and dry seeds and showed spatial as well as temporal regulation. As the cotyledons and axes have very different physiology it was proposed that different sets of genes may be expressed in the two tissues. Harada and co-workers found that although many post-germination abundant messages were present in both tissues, some accumulated preferentially in the cotyledons and others in the axis. Two cloned axis-abundant genes showed different patterns of accumulation to each other; however, the cotyledon-abundant genes appeared to be regulated coordinately. The messages of cotyledon-abundant genes accumulated in late maturation (35-40DAF) but only to fairly low levels, due either to some transcriptional or post-transcriptional control, or to the

general fall in RNA and protein synthesis during maturation drying. Their transcript level then increased within 8-16HAI, reaching maximal expression within 1DAI. Two members of this set were identified as isocitrate lyase and malate synthase, both of which show the same pattern of accumulation and tissue distribution as each other and the rest of the cotyledon-abundant set. In contrast, the axis-abundant messages accumulated in young embryos, falling to minimal levels at 35-40DAF and not increasing again until imbibition. This pattern of expression may suggest an involvement of such messages with division and differentiation, both of which occur during early embryogenesis, are absent during the rest of seed development, then occur again in germination and growth, the messages presumably being down-regulated when not needed but maintained at low levels to enable the rapid resumption of division upon imbibition. The accumulation of messages necessary for germination in the axes during early embryogeny may explain the ability of very young barley and wheat embryos to germinate when isolated from the whole seed and allowed to imbibe (Bartels, Singh and Salamini, 1988). Although germination- and post-germination specific messages have been isolated from embryonic polysomes in developing seeds, the proteins do not accumulate and no germination-related events occur until after imbibition (Harada *et al*, 1988). Thus the results indicate that post-germination genes have more than one mechanism of regulation which coordinates the expression of both genes and proteins with respect to the

seed's developmental stage and the function of the tissue. Harada, Baden and Comai (1988) suggested that a 'cascade' of events is required for germination and is initiated only on imbibition of a suitably developed seed.

1.2.5.3 The Effect of Desiccation on mRNA and Protein Populations

As stated before, the exposure of developing seeds to premature desiccation causes accelerated maturation with a coincidental change in the seed mRNA population to give a pattern very similar to that seen in naturally matured seeds (Kermode, Pramanik and Bewley, 1989). As it has been established that the messages for germination and post-germinative growth accumulate upon maturation and desiccation and also upon premature desiccation, it has been suggested that desiccation may act as a cue for the preparation of the seed for germination (Dasgupta and Bewley, 1982; Kermode, Pramanik and Bewley, 1989; Rosenberg and Rinne, 1986, 1988; Saab and Obendorf, 1989; reviewed by Bewley, 1979). However, it has been noted that changes in mRNA and protein synthesis in prematurely dried seeds occur after rehydration, and the imbibition of water following germination may also be important in the 'switch' from development to germination (Bewley, Kermode and Misra, 1989); it is at least clear that one or both of these factors is of greater importance than the developmental age of the seed as imposed desiccation automatically alters the natural developmental programme of

immature seeds to the germinative program usually only seen later in ontogeny (providing desiccation tolerance has been achieved). Lalonde and Bewley (1986) examined the effect of desiccation then rehydration during the germination of pea axes. The treatment enhanced the degradation of some developmental-specific mRNAs upon rehydration, implying a control by desiccation/rehydration of specific degradation mechanisms. However, another set of residual mRNAs, usually lost on imbibition, was stimulated by desiccation and rehydration; this group could possibly represent the stress proteins which are thought to be expressed during late embryogenesis for protection against desiccation (see section 1.3), or for repair of damaged proteins and cellular structures. The set of messages whose level usually increases upon germination continues to increase after the desiccation/rehydration treatment but the germination process itself takes longer, possibly due to the need to repair damage caused by the treatment. Further, Lalonde and Bewley isolated at least one message after the treatment which was not detectable in the mature dry seed or in early germination.

1.2.5.4 The Regulation of mRNA and Protein Levels during Seed Development and Germination

The regulation of gene expression in the seed is tightly controlled and conserved; however, it can also change in response to environmental stimuli such as water loss or

low sulphur - the control mechanisms therefore need to be able to adapt, both to internal developmentally programmed changes and to external factors, and still produce a viable embryo (reviewed by Gatehouse *et al*, 1986 and Goldberg, Barker and Perez-Grau, 1989). Most major seed-specific genes (e.g. storage protein genes, lectins, protease inhibitors) show a general increase in expression during development to a peak, then fall; however this pattern of expression varies considerably among the proteins, starting at different points in development, having varying lengths of peak expression and declining at different stages - for example, the expression of the pea vicilin storage protein starts before that of the legumin (reviewed by Gatehouse *et al*, 1986). Apart from their varied temporal regulation, most prominent seed proteins appear to exhibit seed specificity, being 'switched' on and off at the beginning and end of seed development. This, however, has been called into doubt by a number of studies. Goldberg *et al* (1981) demonstrated that although the superabundant messages were not detected in soybean leaf polysomes, the majority of other messages expressed in the seed were; despite their presence they were expressed at very low levels - lower than the leaf's own rare messages - implying severe 'down' regulation, but it appears from these results that very few genes were actually 'shut off' after embryogenesis. Boulter *et al* (1990), looking at the expression of major pea seed protein genes in run-off transcription assays from leaf nuclei, found that although

Leg A messages could not be detected, it seemed that *Leg J*, *S* and pea lectin were expressed at very low levels. Walling *et al* (1986) found similar results in soybean; the messages for storage proteins beta-conglycinin, glycinin and a 15K protein were shown to be present in leaf nuclei, but transcription rates were 50-100 fold less than those in embryo nuclei. Further, the steady state levels of these sequences in leaf nuclear RNA were over 10,000 fold lower than in embryonic nuclei, implying that the down-regulation of 'seed-specific' genes occurs at both transcriptional and post-transcriptional levels. Perhaps the most convincing proof for the expression of usually seed-specific genes in non-seed tissues is the ability of many plants to form somatic embryos, with morphological and genetic events equivalent to those in zygotic embryos, including the expression of seed storage protein genes, despite the absence of a seed environment (reviewed by Goldberg, Barker and Perez-Grau, 1989). These results indicate that although certain seed-specific genes may not be expressed in non-seed tissues they are at least present in a state where transcription can be induced, while others are transcribed and accumulate at very low levels but can be 'up-regulated' if necessary. A further feature of the regulation of seed genes is the spatial regulation reviewed by Harada *et al* (1988) and Harada, Baden and Comai (1988), indicating that specific genes or gene sets accumulate to different levels in tissues with different physiological functions and that this accumulation is controlled in varied

fashions. Further, as well as spatial regulation with respect to organs, there is also some evidence that messages can accumulate to different levels within specific cells (reviewed by Goldberg, Barker and Perez-Grau, 1989).

Many groups have concluded that the major control of protein expression during seed development, germination and growth is at the level of transcription (reviewed by Gatehouse *et al*, 1986, Goldberg, Barker and Perez-Grau, 1989 and Boulter *et al*, 1990). However, post-transcriptional control in the form of processing, transport, transcript stability and specific degradation also plays an important part in the response to maturation drying; after desiccation and imbibition any residual developmental mRNAs are selectively degraded (Kermode, Pramanik and Bewley, 1989; reviewed by Goldberg, Barker and Perez-Grau, 1989). Post-transcriptional control also appears to be important during development: Morton *et al* (1983) found that of the mRNA transcripts present in the nuclei of pea seeds during early development (9DAF) only 50% were represented in the cytoplasm; Gatehouse *et al* (1986) found that although the levels of seed protein mRNAs and the amount of protein synthesis both increased during early to mid-development they did not do so in parallel, the rate of increase of protein synthesis being less than that of transcription; in soybean, Walling *et al* (1986) showed that although seed-specific protein and non-seed protein genes had similar transcription rates, the accumulation of the two types of mRNA varied by up

to 10,000 fold. Post-transcriptional regulation has also been cited as the level at which seeds respond to environmental conditions; Beach *et al* (1985) found that peas on low sulphur did not accumulate legumin (a sulphur-containing protein), although the rate of transcription was not altered. There is also some indication of translational or post-translational control in the prevention of accumulation of germination-specific and growth-related proteins (despite the presence of their messages on cotyledonary polyribosomes) during development and in the post-translational phosphorylation of ABA-inducible maturation polypeptides in maize (Harada *et al*, 1988; Goday *et al*, 1988).

The isolation of conserved sequences which are specific to gene families in the promoters of seed protein genes suggest the presence of elements for transcriptional control (reviewed by Gatehouse *et al*, 1986). Several groups have demonstrated that 5' flanking regions of seed protein genes contain all of the information necessary to direct not only seed-specific expression but also temporal regulation (see section 1.4.1; reviewed by Goldberg, Barker and Perez-Grau, 1989). Further work has isolated putative *trans*-acting factors - DNA binding proteins - which may associate with these *cis* regulatory elements (Meakin and Gatehouse, 1990; reviewed by Goldberg, Barker and Perez-Grau, 1989). Such results, along with the isolation of ABA-responsive elements, support the view that transcriptional control is a major part

of the differential expression of genes during development and growth and is probably involved in effecting the switch from developmental to germination and growth related processes. The timing and mechanism of this switch have not yet been elucidated: as mentioned before, the ability of many seeds to germinate and grow precociously, with or without a desiccation treatment, indicates that the required factors for germination and growth are present at least by the point of physiological maturity (Dasgupta and Bewley, 1982; Kermode and Bewley, 1985; Rosenberg and Rinne, 1988; Ishibashi and Minamikawa, 1989; Oishi and Bewley, 1992); but in most cases the switch does not seem to be actually effected until after rehydration following a desiccation treatment (Comai and Harada, 1990; Bewley, Kermode and Misra, 1989;). The results obtained from isolated, non-dried embryos further suggest that there may be two switches - one to allow germination to proceed, which may or may not give complete inhibition of developmental processes and a second to activate the post-germinative programme to give enzyme synthesis and seedling growth (Rosenberg and Rinne, 1986; Kermode and Bewley, 1988). The results so far indicate that desiccation, rather than developmental age, is the primary inducer of the full change in direction of seed metabolism and morphology. Bewley, Kermode and Misra (1989) suggested that the desiccation stimulus could be effected by a minimal drying to a critical water content, rather than the extensive dehydration usually seen during seed maturation; however the demonstration by

Oishi and Bewley (1992) that this minimal drying treatment is ineffective in maize kernels suggests that different species may require different levels of water loss.

1.3 LEA PROTEINS AND DEHYDRINS

A set of messages and proteins which are expressed primarily from mid-maturation through late embryogenesis has been isolated from a wide range of both monocotyledons and dicotyledons. The polypeptides remain in the dry seed and are degraded upon imbibition; they have been termed variably 'maturation polypeptides' (Rosenberg and Rinne, 1986; Bewley, Kermode and Misra, 1989; Kermode, Pramanik and Bewley, 1989; Blackman *et al*, 1991) and more recently LEA, *rab* or dehydrin proteins. The last two have been more commonly isolated from stressed seedling tissue, but were subsequently shown to exhibit late embryogenesis abundance in seeds and to share significant sequence homology with LEA genes isolated from seeds. Proteins exhibiting this expression pattern in seeds are thought to occur in all plants (reviewed by Dure *et al*, 1989, Goldberg, Barker and Perez-Grau, 1989 and Skriver and Mundy, 1990). As a group, LEA and dehydrin proteins are heat-stable, strongly hydrophilic and have been shown to be synthesised in response to ABA and osmotic stress, applied to the immature or germinating seed; seedlings or (less frequently) adult plants. Cotton LEA genes are also transiently expressed in immature seeds upon excision, before

premature germination begins (Baker, Steele and Dure, 1988). The rapid response of the genes to such external stimuli, together with the induction of certain LEAs by cold, salt and heat suggests that they function as stress proteins (Morris *et al*, 1990; Hong, Barg and Ho, 1992). The response to osmotic stress, their hydrophilicity and the structure of their polypeptides have led to the general conclusion that most of the LEA proteins and dehydrins are involved in the protection of plant tissues during dehydration, functioning primarily in the seed and either absent or present at lower levels in the mature plant, but inducible in seedlings and in some cases vegetative tissues on osmotic stress (Gomez *et al*, 1988; Harada *et al*, 1989; Dure *et al*, 1989; Blackman *et al*, 1991). LEAs and dehydrins have been isolated from cotton (Baker, Steele and Dure, 1988); rice (Mundy and Chua, 1988); wheat (Williamson, Quatrano and Cuming, 1985); barley (dehydrins, Close, Kortt and Chandler, 1989; Hong, Barg and Ho, 1992; Espelund *et al*, 1992); maize (dehydrin, Close, Kortt and Chandler, 1989; Gomez *et al*, 1988); rape (Harada *et al*, 1989); radish (Raynal *et al*, 1990); *Craterostigma plantagineum* (Bartels *et al*, 1990); tomato (Bray, 1991; Cohen and Bray, 1992); *Arabidopsis thaliana* (dehydrin, Rouse, Gehring and Parish, 1992); carrot (Goupil *et al*, 1992 and see Dure *et al*, 1989) and pea (dehydrin, Robertson and Chandler, 1992; Barratt and Clark, 1991). In 1989, Dure *et al* divided several of the LEAs that had been characterised at that time into three groups on the basis of amino acid sequence

homology or, in the cases of groups 2 and 3, on the basis of possession of areas of conserved homology.

1.3.1 Group 1 LEAs

To the original two members of this group - the wheat Em message (Williamson, Quatrano and Cuming, 1985) and the cotton D19 (Baker, Steel and Dure, 1988) can be added the other Em genes from maize and rice (see section 1.4.2); three messages from barley - B19.1, B19.3, B19.4 (Espelund *et al*, 1992), cotton D132, maize Emb564, carrot EMB1 and radish p8B6 (reviewed in Espelund *et al*, 1992). All of these share very strong amino acid homology, including the characteristic hydrophilicity and a lack of cysteine (Cys) and tryptophan (Trp) residues. The three barley messages are classified by the number of repeats of a 20 amino acid motif that they contain (B19.4 having four repeats, B19.3 three, etc.); this motif is the most hydrophilic region of the polypeptides and is found twice in cotton D132, but is only present as a monomer in cotton D19 and the other members of this group. Like most LEAs, the B19 messages are encoded by a multigene family; however, the three members show differential expression (Espelund *et al*, 1992). They share the same temporal expression during normal seed development, but B19.1 accumulates to levels at least 10-fold greater than B19.4 and B19.3 messages (the latter being the least abundant) and this pattern is also seen in the dry seed. In the immature embryo and germinating seeds all three messages

are responsive to ABA or mannitol, but to different levels; further, while the level of B19.1 mRNA is noticeably increased by treatment of immature embryos with 200mM NaCl, the B19.3 message is unaffected and the B19.4 transcript shows only a slight rise. This response to salt is not seen in germinating seeds and so appears to be developmentally dependent. B19.1 mRNA also seems to accumulate to higher levels upon cold stress, but it is thought that the effect of the cold stress is mediated by stabilisation of the transcripts and not by enhanced transcription of messages, as seen in certain other LEAs (e.g. rice RAB21, group 2) (see Espelund *et al*, 1992). The response to salt and cold can therefore be separated from that to mannitol and ABA, despite the fact that ABA has often been proposed as the mediator of responses to such stresses. Unlike many LEAs, the B19 family are not expressed in non-embryonic tissue, being inducible only in zygotic and somatic embryos and young germinating seedlings; this organ-specificity has also been described in similar genes in barley and cotton (see Espelund *et al*, 1992).

1.3.2 Group 2 LEAs

As group 3, this set of LEA and dehydrin messages exhibit considerably less amino acid homology to each other than the more closely related group 1 messages, instead sharing certain conserved sequence motifs. Dure *et al* (1989)

first placed in this group the cotton LEA D11 (Baker, Steele and Dure, 1988) and the rice RAB21 (Mundy and Chua, 1988); it can be extended to contain the barley and maize dehydrins (Close, Kortt and Chandler, 1989), an *Arabidopsis* dehydrin (Rouse, Gehring and Parish, 1992), pea dehydrin B12 (Robertson and Chandler, 1992), a dehydrin from *Craterostigma* (Bartels *et al* 1990), seed LEAs from radish (Raynal *et al*, 1990), tomato LE4 (Bray, 1991) and maize RAB17 (see Raynal *et al*, 1990). All have a high glycine content, no Trp residues and conserved sequence motifs which include a lysine-rich block and a stretch of from 7-9 serine residues, the only exception being pea dehydrin B12 which has no serine run. All have been shown to be inducible by ABA and water stress in seedlings (e.g. Close, Kortt and Chandler, 1989; Robertson and Chandler, 1992), immature seeds (e.g. D11, Baker, Steele and Dure, 1988), leaves (e.g. Raynal *et al*, 1990), shoots and roots (Bray, 1991). Interestingly, in water-stressed seedlings B12 pea dehydrin protein only accumulated in cotyledons and shoots, even though the message is available in roots; Robertson and Chandler (1992) suggested that this may be purely an effect of the rapid cessation of protein synthesis in roots upon desiccation.

1.3.3 Group 3 LEAs

Members of this group include cotton LEA D7 (Baker, Steele and Dure, 1988), rape pLEA76 (Harada *et al*, 1989), barley pHVA1 and carrot DC3 and DC8 (see Dure *et al*, 1989).

Again the deduced proteins, which cover a range of sizes, are highly hydrophilic; but in this case they contain a varying number of conserved 11 amino acid repeats which may allow the formation of an amphipathic alpha helix (Dure *et al*, 1989). All are Cys and Trp free (except carrot DC3 which has one Cys residue). The accumulation of rape pLEA76 mRNA in seeds occurs around the time of peak ABA levels and can also be induced by ABA in immature embryos and by water stress in seedlings (Harada *et al*, 1989). The carrot LEA DC8 is also inducible by ABA in the immature seeds but cannot be induced in seedlings or vegetative tissue, this strict seed specificity being retained in transgenic tobacco (Goupil *et al*, 1992). Induction of the barley HVA1 message and protein by drought stress in seedlings also decreases as the seedlings age, suggesting developmental regulation; however, another smaller protein showing some similarity to HVA1 is induced in seedlings by ABA but not in seeds, suggesting differential developmental regulation of either a gene family or similar genes (Hong, Barg and Ho, 1992). HVA1 is also induced by salt, cold or heat stress.

1.3.4 Other Unclassified LEAs

Cotton LEA D113 (Baker, Steele and Dure, 1989) and tomato LE25 share 50% amino acid homology and areas of homology in their 5' flanking region; these include a proposed ABA-responsive element (5'-TACGTGGC-3') that has

been found in several other ABA-responsive genes, and a novel 11bp sequence (5'-TACTGAAAAAC-3'), which are thought to be involved in regulation (Cohen and Bray, 1992). Again, the proteins are very hydrophilic with no Cys or Trp residues. Baker, Steele and Dure (1989) proposed that the frequency of glycine and hydroxylated residues in D113 (also found in D11, group 2) favoured the existence of the polypeptide as a random coil, similar to that suggested for the Em protein by Dure *et al* (1989). Both D113 and LE25 are expressed in the seed and inducible by exogenous ABA and drought stress in seedlings, although LE25 message is expressed at much higher levels during maturation than in drought stressed shoots and roots (Bray, 1991).

Other LEAs which show no homology to group 1, 2 and 3 LEAs are the two ABA-responsive albumins isolated from pea by Barratt and Clark (1991) - ABR17 and ABR18. These bear no relation in sequence or protein structure to other identified LEAs, although they are expressed in fairly high levels in the seed during late embryogenesis and their levels can be increased in young seeds by incubation on low osmotica or with ABA (Barratt, Domoney and Wang, 1989a). Instead, they appear to be related to a set of proteins involved in disease reponse in pea and to a birch pollen allergen. ABR18 expression follows the peaks of ABA expression, first in the testa and then in the embryo and although the protein usually disappears on germination it can be induced in germinating seedlings by exogenous ABA. In contrast ABR17 was absent

from the testa, showed typical late embryogenesis abundance in the embryo and was present throughout germination and in non-stressed leaves (Barratt and Clark, 1991). This differential expression of closely related proteins could imply differential developmental regulation of a gene family, as in barley (Hong, Barg and Ho, 1992).

1.3.5 Induction and Regulation of LEAs

In most cases, LEA and dehydrin polypeptide levels parallel the mRNA levels fairly closely, suggesting that the protein levels are regulated at the level of message accumulation rather than at the level of translation (Baker, Steel and Dure, 1988; Close, Kortt and Chandler, 1989). There is some control at the level of the rate of transcription (Williamson and Quatrano, 1988; Comai and Harada, 1990); several conserved elements have been isolated in the 5' flanking regions of LEA genes (e.g. Cohen and Bray, 1992; Espelund *et al*, 1992) and Oeda, Salinas and Chua (1991) isolated a nuclear factor found in rice and tobacco that shows specific binding to a motif I sequence (5'-GTACGTGGCG-3') found in rice RAB and cotton LEA genes. Post-transcriptional regulation also seems to play a part, especially after ABA treatment; stabilisation of transcripts has been indicated as part of the regulatory mechanism in response to ABA, desiccation and cold stress (Williamson and Quatrano, 1988; Marcotte, Russell and Quatrano, 1989; Espelund *et al*, 1992). There are many examples of

differential regulation of LEAs in the same plant. Baker, Steele and Dure (1988) demonstrated different temporal accumulation of two sets of cotton LEAs. Different spatial accumulation has been noted, for example of tomato LE4, LE16 and LE25 in seedlings (Bray, 1991) and pea ABR17 and ABR18 in the seed tissues (Barratt and Clark, 1991). The ABR17 and ABR18 proteins also exhibit differential developmental control (Barratt and Clark, 1991), a feature seen in the expression of several other LEA proteins {e.g. the HVA1 message and protein in barley (Hong, Barg and Ho, 1992) and carrot DC8 (Goupil *et al*, 1992)}. Accumulation of messages and proteins to different levels in normal and treated (with environmental stress or ABA) seeds has also been documented, for example barley B19 (Espelund *et al*, 1992) and rice RAB21 (Mundy and Chua, 1988). The basis of this differential expression is not yet understood, although both transcription and stabilisation of transcripts may be involved (Espelund *et al*, 1992). Seed specificity seems to vary, being absolute in some LEAs and absent in others; Goupil *et al* (1992) suggested that the specificity to embryonic tissues exhibited by carrot DC8 was due to the absence in vegetative nuclei of a certain embryo nuclear binding factor which associates with part of the DC8 promoter, indicating that transcriptional control through transacting factors may play a major role in seed specificity.

The role of ABA in the induction of LEAs and dehydrins is confusing. It has been suggested that ABA is

involved in the response of the LEAs to dehydration. This has been supported in some cases with the use of ABA-deficient mutants; for example, tomato LE4, LE16 and LE25 are responsive to drought stress in wild type tomato but not in the tomato mutant *flacca*, which does not accumulate ABA (Bray, 1991). ABA-responsive elements have been isolated in several LEA genes (Williamson and Quatrano, 1988; Mundy and Chua, 1988; reviewed by Skriver and Mundy, 1990), along with putative transacting factors which show affinity to these elements (see Espelund *et al*, 1992), suggesting a definite role for ABA in at least part of the expression of LEA genes. ABA could merely account for the activation of such genes, later expression being controlled by other factors, as in most plants LEAs begin to accumulate at or just after the peak of embryonic ABA is reached (Goldberg, Barker and Perez-Grau, 1989). Sensitivity to ABA, the balance with antagonistic GA levels and the presence of ABA metabolites such as phaseic acid may also be involved in the gene response to ABA (Kermode, Dumbroff and Bewley, 1989; Benech Arnold, Fenner and Edwards, 1991; Robertson and Chandler, 1992). ABA may also merely act as a general stress transducer, as with the putative disease response proteins ABR17 and ABR18 (Barratt and Clark, 1991).

1.3.6 Proposed Roles of LEAs

The majority of LEAs isolated to date are thought to be involved in the protection of plant cells during

dehydration, by association with membranes or proteins in the cytosol (Harada *et al*, 1989; Dure *et al*, 1989; reviewed by Skriver and Mundy, 1990). Baker, Steele and Dure (1988) suggested that their high solubility and high concentration during late embryogenesis, together with their non-compartmentalisation, made a structural role unlikely but allowed functions such as the 'solvation' of cytosolic structures and possible association with highly charged proteins by salt bridges. Those LEA proteins with virtually unbroken hydrophilicity are unlikely to be able to span membranes, although they may stabilise certain membrane structures; however, those polypeptides which can exist as a random coil (e.g. Em, D11, D113 etc.) could intercalate with macromolecules to maintain hydration of the cytosol in protein-water-protein matrices, protecting against both water loss and the potential damage caused by rehydration (Dure *et al*, 1989; Morris *et al*, 1990; Robertson and Chandler, 1992). It has been proposed that the conserved sequences found in many LEAs may be important in their ability to act in dehydration protection of proteins and cellular structures, notably those which allow formation of amphipathic alpha helices and further structures (Dure *et al*, 1989; reviewed by Skriver and Mundy, 1990). Further, a recent review by Dure (1993) proposed that the those LEAs in Group 3, containing tandem repeats of an 11 amino acid sequence, may act as 'ion carriers', sequestering ions to prevent precipitation and/or crystallisation in the high salt conditions produced by

desiccation; such a function may explain the relatively high levels of these proteins during late embryogenesis and the conservation of amino acid sequence between species. LEAs may also be involved in other stress responses to factors such as salt, cold, heat and disease (Barratt and Clark, 1991; Hong, Barg and Ho, 1992), while Morris *et al* (1990) suggested that, due to their expression pattern in seeds and relationship to ABA, some may be involved in the maintenance of development and the inhibition of germination.

1.4 GENES AND PROTEINS USED IN EXPRESSION STUDIES

1.4.1 Legumin Storage Protein

Seed storage proteins are synthesised during embryo development and accumulate to form the most abundant form of seed protein, comprising up to 80% of the total protein (Boulter *et al*, 1990). They are important not only economically, controlling the nutritional value and content of seeds, but also as a model for the study of gene expression, by virtue of their seed-specific and developmentally-regulated accumulation. One of the proteins which has been investigated extensively is legumin, an 11S protein that is one of the major storage proteins of pea and is related to a range of legumin-type proteins found in other legumes and some non-legumes (Gatehouse *et al*, 1988). Legumin is a heterogeneous hexameric protein composed of 6

disulphide-bonded subunit pairs, of which there are several species, their arrangement within the final protein giving its heterogeneous nature. Each subunit pair contains an acidic (alpha, MW ca.40,000) and a basic (beta, MW ca.20,000) polypeptide which are specifically linked, being synthesised at the endoplasmic reticulum (ER) as a precursor of ca.60,000 MW, then transported through the Golgi apparatus to vacuolar protein bodies where the disulphide bond between alpha and beta subunits is formed before separation of the polypeptides (reviewed by Croy and Gatehouse, 1985). The subunit pairs have been divided into two types, 'major' and 'minor', depending on their relative abundance in the legumin protein (Casey, 1979). The 'major' class, L4, are thought to be coded for by a small, highly homologous gene sub-family of 5 genes, *A-E*, several of which have been characterised and sequenced (reviewed by Croy and Gatehouse, 1985); these genes have been mapped to a single locus, *Lg-1* on chromosome 7 (Domoney, Ellis and Davies, 1986). The 'minor' subunit pairs, designated L1-L3 and L5 by Matta *et al* (1981), can be divided into at least two classes: L1-L3 (minor 'big') are thought to be the products of the *Leg J* subfamily, *J-L* (Gatehouse *et al*, 1988, Thompson *et al*, 1991) and map to *Lg-2* on chromosome 1 (Domoney, Ellis and Davies, 1986); other 'minor' subunit polypeptides map to a different locus and may be the product of more than one gene subfamily (March *et al*, 1988, Thompson, 1989).

1.4.1.1 The *Leg A* Subfamily

Five highly homologous members of the *Leg A* subfamily, designated A to E, have been isolated and sequenced {Lycett *et al*, 1984 (*Leg A*); Lycett *et al*, 1985 (*Leg B* and *C*); Bown *et al*, 1985 (*Leg D*); Rerie *et al*, 1990 (*Leg E*)}. All are present in the same area of the chromosome, four occurring as a direct tandem repeat (Casey, Domoney and Ellis, 1986). Transcripts have been detected for all of the genes except *Leg D*; the sequence of this gene, which is situated ca.1.3kb 3' of the *Leg A* gene in the same orientation, shows that *Leg D* is a pseudogene (Bown *et al*, 1985). The complete sequence of the *Leg A* gene isolated by Lycett *et al* (1984) included the 5' flanking sequence, coding sequence for a signal peptide, a 36.44K alpha polypeptide, a 20.19K beta polypeptide and a 3' flanking sequence with three alternative polyadenylation sites. Lycett *et al* (1985) showed that *Leg A*, *B* and *C* have highly homologous sequences in the immediate 5' flanking sequences and 100% homology in the early coding sequence until 300bp from the start site, the complete homology between *B* and *C* continuing for another 550bp. *Leg E* has similar conservation of sequence in the 5' flanking and coding region (Rerie *et al*, 1990). All of the members of this subfamily have three introns found in the same positions and showing partial homology. Less homology is found between the 3' flanking sequences of the genes, but all have multiple polyadenylation sites (Lycett *et al*, 1984; Lycett *et al*, 1985; Bown *et al*, 1985; Rerie *et al*, 1990).

The major legumin polypeptide is fairly rich in sulphur, containing five cysteine and four methionine residues per polypeptide.

1.4.1.2 The *Leg J* Subfamily

Of the three proposed members of the *Leg J* subfamily, only *J* and *K* have been fully characterised and sequenced. Gatehouse *et al* (1988) isolated a genomic clone of *Leg J* and part of *Leg K* which demonstrated that the genes were arranged in tandem in the same orientation, separated by 6.5kb. The *Leg J* gene contained 616bp 5' flanking sequence, 611bp 3' flanking sequence and a coding region of 1742bp containing two introns. The predicted polypeptide of 503 amino acids exhibits the typical legumin subunit characteristics: a leader sequence (22 amino acids); an acidic alpha polypeptide (300 amino acids) containing a Cys residue at position 87 for formation of the disulphide bond, a hydrophilic and glutamate-rich C terminus and an Asp residue at the C terminal end, the probable point of proteolytic cleavage; a basic beta polypeptide (181 amino acids) with a Cys residue at position 7. The coding sequence of *Leg J* is 97% homologous to that of *Leg K*, confirming them as members of the same subfamily. This homology extends for ca.250bp of the 3' flanking region after the stop codon, including at least four potential polyadenylation signals, before homology is lost. Full characterisation of the *Leg K* gene from cv. Feltham First (Thompson *et al*, 1991)

demonstrated that there was also significant homology between *Leg J* and *Leg K* in the area of 5' flanking sequence isolated (up to ca.-560), including conserved sequence elements such as the Leg box (Gatehouse *et al*, 1986). The full sequence of *Leg K* from cv. Feltham First showed that the gene was translationally ineffective, due to a mutation of the ATG start codon to GTG (valine), and the *Leg K* message did not accumulate to detectable amounts (Thompson *et al*, 1991). However, Domoney and co-workers have cloned a functional *Leg K* cDNA from *Pisum sativum* cv. Birte, indicating that *Leg K* is expressed in other pea lines (Domoney and Casey, 1987).

Although the products of *Leg A* and *Leg J* gene subfamilies react with the same anti-legumin antibodies, their transcripts do not cross-hybridise and they have an overall homology of only 48% in the coding region, although this homology is stronger in some areas than others (Gatehouse *et al*, 1988). The introns of *Leg J* and the second two introns of *Leg A* show only a weak homology, as do the 3' flanking sequences; however, significant homology is seen in the 5' untranslated and flanking region. As the lack of homology between the two subfamilies indicates an early evolutionary divergence, with a much later divergence of genes within subfamilies, the strong conservation of the 5' flanking region suggests a functional constraint (Gatehouse *et al*, 1988). The *Leg J* gene also exhibits significant homology to the *Vfa LeB4* legumin gene from *Vicia faba*, from -450 in the 5' flanking sequence to 265bp 3' of the *Leg J*

stop site, with approximately 90% homology within exons 1 and 3 of the coding sequence (Gatehouse *et al*, 1988). The second exon codes for the C terminal region of the alpha polypeptide, which is the most variable portion of the sequence between legumin genes and shows only ca.75% homology between pea *Leg J* and *Vfa LeB4*. *Leg J, K* and *Vfa LeB4* also show 60% homology to two B-type (A- and B-type being analagous to 'major' and 'minor' respectively) legumin genes from soybean (Gatehouse *et al*, 1988).

1.4.1.3 Expression of Legumin Genes and Proteins

The general pattern of legumin synthesis is typical of seed storage proteins (Gatehouse *et al*, 1986): deposition of the protein is observed from early cotyledon expansion (9 DAF), with abundant synthesis occurring at mid-development and continuing until the legumin fraction forms 40-50% of the total protein (reviewed by Croy and Gatehouse, 1985, Boulter *et al*, 1990). However, within the legumin pattern of accumulation, the expression of different gene products may vary (Domoney and Casey, 1987, Gatehouse *et al*, 1986). Domoney and Casey (1987), using *Pisum sativum* cv. Birte and dot blot analysis, showed that the level of a message homologous to the *Leg A* gene (pCD43) increased rapidly between 15 and 19DAF, was comparatively constant until 22DAF, then increased again to a maximum at 24DAF before decreasing, showing a 13-fold increase in message level between 15 and 24DAF. The results from Northern analysis confirmed the rise

to around 18DAF and a slight increase at 24DAF. These results contrast with those of Thompson *et al* (1989), using cv. Feltham First, who demonstrated a peak of *Leg A* expression, as measured by dot blot analysis, at 18DAF and a peak at 18-20DAF by Northern analysis, with no peak at 24DAF. These workers noted a 35-fold rise in *Leg A* message from early cotyledon expansion (8-9DAF) to the peak at 18DAF, when the transcript formed 0.3% of the total RNA; however, they also found large differences in message levels between different batches of the same line (e.g. 30-40% difference in the level of *Leg A* transcript at 18DAF between two batches) which makes quantification of results unreliable. Further experiments by Thompson (1989) and Boulter *et al* (1990) looking at steady state levels of mRNA also found the peak of *Leg A* expression at 18DAF looking at seeds aged from 8-22DAF, although a second batch of 18-28DAF seeds showed a second possible peak at 22DAF. It was suggested that the disagreement between these results came from the use of different pea lines and/or environmental growth conditions.

Using a probe of a *Leg K* cDNA (pCD40), Domoney and Casey (1987) noted an increase in message level to a peak at 19DAF, after which the level fell slightly before a second possible peak at 24DAF, another slight fall and the level was then maintained until the last reading at 29DAF. These dot blot results were contradicted by Northern analysis where the message level rose from 14DAF to around 18DAF but appeared to be significantly lower than 14DAF at 24DAF. The increase

from 15-19DAF by dot blot analysis was only 2.5-fold, much less than that seen with *Leg A*. Thompson *et al* (1989), looking at *Leg J* expression, also noted that the level of 'minor' legumin message expression was much less than that of 'major'. Their dot blots showed an increase of expression from a low level at early cotyledon expansion (8-9DAF) to a peak at 16DAF; the level then fell slightly before a second peak at 22DAF, then decreased as the seed continued to desiccate. Measurement of the steady state levels of the *Leg J* subfamily again showed peaks at 16 and 22DAF, followed by a decline in message level during later desiccation; further dot blots with a different batch of peas gave peaks at 18 and 24DAF but the same pattern of expression, the actual timing possibly varying between batches (Thompson, 1989). Northern blots gave a more detailed picture of *Leg J* subfamily expression in cv. Feltham First, the probe pCD40 hybridising to two bands of ca.1860 and 1660b which increased equally to a peak at 16DAF; the upper band seemed to maintain its level until 20DAF then decreased slowly, while the lower band continued to increase in intensity to a peak at 24DAF, after which the level fell (Thompson *et al*, 1989). Thompson *et al* (1989) suggested that this biphasic pattern of expression may be due to the combination of different gene expression patterns from different members of the *Leg J* subfamily. The pCD40 probe used by Domoney and Casey (1987), Thompson (1989) and Thompson *et al* (1989) contains 0.91kb of the coding and 3' flanking region of *Leg K* and so should be homologous to

all three members of the subfamily, *J*, *K* and *L*. (Domoney, Ellis and Davies, 1986). Thompson *et al* (1991) investigated this further by treating the RNA samples (12-28DAF) with RNase H to reduce poly(A) tails and decrease the variation in transcript length. Hybridisation to pCD40 showed three clear bands and one faint band; densitometric analysis divided these into two groups by their expression patterns. The low mobility bands (ca.68 and 70mm) increased equally to a peak at 16DAF (late cotyledon expansion); the higher mobility bands (ca.72 and 74mm) also maintained an equal ratio but reached their maximum level at 24DAF, during desiccation, the level having fallen by 28DAF. Members of the subfamily were identified using oligos of *Leg J* and *Leg K* coding regions from an area of sequence divergence and a second *Leg J* oligo of the 3' flanking sequence (+1988 to +2010). The *Leg J* coding oligo hybridised only to the high mobility bands, 72 and 74mm, while the *Leg K* oligo showed no hybridisation above background, explained by its low steady state level in this line (Thompson *et al*, 1989). They concluded that the transcripts whose levels increased during late development and desiccation were from *Leg J*, and assumed that the lower mobility bands, which showed a more typical storage protein gene expression pattern of a rise to late cotyledon expansion then a fall at maturation, corresponded to the as yet uncharacterised *Leg L*, on the basis of cross-hybridisation to the pCD40 probe. This would explain the apparent biphasic accumulation of the *Leg J* message, the two peaks representing

peaks in message level of transcripts from different members of the *Leg J* subfamily. The second *Leg J* oligo from the 3' flanking region hybridised only to the 72mm band; from this it was concluded that the messages in the shorter 74mm band used the first polyadenylation site at 1869-1874, while the messages with 72mm mobility used the last site at 2015-2020. This use of alternative polyadenylation sites does not seem to be related to the developmental control of *Leg J* expression, as the ratio between the two bands remained constant as the seed aged.

Domoney and Casey (1987) and Thompson *et al* (1989) also examined the expression of another 'minor' legumin gene, termed *Leg S*. Domoney and Casey, using cv. Birte, showed a peak of expression at 19DAF, giving a 4-fold increase from 15DAF, followed by an irregular expression pattern with smaller apparent peaks at 24 and 27DAF, then an obvious decrease to 29DAF. Again, the level of expression of this message was markedly less than that of *Leg A*, although it appeared to be greater than that of the *Leg J* subfamily, even though it is thought to be coded for by only one gene. Thompson *et al* (1989), using cv. Feltham First, also found the expression level of *Leg S* to be less than *Leg A* and slightly greater than the *Leg J* subfamily, but showed a peak of expression at 16DAF followed by a fall in message level encompassing a second possible peak at 24DAF; this pattern of expression suggested by dot blot analysis was, however, markedly different to that seen by Northern analysis, which

showed the *Leg S* message level increasing from 12DAF to a maximum at 14, 16 and 18DAF, then falling rapidly. In essence, it showed the same expression pattern as *Leg A* but at an earlier stage, with a shorter period of maximum expression ending at late cotyledon expansion.

1.4.1.4 Control of Legumin Gene Expression

Several groups have demonstrated that control of legumin expression occurs at both transcriptional and post-transcriptional, and possibly translational, levels (Gatehouse *et al*, 1986; Thompson *et al*, 1989; Boulter *et al*, 1990). Domoney *et al* (1980) detected legumin message at very low levels very early in seed development, before any noticeable increase in transcript level. A measurement of transcription rates showed that from 8-12DAF both *Leg A* and *Leg J* transcription rates increased rapidly, suggesting that the 'up-regulation' of legumin gene expression occurred at the level of transcription (Boulter *et al*, 1990); however, between 12 and 16DAF the transcription rates of *Leg A* continued to rise while those of *Leg J* and *S* remained constant, the transcription rates showing little agreement with the steady state mRNA levels and so suggesting post-transcriptional regulation (Thompson *et al*, 1989). Post-transcriptional regulation was also noted in response to environmental conditions by Beach *et al* (1985) who observed a fall in steady state levels of legumin mRNA in plants exposed to sulphur deficiency, despite the maintenance of the normal

transcription rate. Thus it appears that the steady state level of legumin messages is controlled mainly by post-transcriptional processes (Thompson, 1989). Transcriptional control, however, is thought to be responsible for the timing of legumin message accumulation and for its tissue specific expression (Thompson *et al*, 1989). It is not clear whether the organ specificity of legumin expression is due to an 'on/off' developmental switch (Gatehouse *et al*, 1986) or to greatly enhanced transcription combined with inhibition of a post-transcriptional control which could prevent message accumulation in non-seed tissue, as suggested by Thompson *et al* (1989).

Several seed storage protein genes contain sequences in their 5' flanking regions which are specific to and conserved within gene families and it has been suggested that these sequences are involved in gene regulation (Gatehouse *et al*, 1986; reviewed by Goldberg, Barker and Perez-Grau, 1989). Croy *et al* (1988) showed that a *Leg A* construct containing 1239bp 5' flanking sequence, the full coding region and 633bp 3' flanking region was sufficient to give not only seed specific expression of legumin message in transformed tobacco but also to ensure accurate protein synthesis at the ER, transport to protein bodies and processing to the mature protein as in pea seeds. The amount of legumin synthesis in such transgenic tobacco plants, as a percentage of total seed protein, was much lower than in pea, suggesting that some other sequence or transacting factor was

necessary to attain the full *Leg A* expression seen in pea seeds (Ellis *et al*, 1988). Using 5' deletion mutants of the *Leg A* gene transferred into tobacco, Shirsat *et al* (1989) found that, as has been noted in other species, a short promoter sequence of 97bp including the CAAT and TATA boxes did not direct legumin gene expression. They showed that sequences between -97 and -549 gave legumin expression, including seed specificity and temporal regulation, while sequences between -549 and -1203 contained enhancer-like elements which increased levels of expression significantly. The sequence between -97 and -549 contains the Leg box - a 28bp sequence highly conserved among 11S legumin-type proteins from a number of species including pea, *Glycine max*, *Vicia faba* and *Helianthus* (Gatehouse *et al*, 1986) - and it was assumed, due to its strong conservation, that this sequence was necessary for gene expression, or at least had a role in the regulation of gene expression. However, later experiments in transgenic tobacco showed that 124bp of the *Leg A* 5' flanking sequence (to the transcription start) including the Leg box did not give gene expression and, further, that no nuclear proteins could be seen to bind to this area (Shirsat, Meakin and Gatehouse, 1990; Meakin and Gatehouse, 1991). Baumlein *et al* (1991) conducted similar experiments with the legumin gene *Vfa LeB4* from *Vicia faba* group B (analogous to 'minor') in transgenic tobacco and showed that while 1kb of the 5' flanking sequence to the transcriptional start gave a high level of expression,

deletion of the Leg box had no effect on gene activity. Like Shirsat *et al* (1989), they also found indications of the presence of enhancer elements in this 1kb fragment, as a deletion mutant of only 200bp of the 5' flanking sequence gave a fall in gene expression to less than 10% of that seen with the 1kb promoter fragment. Bogue *et al* (1990) also proved the capability of a 2.4kb fragment from the 5' flanking region of the major sunflower 11S seed protein helianthenin, which is structurally similar to other legumin-type proteins, to direct seed specific and temporally-regulated expression in transgenic tobacco.

Further examination of the association of the *Leg A* promoter with proteins has isolated two areas, -316 to -549 and -582 to -833 which show strong binding to seed nuclear proteins, but not to leaf nuclear proteins; one putative transcription factor, LABF1, was thought to interact with both of these regions (Meakin and Gatehouse, 1991). The activity of LABF1 is specific both to the seed and to the developmental period when *Leg A* mRNA synthesis is increasing, although *Leg A* transcription does occur when LABF1 appears to be absent or inactive; Meakin and Gatehouse (1991) suggest from these results that LABF1 acts as an enhancer of transcription.

1.4.2 The Wheat Em Late Embryogenesis Abundant Protein

The Em (early-methionine labelled) polypeptide was first characterised as an acid-soluble protein which is

abundant in dry wheat embryos and synthesised at the beginning of germination (0-3HAI), after which the level of both protein and message decreased rapidly (Thompson and Lane, 1980; Grzelczak *et al*, 1982). Later experiments showed that the Em message is detectable at very low levels in immature embryos during early embryogenesis (around 14dpa), then increases significantly after ca.17dpa to give high sustained levels during late embryogenesis and in the mature seed (at 60dpa) (Williamson, Quatrano and Cuming, 1985; Morris *et al*, 1990); the message then becomes undetectable after three days of germination (Cuming, 1984; Williamson, Quatrano and Cuming, 1985). The levels of Em protein parallel those of the message, the protein beginning to accumulate by 21dpa (Marcotte, Bayley and Quatrano, 1988), being fairly abundant during late embryogenesis and the dry seed, but is undetectable in immature embryos and germinating seeds after 12 hours of imbibition (Cuming, 1984; Williamson, Quatrano and Cuming, 1985). The mature message is approximately 780bp in length (Williamson, Quatrano and Cuming, 1985) and is coded by a gene family of about 10 members which exhibit considerable homology to each other (Litts *et al*, 1987). The deduced polypeptide from the 282bp coding region has a relative molecular weight of ca.9.9K, is highly hydrophilic and is very rich in glycine and both basic and acidic residues (Litts *et al*, 1987). The high glycine content, the abundance of the protein in dry embryos and its apparent embryogenesis specificity led to suggestions that Em

functioned as a major storage protein; however, Thompson and Lane (1980) discovered that it is found not in protein bodies (as the majority of storage proteins are) but in the cytosol, and, further, it appears to have no signal peptide (Cuming, 1984; Litts *et al*, 1987). McCubbin, Kay and Lane (1985) suggested that the Em polypeptide may act as a cryptobiotic protein to protect the wheat embryo during desiccation. Their hydrodynamic and optical data indicate that the polypeptide is present in the cytoplasm as a random coil whose hydrophilic residues allow it both to bind water, forming a protein-water-protein matrix to maintain a certain level of hydration during desiccation, and also to intercalate between macromolecules conferring protection against the damage caused by water loss (Morris *et al*, 1990). Although Em may still function as a storage protein in these circumstances, Marcotte, Bayley and Quatrano (1988) have shown that it is not essential for seedling production, as immature embryos will germinate precociously into normal seedlings when removed from the grain, without any accumulation of Em.

Several workers have demonstrated enhancement of the Em message and protein levels by ABA (Williamson, Quatrano and Cuming, 1985; Marcotte, Bayley and Quatrano, 1988; Williamson and Quatrano, 1988). Their results suggest that regulation by ABA occurs at the level of both transcription and specific message stability (Williamson, Quatrano and Cuming, 1985; Williamson and Quatrano, 1988);

the correlation between message levels and protein synthesis indicates that translational control is unlikely, but Williamson, Quatrano and Cuming (1985) found that there was a certain amount of regulation of protein stability in immature embryos cultured without ABA which, though synthesising the Em protein, did not accumulate it to detectable levels. They also concluded that ABA was probably not involved in the observed embryo-specific expression of the Em message. Examination of the Em promoter in rice protoplasts isolated a 650bp fragment which conferred ABA-regulated expression to a gene construct including GUS (beta-glucuronidase), exposure to 10^{-4} M ABA giving a rapid (within 1 hour) 30-fold increase in GUS activity (Marcotte, Bayley and Quatrano, 1988). Further analysis of the 5' flanking region of the Em message by Marcotte, Russell and Quatrano (1989) identified an ABA-responsive element (-152 to -103), active in either orientation in rice transient assays, and an element from +6 to +86 that increased the ABA response 10-fold.

Despite the evidence that ABA specifically regulates Em message and protein levels in culture, the rise in endogenous ABA during development *in planta* is not sufficient to give the observed increase in Em mRNA based on the responses seen *in vitro* (Morris *et al*, 1990). Morris and co-workers (1990) investigated the effect of osmotic stress on the levels of Em message and demonstrated a rapid (within 24 hours) accumulation of message after exposure to osmotic stress, the level of induction rising with increasing

severity of stress. No corresponding rise in endogenous ABA was seen upon this treatment, but the presence of ABA was shown to be necessary to maintain high levels of the Em transcript, possibly by specific stabilisation of messages (Williamson and Quatrano, 1988; Morris *et al*, 1990). Morris *et al* (1990) also established that the message could be induced after embryogenesis in early germination by osmotic stress or by ABA; Em transcript expression in non-embryogenic tissue had also been demonstrated in 5-10 DAI wheat seedlings, and in rice suspension cultures treated with exogenous ABA (Marcotte, Bayley and Quatrano, 1988). This suggests that the tissue-specific expression of the Em message can be overridden by ABA and osmotic stress, lending further support to the theory that the Em protein is important in protecting plant tissues from desiccation (Morris *et al*, 1990).

Litts *et al* (1992) recently isolated a rice Em gene which exhibits a very similar pattern of expression to the wheat Em and contains sequences in its 5' flanking region which bear a close relationship to the ABA responsive elements in wheat Em and another rice gene (Skriver and Mundy, 1990). Em has also been found in maize, where it follows the same developmental pattern as that seen in wheat, and can be induced rapidly and substantially by osmotic stress or ABA during embryo development or in early germination (Butler and Cuming, 1991). As in wheat the presence of ABA is necessary for the response to osmotic

stress; there is also an absolute requirement for the perception of ABA in maize, as a mutant with an inability to respond to ABA (*vp1*) produced no Em mRNA, despite application of exogenous ABA or osmotic stress (Butler and Cuming, 1991). More recent research (see section 1.3) has identified polypeptides showing significant amino acid homology to the Em sequence in several other plants including cotton, carrot and radish (reviewed by Espelund *et al*, 1992); these findings, together with the isolation of a range of proteins analogous in their expression patterns and inducibility by ABA and water stress (see section 1.3) imply not only that Em polypeptides may be found in a range of plants, both monocotyledons and dicotyledons, but also indicate a possible range of function-related proteins associated with protection from desiccation stress.

1.4.3 Ubiquitin

Ubiquitin is a small (76 amino acids) heat-stable protein whose presence has been demonstrated in both the nucleus and cytoplasm of all eukaryotic cells examined (reviewed by Hershko and Ciechanover, 1982 and by Finley and Varshavsky, 1985). Although the majority of work has been on animal systems, investigations of higher plant systems have revealed that the ubiquitin amino acid sequence, which is totally conserved among animals, differs from the characterised conserved plant sequences by only three amino

acids, with the yeast ubiquitin amino acid sequence differing from the plant sequence by two amino acids (Gausling and Barkardottir, 1986). This strong conservation of sequence argues a corresponding similarity of function for the ubiquitin protein in animal, plant and fungal systems and in eukaryotic micro-organisms.

In the cytoplasm ubiquitin exists either free or conjugated to a range of cytoplasmic and cell surface proteins, while in the nucleus ubiquitin has been found in association with histones; the functions of ubiquitin appear to revolve around this ability to conjugate to eukaryotic proteins, a primary function being the degradation of damaged or short-lived proteins (reviewed by Hershko and Ciechanover, 1982 and by Hershko, 1991).

1.4.3.1 Ubiquitin-Protein Conjugation and Protein Degradation

The removal of unwanted or damaged proteins to prevent incorrect associations and recycle free amino acids is an important process for the maintenance of cell viability. Research has shown that one pathway for the proteolysis of selected proteins involves the covalent linkage of a ubiquitin polypeptide to the target protein, followed by the formation of a multiubiquitin chain in a pathway elucidated by Hershko and Ciechanover (1982) and recently reviewed by Hershko (1991). In the first step of the pathway, ubiquitin is 'activated' by a specific enzyme, termed E_1 , which facilitates the transfer of ubiquitin to one

of a variety of ubiquitin carrier proteins (E_2 s). The E_2 carrier protein can transfer the activated ubiquitin to target proteins either directly or, more generally, in conjunction with a ubiquitin-protein ligase, E_3 . Examination of one E_3 ubiquitin protein ligase from yeast (E_3 alpha) suggests that it has a variety of specific binding sites, including ones for ubiquitin, E_2 carrier proteins and at least three types of protein substrate, while other ligases recognise further protein substrates; it is proposed that this system is important in selecting proteins for degradation (reviewed by Hershko, 1991). In the final step of ubiquitin conjugation, the ubiquitin molecules form covalent attachments to the target protein by isopeptide links from the ubiquitin's C-terminal glycine residue to ϵ -amino groups on the substrate protein's lysine residues. For the degradation process this is followed by the formation of a multiubiquitin chain. Recent research by Johnson *et al* (1992), using a constructed fusion protein of ubiquitin-proline-beta-galactosidase in yeast (*Saccharomyces cerevisiae*), has demonstrated that the presence of a single N-terminal ubiquitin polypeptide is a sufficient signal for the formation of a multiubiquitin chain (with the help of at least one of several ubiquitin-conjugating enzymes) and subsequent degradation of the protein. The degradation process itself has not been fully elucidated. It is thought to involve a 26S ATP-dependent protease complex, composed of three units termed conjugate degrading factors; one of these

has been identified as a 20S 'multicatalyst' protease component (reviewed by Hershko, 1991). After proteolysis it is thought that the protein monomers are regenerated for reuse by the action of isopeptidases, although the enzymes and mechanisms remain unclear.

Ubiquitin is thought to be involved in a number of roles other than selective proteolysis: an examination of the seven carrier proteins isolated from yeast indicates that only three E_2 s are thought to be associated with protein degradation; two others have been shown to be the products of genes RAD6 and CDC34, involved in DNA repair and in cell cycle progression respectively (reviewed by Finley, Bartel and Varshavsky, 1989). Ubiquitin has also been implicated in cell surface recognition and gene regulation (reviewed by Burke, Callis and Vierstra, 1988) and it has further been proposed that a reversible association of ubiquitin with an acceptor protein may modify the function of that protein without producing degradation or denaturation (reviewed by Finley, Bartel and Varshavsky, 1989 and by Hershko, 1991). That the ubiquitin conjugation and degradation processes are an essential part of cellular activity has been demonstrated by the use of mutants in yeast: deletion of the genes for E_1 , or the three degradation-related E_2 s, or the 26S protease complex are all lethal (reviewed by Hershko, 1991). Obviously intracellular proteolysis must be very tightly controlled; damaged proteins must be removed quickly while active proteins are not affected, while undamaged but

unwanted proteins must be degraded when their usefulness is at an end. For example, ubiquitin is thought to be instrumental in the degradation of cyclins, whose synthesis and removal is carefully controlled to allow the cell cycle to proceed normally. The control of ubiquitin activity can occur at many points - at the presence or activity of the many enzymes associated with ubiquitin activation, at ubiquitin-protein conjugation, protein selectivity, ubiquitin chain formation and also at the levels of free usable ubiquitin. Some or all of these factors may respond to cellular and developmental conditions to allow complex regulation of ubiquitin activity, both in degradation and in its other proposed functions. Further information of the activity, functions and possible regulation of ubiquitin may be provided by an examination of ubiquitin coding genes.

1.4.3.2 Ubiquitin Multigene Families

In all systems examined ubiquitin is coded for by multigene families, producing messages of various sizes (Gausling and Barkardottir, 1986; Ozkaynak *et al*, 1987; Burke, Callis and Vierstra, 1988; Baker and Board, 1991). Investigation of the ubiquitin multigene families in a range of organisms has led to the division of ubiquitin genes into two types: polyubiquitin, where a variable number of precise repeats of the 228bp ubiquitin coding sequence are present in tandem, head to tail with no spaces; and hybrid ubiquitin extension genes, consisting of a single ubiquitin sequence

followed by an unrelated C-terminal tail (Ozkaynak *et al*, 1987, reviewed by Neves *et al*, 1991 and by Baker and Board, 1991). These findings led Ozkaynak and co-workers (1987) to describe ubiquitin genes in yeast as being a 'family of natural gene fusions' with free ubiquitin polypeptides being produced only upon post-translational processing of precursors.

The number of ubiquitin repeats in the polyubiquitin genes varies both between species and within multigene families of species which have more than one polyubiquitin gene; the nucleotide sequences of the units also varies slightly within the same polyubiquitin gene, although the amino acid sequence is completely conserved (Ozkaynak *et al*, 1987; reviewed by Gausing and Barkardottir, 1986). All polyubiquitin genes contain 1-3 extra unrelated amino acids at the end of the final ubiquitin repeat (reviewed by Pollmann, Kampen and Wetter, 1991). Ozkaynak *et al* (1987) suggested that the extra residue(s), which vary between species, may prohibit the association of the ubiquitin polyprecursor with proteins for degradation, or other ubiquitin functions, until the polyprecursor has been processed.

Ozkaynak and co-workers (1987) first described the hybrid ubiquitin fusion proteins in yeast, from which they isolated three hybrid (UBI1-3) and one polyubiquitin gene (UBI4). Examination of the genes showed a total conservation of ubiquitin amino acid sequence, despite divergence of

nucleotide sequence. The C-terminal 'tails' of UBI1 and 2 consisted of 52 amino acids coding for the same protein, while the UBI3 tail was 76 amino acids in length and bore no significant homology to the 52 residue tail, apart from sharing certain common features. Both tails are highly basic, UBI1 containing 31% lysine and arginine and UBI3 29% Lys and Arg. Some of these basic residues go to form a consensus nuclear localisation signal, present at the end of the 52 amino acid and near the beginning of the 76 amino acid tails, of the forms Arg-Pro-Lys-Lys-Lys-Leu-Lys (UBI1 and 2) and Lys-Lys-Arg-Lys-Lys-Lys-Val (UBI3). Further, both tails contain a cysteine-rich consensus sequence forming a putative zinc-binding, nucleic acid binding domain, sometimes termed a 'zinc finger', of the form Cys-X₂₋₄-Cys-X₂₋₁₅-Cys-X₂₋₄-Cys (where X is any amino acid residue). The apparent localisation of these proteins to the nucleus, their basic nature and the proposed ability to associate with nucleic acids led to the suggestion by Ozkaynak et al (1987) that these ubiquitin extension proteins could be involved in the regulation of gene expression, either as nucleic acid binding proteins, with or without ubiquitin, or by enhancing the association of ubiquitin at specific sites on the chromosome; alternatively, these workers also proposed that the tail may act to regulate ubiquitin conjugation or degradation, or to conjugate itself to proteins in a similar manner to ubiquitin. However, later work by Finley, Bartel and Varshavsky (1989), using a range of yeast deletion mutants,

demonstrated that the fusion protein is cleaved to produce free ubiquitin and tail proteins, which play an important role in ribosome biogenesis, the UBI1 and 2 tails being associated with the large (60S) ribosomal subunit and the UBI3 tail with the small (40S) subunit. As the binding between the tail proteins and their respective ribosomal subunits has been shown to be very strong in the active ribosome, Finley and co-workers concluded that the tail proteins are actually components of the ribosome rather than factors involved in assembly or processing, with the putative nucleic acid binding area being associated with ribosomal RNA. The presence of ubiquitin in the original fusion protein is not essential for the action of the ribosomal tail proteins; however, it does give more efficient ribosome biosynthesis. Thus, rather than the tail proteins facilitating the transport or action of ubiquitin it transpires that the ubiquitin may act as a 'chaperone', enhancing the transport or assembly of the tail proteins into ribosomes (Finley, Bartel and Varshavsky, 1989).

Work in other species has demonstrated that ubiquitin extension protein tails are highly conserved between a variety of organisms. The human 52 and 80 residue tail proteins have also been identified as ribosomal proteins (in the large and small subunits respectively), with the ubiquitin part of the extension protein again increasing the efficiency of the association of the tail proteins into ribosomes (Baker and Board, 1991). The promoter and 5'

flanking sequences of the 52 amino acid tail extension protein gene have been shown to hold several components commonly found in mammalian ribosomal protein genes (Baker and Board, 1991). Yeast ubiquitin extension genes have also been shown to contain points of similarity with yeast ribosomal genes, including a T-rich stretch and areas bearing some resemblance to the ribosomal consensus upstream activation site in the promoter, and the presence of introns, which are rare in non-ribosomal yeast genes (Ozkaynak *et al*, 1987; Finley, Bartel and Varshavsky, 1989). Other highly homologous ubiquitin extension tails of 52-53 amino acids and 76-81 amino acids, all containing the characteristic basic nature, nuclear localisation site and cysteine-rich putative zinc-binding nucleic acid binding site (both in conserved positions), have been described in several other species. These include a 53 residue tail in the protozoa *Tetrahymena pyriformis* (Neves *et al*, 1991) and 52 residue tails in the alga *Chlamydomonas reinhardtii* (Pollmann, Kampen and Wetter, 1991), *Drosophila* (Cabrera y Poch, Arribas and Izquierdo, 1990) and *Arabidopsis thaliana* (Callis, Raasch and Vierstra, 1990). A similar tail sequence has also been identified in tobacco, but is unusual in having an extra 20 residues inserted into the 52 residue sequence (Genschik *et al*, 1990). Tails having a very similar sequence to the yeast 76 amino acid and human 80 amino acid ubiquitin extension tail proteins have been described in potato, tomato and *Arabidopsis* (reviewed by Garbarino, Rockhold and Belknap,

1992). Because of the similarity of tail protein amino acid sequence between species, the inevitable fusion to ubiquitin, and the conservation of putative nuclear localisation and binding sites it is generally assumed that the role in ribosome biogenesis discovered in yeast and humans is common to all species in which the tail proteins have been identified, and possibly to all organisms (Ozkaynak *et al*, 1987; Baker and Board, 1991; Pollmann, Kampen and Wettern, 1991). However, despite the similarity of ubiquitin-coding genes within and between species, investigations of their expression between different tissues and in response to stress has indicated that genes which code for the same ubiquitin polypeptides may have very different patterns of expression and possibly gene regulation.

1.4.3.3 Differential Expression of Ubiquitin Genes

In several of the organisms examined ubiquitin has been shown to respond to heat shock, while the use of yeast mutants has demonstrated that ubiquitin is involved in the response of cells both to heat shock and to other stresses (Finley and Varshavsky, 1985). Work by Ozkaynak *et al* (1987) on the ubiquitin genes of yeast showed that the polyubiquitin gene (UBI4) is specifically required for the ability of yeast cells to resist stresses (including high temperature and starvation) and is also strongly induced by these stresses, although it was shown to be not essential for the existence of non-stressed yeast cells. Examination of the UBI4

promoter showed the presence of a sequence with strong similarity to the consensus 'heat shock box' sequence; this box, when positioned upstream of heterologous promoters, has been shown to confer stress inducibility (Ozkaynak *et al*, 1987). Heat shock boxes have also been found in the promoters of other polyubiquitin genes (reviewed by Pollmann, Kampen and Wetter, 1991). In contrast, in *Chlamydomonas reinhardtii* both photoinhibition and heat shock led to a large fall in the level of RNA hybridising to a 52 residue tail ubiquitin extension probe, the repression of message level continuing for a longer period after removal of stress than that required by the alga to regain normal photosynthesis (3.5 hours); it was proposed that this decrease occurred to balance the greatly induced expression of algal polyubiquitin upon stress (Pollmann, Kampen and Wetter, 1991). The 53 residue tail ubiquitin extension protein from the protozoa *Tetrahymena pyriformis* was found to be unaffected by heat stress (Neves *et al*, 1991); similarly, the 80 amino acid ubiquitin extension protein message from potato tubers is not induced by heat shock, although it does respond to wounding or ethylene treatment (Garbarino, Rockhold and Belknap, 1992). An identical 80 amino acid tail extension message from tomato shows lowered transcript levels after heat stress, while an 81 residue tail extension protein message from *Arabidopsis* is unaffected (reviewed by Garbarino, Rockhold and Belknap, 1992). The results would appear to imply that ubiquitin from the polyubiquitin genes is

responsible for the response to heat and some other stresses. However, investigation of organisms with more than one polyubiquitin gene indicates that the induction upon stress is not universal. In potato, three polyubiquitin genes show very different responses to stress: one (*ubi7*, 1.6kb, 6 ubiquitin repeats), which is not detectable in unstressed tubers, is induced by ethylene treatment, heat shock and wounding; a second (*ubi16*, 1.8kb, 7 repeats) is also undetectable in unstressed tubers, induced slightly by ethylene treatment and strongly induced by heat shock and wounding; the third (*ubi9*, 1.8kb), although having the same number of ubiquitin repeats as *ubi16*, is abundant in control tubers but undetectable in stress tissue (Garbarino, Rockhold and Belknap, 1992). *Arabidopsis* has a multigene family of about 11 members and at least four transcript size classes (800-1900b); one of these messages (1.7kb) increases after two hours of heat shock, but another (1.35kb) shows a fall in transcript level upon the same stress (Burke, Callis and Vierstra, 1988). However, the fall in this smaller message occurs only in the flowers and buds, not in the leaves (Burke, Callis and Vierstra, 1988). A similar disparity of response to stress in different tissues is seen in rice: Borkird *et al* (1991) identified a rice polyubiquitin message (ca.2.4kb) which was responsive to heat shock and water stress in the leaf blades, where its expression was usually undetectable; very strongly induced by heat shock only in the roots, where the expression was usually low to undetectable;

and to heat shock, drying, salt and PEG treatments in the leaf sheath, which usually has a low level of expression, with heat shock being the most effective inducer. There are many other examples of differential expression of ubiquitin genes without stress in different tissues. Transcripts from a polyubiquitin gene in *Arabidopsis* accumulate to higher levels in flowers, buds, germinating tissue and etiolated tissue than in mature leaves, stems and roots (Burke, Callis and Vierstra, 1988). In yeast, although all four ubiquitin genes are expressed during exponential growth, the levels of the 52 amino acid tail ubiquitin extension messages (UBI1 and 2) are repressed in the stationary phase, while the level of the polyubiquitin message (UBI4) is increased (Ozkaynak et al, 1987). This phenomenon is also seen in barley leaves, where Gausing and Barkardottir (1986) found that although the large transcripts of the ubiquitin multigene family (which produces messages from 700-2000 nucleotides) were expressed constitutively, the smaller 700b messages, presumably representing ubiquitin extension messages, were found only in cells undergoing active division. Similar differential expression of ubiquitin extension transcripts between actively dividing and mature tissue has also been noted in tomato, *Arabidopsis* and potato (reviewed by Garbarino, Rockhold and Belknap, 1992). Garbarino and co-workers (1992) suggested that the observed increase in ubiquitin extension protein expression upon wounding or ethylene treatment was due to an increase in metabolic activity in the treated

tissue; they corroborated this proposal with evidence that the stress treatments also induced a noticeable increase in the transcription of ribosomal RNA. This indicates that ubiquitin extension protein expression is highest in tissues with high ribosome biogenesis and activity, either natural or induced, which agrees with the function of the extension tail protein. Garbarino, Rockhold and Belknap (1992) suggested that the ubiquitin polypeptides coded for by the ubiquitin extension genes were expressed specifically for facilitating the assembly of the co-translated ribosomal proteins, while the polyubiquitin genes provided the ubiquitin proteins for general cellular activity. This does not appear to be true in yeast, where the single polyubiquitin gene is not essential for cell viability, the three ubiquitin extension proteins producing sufficient free ubiquitin for cellular activity; however, the polyubiquitin gene is essential for resistance of the cells to stress, when it is responsible for producing large quantities of free ubiquitin (Ozkaynak *et al*, 1987). It is possible that the ubiquitin produced by both types of genes has a variety of roles in the cell, the proteins originally used for 'chaperoning' the tail ribosomal proteins being recycled and used for other processes in the nucleus, and ubiquitin from polyubiquitin genes being recycled after protein degradation/conjugation. Alternatively, the observed differential expression of genes within a multigene family, both between tissue types and in response to stress may suggest that in those organisms with

large multigene families the regulation of genes could be related to their function, some being directed to rapid ubiquitin synthesis for stress response, others providing ubiquitin for other roles. Ozkaynak *et al* (1987) noted that although the two yeast 52 residue tail ubiquitin extension genes coded for identical proteins their flanking sequences were largely non-homologous, perhaps suggesting differential regulation; for example, the upstream region of UBI2 contains stretches of poly(dA).poly(dT), also present in UBI3 but absent in UBI1, which are thought to be related to constitutive expression in yeast. Similarly, two polyubiquitin genes from potato which both contain 7 ubiquitin repeats (*ubi9* and *ubi16*) have completely different patterns of expression in response to stress, despite coding for the same polyprotein. Obviously the ubiquitin system, though common to a wide range of organisms, is very complicated. Examination of the flanking regions and regulation of the members of ubiquitin multigene families may aid not only the understanding of the extent of ubiquitin's role in the cell and how the precise controls required for protein degradation and stress response are effected, but also possibly the mechanisms of diversification of genes within a subfamily and between related organisms and species.

It is clear that ubiquitin plays an important role in stress response and it would be interesting to examine its expression in the developing seed, a system which should undergo increasing stress as it desiccates, then is liable to

further cellular and protein damage due to rapid influx of water into desiccated tissue upon imbibition. The isolation of two polyubiquitin cDNAs from a pea leaf library by Watts and Moore (1989) provided the means with which to investigate the nature of polyubiquitin in pea and the expression of ubiquitin coding messages during seed maturation and desiccation.

1.4.4 Chaperonin 60 - A Molecular Chaperone

Molecular chaperones are a group of unrelated proteins which engineer the correct assembly and transport of other polypeptides, but are not part of the resulting mature protein; it is generally agreed that this association with polypeptides protects potentially active sites exposed during folding and oligomerisation of the polypeptide structure, thus preventing incorrect associations within the polypeptide, or with other molecules, which would render the polypeptide biologically unfunctional (reviewed by Ellis, 1987 and 1990). The binding of the chaperones to the target polypeptide has been shown to be specific and non-covalent and is reversed when conditions are favourable for the formation of the functional structure, the release of the chaperone sometimes requiring the hydrolysis of ATP. Apart from the protection of active sites, both in the mediation of correct protein assembly and during polypeptide transport within the cell, chaperones have also been implicated in the disassembly of protein aggregates - either for rearrangement

during cellular cycles or after stress, or for degradation of unwanted protein structures (reviewed by Ellis, 1987 and by Gething and Sambrook, 1992). Many heat shock proteins have later been identified as chaperones (e.g. hsp60, hsp70, hsp90) and the increase in chaperone levels seen after such environmental stresses serves both to protect vulnerable polypeptides and also to disassemble incorrect structures or aggregates caused by the stress (Ellis, 1990).

The majority of molecular chaperones belong to three highly conserved families, found in bacteria, plants and animals - 'stress-70', 'stress-90' and chaperonins; although the protein structures are unrelated between families, many show a similarity of function (Gething and Sambrook, 1992). The chaperonins are a set of sequence-related chaperones which have been found in all bacterial, mitochondrial and plastid systems examined. Their expression is constitutive, giving high levels of protein, but can also be increased by stresses such as heat shock (reviewed in Ellis, 1990 and Gething and Sambrook, 1992 and references therein). The main group of chaperonins contains those that show homology to the *E.coli* GroEL subfamily, also called hsp60 and now defined as chaperonin 60 (Ellis, 1990). Chaperonin-60 proteins are large oligomers formed of fourteen 60K subunits in two rings of seven subunits. In *E.coli* they function mainly in protein folding and oligomer assembly; however they have also been shown *in vitro* to bind unfolded precursors of secretory proteins and, *in vivo*, are necessary

for the secretion of beta-lactamase (reviewed by Gething and Sambrook, 1992). Chaperonin-60 has been isolated in mitochondria of maize, animals and yeast cells. The yeast hsp60 is encoded by a nuclear gene and has been shown to be necessary for the correct assembly of certain proteins in the mitochondrial matrix; chloroplast chaperonin-60, which functions in the formation of plant Rubisco, is also nuclear-encoded. Unlike bacterial and mitochondrial chaperonin-60, chloroplast chaperonin can be separated into two subunits of 61K (alpha) and 60K (beta); however, the chaperonins retain functional homology, as demonstrated by the association of the wheat Rubisco small subunit with the bacterial GroEL (reviewed in Gething and Sambrook, 1992).

1.4.5 *PsMT_A* - A Putative Metallothionein Gene

Metallothioneins are seemingly ubiquitous proteins which have been shown to occur throughout the animal kingdom and more recently have been demonstrated in higher plants, eukaryotic micro-organisms and some prokaryotic micro-organisms (reviewed by Kagi and Schaffer, 1988). They commonly bind Cd, Cu and Zn and some also associate with other heavy metal ions such as Hg, Bi, Ag, Au and Pt; this ability, their high conservation of structure in metal-binding regions, their apparently universal occurrence and their ability to be induced by the presence of metal ions, hormones, cytotoxic agents and physical or chemical stress

all indicate that metallothioneins play an important part in the metal relations of the cell.

It has been suggested that metallothioneins may act to sequester or dispense essential metal ions (e.g. Zn and Cu), either unspecifically as a metal-buffering ligand or specifically, performing a specialised function in cellular maintenance, differentiation and repair. It is also assumed that they act in short-term detoxification of heavy metal ions, although the extent of their usefulness in this capacity has been questioned. A third possible role is as a metalloregulatory unit, functioning by association with DNA either directly or via certain DNA binding proteins which have been shown to have similar Zn-binding motifs (functions reviewed by Kagi and Schaffer, 1988).

Metallothioneins have been divided into three classes: the first two classes (I and II) occur as single chain proteins and are characterised by an abundance of Cys and the motif Cys-X-Cys (where X is any amino acid other than Cys). Class III is atypical, its polypeptides containing gamma-glutamylcysteinyl units. An examination of the metallothionein genes and expression has demonstrated that they are inducible by a variety of factors (as before), the induction occurring at the level of transcription initiation and giving a maximum concentration of the metallothionein within 24-48 hours of induction. Further, metal responsive elements have been identified in the promoters of some animal metallothioneins. There is little information available on

the presence of metallothioneins in the seeds of higher plants; however, recently the wheat Ec (early cysteine labelled) polypeptide has been identified as a Zn-binding class II metallothionein, containing the characteristic Cys-X-Cys sequence (Lane, Kajioka and Kennedy, 1987). The Ec message accumulates during embryogenesis (as do mammalian Zn metallothionein genes) with the highest levels occurring in immature embryos (15dpa) and decreasing as the embryo matures, being conserved in the mature embryo but undetectable in germinated seeds (Kawashima *et al*, 1992). The gene has been shown to have a known ABA-responsive element in its promoter (CACGTGCA); however, it has no metal responsive sequences like those seen in animal systems and the Ec message is not induced by the addition of Zn to germinating mature wheat embryos (Kawashima *et al*, 1992). The same workers suggested that, due to this lack of metal responsiveness and the fact that only ca.5% of the total Zn in a mature wheat embryo is associated with the Ec protein, it may act in Zn homeostasis rather than as a detoxicant.

Evans *et al* (1990) isolated a gene from pea roots (*Pisum sativum* cv. Feltham First) with an open reading frame that coded for a putative protein of 75 amino acids and had some similarity to metallothioneins, although the protein was not isolated. Further investigation showed that the gene transcript, which was present in the root as a single, fairly abundant message of ca.640 bases, was also present in both green and etiolated leaves and in cotyledons during early

development (14-15DAF). In both cases the transcript was less abundant in these tissues and also varied in size - ca.670b in leaf and ca.600b in developing seeds (L.Gatehouse, unpublished); results from hybridisations to the restricted pea genome suggested that the messages came from a small multigene family. The predicted protein was not isolated from pea but, by computer analysis, was shown to have significant homology to metallothioneins. It shared two regions of homology, containing the characteristic Cys-X-Cys motif, with a class I metallothionein from *Neurospora crassa*, the homologous area being on either side of a sequence with no Cys residues and no relationship to other metallothionein coding regions. The metal-binding properties of the gene product were confirmed by Tommey *et al* (1991) when they cloned the isolated gene into *E.coli* as a carboxyterminal extension of glutathione-S-transferase; after growth of the transformed bacteria in metal-supplemented media (Cd, Cu or Zn) they were able to isolate the expressed fusion protein and establish that there was a greater association of ions to the fusion protein than to the original GST protein. They also demonstrated the association of Zn with the *PsMT_A* portion after cleavage of the fusion protein. Their results indicated that *PsMT_A* has relatively high affinities for some metal ions when expressed in *E.coli*.

1.4.6 CpTI - Cowpea Trypsin Inhibitor Protein

Plant protease inhibitors function in the protection of

plant tissues against insect attack, affecting the metabolism of the insects (reviewed by Richardson, 1977). They are present in the leaves of a wide range of plant species and, in some cases, have been shown to be inducible by insect attack (Hilder *et al*, 1987). A large class which inhibits serine proteases (trypsin, chymotrypsin, elastase etc.) are common in various species, both mono- and dicotyledons, and are especially abundant in seeds and storage organs (reviewed by Richardson, 1977). Plant serine protease inhibitors are primary gene products and fall into three main structurally distinct families: wound induced inhibitors I and II from tomato and potato, whose expression can be induced upon wounding or insect attack; Kunitz type protease inhibitors; and Bowman-Birk type inhibitors, into which category the trypsin inhibitor isolated from cowpea falls (Hilder *et al*, 1989).

Bowman-Birk type protease inhibitors have been isolated from seeds of a range of legumes, and related proteins have been isolated from cereals (reviewed by Hilder, Gatehouse and Boulter, 1990). The protease inhibitors are 'double-headed', being able to inhibit two protease enzymes simultaneously. The Bowman-Birk inhibitors can therefore be divided into several types, inhibiting trypsin/trypsin, trypsin/chymotrypsin or trypsin/elastase (Hilder *et al*, 1989). The proteins are small, ca.8K, and share a typical core area with a high cysteine and serine content and considerable disulphide cross linking. There is a lower

degree of homology between proteases of different types in the amino- and carboxy-terminal regions; however, the proteins are synthesised as precursors with a leader sequence which is highly conserved between different members of the Bowman-Birk type family. As the proteins function mainly in the cytosol this leader sequence is thought to be associated with the action of the protein rather than its transport and localisation in the cell; Hilder *et al* (1989) suggested that the cleaving of the leader sequence may occur on insect attack and so activate the protease inhibitor, a function which would explain its high conservation.

In cowpea, Bowman-Birk type inhibitors are encoded by a small gene family containing at least four active genes producing three trypsin/trypsin and one trypsin/chymotrypsin inhibitors (reviewed by Hilder, Gatehouse and Boulter, 1990). The cowpea trypsin inhibitor proteins isolated have been shown to affect the metabolism of a wide range of insects, including a number of important pests; further, the gene *CpTI*, encoding a trypsin/trypsin protease inhibitor, has been shown to remain active when cloned into transgenic tobacco, the transformed plants producing a functional CpTI polypeptide which confers insect resistance to a major insect pest of tobacco (Hilder *et al*, 1987). This ability, together with its non-toxicity to mammals and the fact that the protease inhibitor acts upon the catalytic sites of more than one enzyme in the insect (and so it is extremely difficult for insects to develop resistance to the protein) makes the



CpTI gene a prime subject for the engineering of insect resistance in a range of plants (Hilder *et al*, 1987; reviewed by Hilder, Gatehouse and Boulter, 1990). J.A. Gatehouse and co-workers (unpublished) found that the *CpTI* protein accumulated towards the end of seed development, perhaps sharing a pattern of expression with LEA genes for seed protection against other stresses, such as the ABA-responsive proteins thought to be associated with disease resistance isolated from pea seeds (Barratt and Clark, 1991). Further investigation of its expression in the developing seed and the whole plant, under both normal growth conditions and insect attack, may aid understanding of the regulation of the plant's natural and inducible defence against insect damage.

1.5 AIMS OF THE PROJECT

In 1989, I.M. Evans and R. Swinhoe (unpublished) conducted an examination of the total protein from pea cotyledons at a range of ages, from 16 to 34DAF, with maturation occurring between 16 and 18DAF and desiccation of the mature seed beginning at 20DAF. They used both Coomassie blue staining, to give an indication of the proteins present in the seed, and pulse labelling with ¹⁴C-containing amino acids to show those proteins which were actively synthesised at each developmental age. They demonstrated that the pattern of protein synthesis changed during late embryogenesis and desiccation, and in particular they noted a

group of proteins whose synthesis appeared at the onset of desiccation and disappeared upon germination. Further examination of the cotyledon albumin fraction using ^{14}C - and ^{35}S - (cysteine and methionine) containing amino acids identified the appearance of at least four new bands on desiccation, at 30K, ca.21K, 17.5K and ca.9K. Protein bands which were fairly rich in methionine and occurred upon desiccation were also seen in a ^{35}S -methionine labelled protein extract from the embryonic axis (e.g. 14.3K, 9K). The presence of a methionine-rich 9K protein synthesised during desiccation suggested that pea may contain a protein with some similarity to the wheat Em polypeptide which exhibits similar properties and expression. These results indicated that pea seeds, in common with seeds of many other plants, exhibit noticeable changes in the proteins synthesised during development and also produce proteins which are abundant during late embryogenesis and desiccation but do not appear to be synthesised earlier in seed development or upon germination. The active synthesis of proteins at a developmental stage when protein synthesis in general is decreasing, the majority of storage proteins have already been accumulated and the proteins for germination and post-germinative processes are not yet expressed indicates that proteins expressed at this time have a definite role in the seed which is not connected with storage or germination.

The aim of this project was to investigate the

hypothesis that the expression of genes in the developing seed (measured both by message and protein levels) changes as the seed ages in parallel with the physiological and functional changes; that the timing of these changes can be influenced by external environmental factors, such as desiccation or ABA; and, further, that included within the messages expressed as the seed ages may be those which code for proteins to protect the seed from the effects of dehydration and rehydration and from other stresses to which the developing and quiescent seed may be exposed.

CHAPTER 2 - MATERIALS

2.1 MATERIALS FOR PLANT GROWTH

Seeds of *Pisum sativum* L. cv. Feltham First were obtained from Sutton Seeds, Reading, Berks, UK.

Phostrogen^R was from Phostrogen Ltd., Corwen, Clwyd, UK.

2.2 GLASSWARE AND PLASTICWARE

All glassware and plasticware for use with DNA or RNA was autoclaved before use. Tubes for precipitation and storage of RNA were siliconised with dimethyldichlorosilane solution (2% in 1,1,1-trichloromethane), after which glassware was rinsed with DEPC-treated water before being autoclaved.

For centrifugation, Corex tubes (Corning Glass Works, New York, USA), Falcon tubes (Becton Dickinson and Co., New Jersey, USA) and Eppendorf tubes were used, depending on volume and speed.

Sterilised CryoTubes (NUNC, Intermed, Roskilde, Denmark) were used for storage of RNA under liquid nitrogen.

2.3 KITS FOR MOLECULAR BIOLOGY

The mRNA Purification Kit (spun column chromatography) was obtained from Pharmacia Biosystems Ltd., Milton Keynes, Bucks, UK.

The PolyAtractTM mRNA Isolation System, the

Proclonone^R Lambda-gt10 System (including host cells C600 and C600Hfl) and the Packagene^R Lambda DNA Packaging System were from Promega Ltd., Southampton, UK.

The cDNA Synthesis Kit was produced by Boehringer Mannheim UK (Diagnostics and Biochemicals) Ltd., Lewes, East Sussex, U.K.

The Qiagen lambda midiprep kit was purchased from Hybaid Ltd., Teddington, Middlesex, UK.

2.4 BIOCHEMICALS, CHEMICALS AND CONSUMABLES

PUC18 and PBR322 vectors, Klenow enzyme, T4 polynucleotide kinase, T4 DNA ligase, Taq polymerase and restriction enzymes and their buffers were all obtained from NBL - Northumbria Biologicals Ltd., Cramlington, Northumberland, UK., as were the lambda-EcoRI/HindIII and lambda-PstI restriction markers and the X-gal.

Alkaline phosphatase, dNTPs and HGT agarose were obtained from Boehringer Mannheim UK.

Competent cells for the cloning of PUC vectors - DH5alphaME and DH5alphaSE - were supplied by Gibco BRL Ltd., Cambridge, UK, as was electrophoresis-grade agarose.

Sepharose^R CL4B, Sephadex^R G50 and Ficoll type 400 were obtained from Pharmacia Biosystems Ltd.

Radiochemicals, Hybond-N nylon membrane and the ECL detection system were supplied by Amersham International Plc, Amersham, Bucks, UK.

Ecoscint scintillation fluid was from National

Diagnostics Ltd. by B.S. and S. (Scotland Ltd.), Edinburgh.

Oligo(dT)-cellulose type 3 was purchased by Universal Biologicals Ltd., London, UK.

Nitrocellulose filters were obtained from Schleicher and Schuell, Dussel, FRG.

3mm paper and GF/C filters were from Whatmann Ltd., Maidstone, Kent, UK.

Hydroxyapatite powder (Bio-Gel HTP) and HRP-conjugated anti-rabbit antibody were obtained from Bio-Rad Laboratories Ltd., Hemel Hempstead, Herts, UK.

Components of bacterial culture media were supplied by Oxoid Ltd., Hamps, UK (yeast extract); Becton Dickinson Microbiology Systems, Cockeysville, MD, USA (trypticase peptone); and Difco Laboratories, Detroit, Michigan, USA (Bacto-agar).

X-ray films were Fuji RX, from Fuji Photo Film Co. Ltd., Japan.

All other chemicals (AnalaR grade) were supplied by BDH Chemicals, Poole, Dorset and Sigma Chemical Co., Poole, Dorset, UK.

2.5 LINKERS, ADAPTORS AND PRIMERS

EcoRI/NotI adaptors, EcoRI linkers and BamHI 12mer linkers were obtained from Pharmacia Biosystems Ltd. and are shown overleaf:

EcoRI/NotI Adaptor - 5'-d[AATTCGCGGCCGCT]-3'
(GCGCCGGCGA)p-3'

EcoRI 8mer linker - 5'-pd[GGAATTCC]-3'

BamHI 12mer linker - 5'-d[CCCGGATCCGGG]-3'

Primers for PCR were homologous to sequences on either side of the lambda-gt10 cloning site and were synthesised on an Applied Biosystems 381A DNA Synthesiser by J. Gilroy

No.288 AGCAAGTTCAGCCTGGTTAAG 21mer; T_m 62°C

No.289 CTTATTGAGTATTTCTTCCAGGGTA 24mer; T_m 66°C

2.6 PLASMID DNA FOR HYBRIDISATION AND SUBCLONING

2.6.1 PJC5.2

The plasmid PJC5.2 was used as a probe for the *Leg J* message. It contains a 1.42kb EcoRI restriction fragment from the 3' transcribed area of a genomic clone lambda-JC5, donated by the John Innes Institute (Norwich, U.K.), subcloned into PUC18 (Gatehouse *et al*, 1988).

2.6.2 PRC3.1

PRC3.1 was provided by Dr. R.R.D. Croy (University of Durham, UK) and contains a 2.4kb HindIII restriction fragment of a genomic clone of *Leg A* in PUC18.

2.6.3 PCU1 and PCU2

PCU1 (Watts and Moore, 1989) contains a 1.6kb EcoRI restriction fragment and PCU2 a 1.05kb EcoRI fragment from pea leaf polyubiquitin cDNA in PUC19; both plasmids were donated by Dr. F.Z. Watts (University of Sussex, UK).

2.6.4 PMuge Kla

PMuge Kla was donated by Dr. A.C. Cuming (University of Leeds, UK) and contains a 670bp KpnI restriction fragment from a wheat Em cDNA, encompassing the full mRNA sequence, subcloned into pBluescript KS⁺.

2.6.5 PASD1

PASD1 contains a 205bp PstI restriction fragment from the CptI (cowpea trypsin inhibitor protein) cDNA in PUC18 and was provided by A.S. Dale (University of Durham).

2.6.6 *PsMT_A*

The probe for *PsMT_A* was provided by Dr. A. Tommey (University of Durham) and contains a 235bp PCR product, amplified from a full length cDNA clone in PUC19, subcloned into the EcoRI site of PGEX3X. The original cDNA was isolated from a pea root cDNA library by L. Gatehouse (University of Durham).

2.7 ANTISERUM

A polyclonal antiserum to pea leaf plastid chaperonin 60 was prepared by Dr. L. Barnett and donated by Prof. R.J. Ellis (University of Warwick, UK).

CHAPTER 3 - METHODS

3.1 GROWTH AND TREATMENT OF PLANT MATERIAL

3.1.1 Growth and Harvesting of Peas

Seeds of *Pisum sativum* cv. Feltham First were germinated for 4-5 days in a dark spray room at 25°C with frequent water misting to give a relative humidity of ca. 100%. Seedlings were grown hydroponically in 2 litre culture bottles of Phostrogen^R (0.55g/litre) in a controlled environment cabinet. The growth conditions consisted of 16 hours of light (including morning and evening phases), temperatures of 25°C and 18°C (day and night respectively) and a relative humidity of c.70-80% (Evans *et al*, 1979). The Phostrogen solution was replaced every two weeks.

Flowers were labelled on the day of full flowering and pods harvested on the required day after flowering (DAF). The cut peduncle was sterilised briefly in 70% ethanol and the required seed organs (embryonic axes or cotyledons) removed aseptically and flash frozen in liquid nitrogen before being stored at -80°C.

3.1.2 Premature Desiccation Treatments

3.1.2.1 Air Drying

Pods were harvested at 14DAF and sterilised briefly in 70% ethanol. They were then air-dried at ambient

conditions in the controlled environment cabinet for 24, 48 and 72 hour periods. Control pods were removed in the same way and placed with their peduncles in sterile water for 24, 48 and 72 hours. Cotyledons were removed and treated as above.

3.1.2.2 Silica Drying

14DAF pods were harvested and dried in a closed container over activated silica gel for 48 and 72 hours and 8 days before removal and storage of the cotyledons.

3.1.3 Treatment with Exogenous ABA

14DAF pods were removed from the pea plants as before and placed in petri dishes with their peduncles immersed in a solution of either 10^{-6} , 10^{-5} or 10^{-4} M ABA in sterile water, or in sterile water for a control. Cotyledons and embryonic axes were harvested together after 24 or 72 hour treatments and stored at -80°C .

3.2 RNA EXTRACTION AND ISOLATION PROCEDURES

3.2.1 Extraction of Total RNA from Pea Cotyledons

Total RNA was extracted from cotyledons (or cotyledons plus embryonic axes) by an adaptation of the method of Hall *et al* (1978). The frozen cotyledons were warmed to -20°C before homogenisation in hot SDS buffer (0.2M

boric acid, 1% SDS, 30mM EGTA, 5mM DTT, pH 9.0, at 100°C, 30ml buffer per 10g tissue) in a Polytron^R (Kinematica, Luzern, Switzerland) for 20 seconds at speed 10. Isoamyl alcohol (iaa) was added to reduce foaming, then the mixture was cooled to below 40°C. The solution was incubated for 1 hour at 37-40°C with 0.3mg/ml proteinase K. The dodecyl sulphate was then removed by precipitation with 2M KCl (final concentration 0.14M) on ice for 30 minutes, followed by centrifugation for 10 minutes at 10,840g (Sorvall^R RC-5B refrigerated superspeed centrifuge, 4x50ml swing-out rotor). The clear supernatant was removed and precipitated overnight with LiCl (final concentration 2M) at 4°C. The resulting pellet was collected by centrifugation as above, washed twice with 5ml cold 2M LiCl (shaking followed by centrifugation) then dissolved in 0.2M potassium acetate pH 5.5. The supernatant from a further spin was precipitated for at least 3 hours with 2.5 volumes of ethanol at -20°C; the RNA was pelleted by another 10 minute spin, then resuspended in sterile distilled water and extracted twice with an equal volume of phenol:chloroform:iaa (25:24:1) before being reprecipitated with 3 volumes of 100% ethanol and 0.1 volumes 3M sodium acetate pH 5.2. The purified RNA was collected by centrifugation for 30 minutes, washed twice with 70% ethanol and vacuum dried, leaving the pellet slightly damp. The RNA was resuspended in sterile distilled water and its concentration measured at 260nm in a PU 8740 UV/VIS scanning spectrophotometer (using a conversion factor of a 1mg/ml

solution of RNA having an optical density of 24 absorbance units at 260nm). The RNA solution was then flash-frozen and stored under liquid nitrogen.

3.2.2 Extraction of Total RNA from Pea Embryonic Axes

Total RNA was extracted from small amounts of tissue (e.g. embryonic axes) by a modification of the guanidine hydrochloride extraction method described by Logemann *et al* (1987). The frozen axes were ground with a pestle and mortar in liquid nitrogen until a fine powder was obtained. This was then homogenised for 1 minute in a Polytron in 4ml extraction buffer (8M guanidine hydrochloride, 20mM MES, 20mM EDTA and 50mM 2-mercaptoethanol at pH 7.0). The homogenate was extracted with an equal volume of phenol/chloroform/iaa (25:24:1) by agitation and centrifugation of the resulting emulsion at 10,840g at room temperature for 45 minutes. The aqueous phase was removed and re-extracted with phenol/chloroform/iaa, the new aqueous phase being collected by a 30 minute spin at 10,840g. The RNA was precipitated overnight at -20°C with 0.7 volumes 100% ethanol and 0.2 volumes acetic acid. The solution was centrifuged for 10 minutes as above; the resulting pellet was then resuspended gently in 3M Na-acetate pH 5.2 at room temperature to dissolve the low molecular weight RNA and polysaccharides. The RNA was collected by centrifugation for 5 minutes as before and the resuspension procedure repeated with fresh sodium acetate. The RNA pellet was washed

carefully with 70% ethanol and centrifuged immediately to avoid dissolution. The RNA was air-dried (leaving the surface of the pellet damp) and then dissolved in DEPC-treated sterile distilled water. A sample was taken for spectrophotometric analysis and the remainder flash-frozen and stored under liquid nitrogen.

3.2.3 Isolation of Poly(A)⁺RNA

3.2.3.1 Oligo(dT)-Cellulose Spun Column Chromatography

Poly(A)⁺RNA was isolated from total cotyledon RNA by 2 rounds of affinity purification with oligo(dT)-cellulose, using a spun column method (Pharmacia mRNA purification kit, according to the method of Jacobson, 1987). The sample, containing up to 2mg total RNA, was heat-denatured at 65°C for 5 minutes then quenched and its salt concentration raised to 0.5M NaCl with sample buffer (10mM Tris-HCl pH 7.4, 1mM EDTA, 3M NaCl). It was then applied to a spun column containing 0.1g oligo(dT)-cellulose equilibrated in high salt buffer (10mM Tris-HCl pH 7.4, 1mM EDTA, 0.5mM NaCl) and left to soak in under gravity. The column was centrifuged at low speed (2 minutes at 350g) to remove the loading buffer; it was then washed twice with high salt buffer and 3 times with low salt buffer (10mM Tris-HCl pH 7.4, 1mM EDTA, 0.1M NaCl), centrifuging between each wash to remove unbound RNA. Bound poly(A)⁺RNA was eluted with four applications of prewarmed (65°C) elution buffer (10mM

Tris-HCl pH 7.4, 1mM EDTA), the eluate again being collected by centrifugation. The eluate was then heat-denatured again and treated as the first sample, using a fresh column. The final eluate was precipitated overnight with 2.5 volumes ethanol, 0.1 volume sample buffer and 0.01 volume glycogen solution (10mg/ml) at -20°C and the poly(A)⁺RNA collected by centrifugation for 30 minutes at 10,840g and 4°C. The pellet was washed with 70% ethanol (10 minutes at 10,840g) and vacuum dried (leaving the pellet slightly damp) before resuspension in a small volume of RNase-free water at 0°C. The concentration of the RNA was measured by spectrophotometry; it was then flash-frozen and stored under liquid nitrogen. The oligo(dT)-cellulose was regenerated by repeated washings in elution buffer, then 0.1M NaOH, then elution buffer again and equilibrated in high salt buffer before being desiccated overnight and stored at -80°C.

3.2.3.2 Isolation of mRNA using Magnetic Particles

Poly(A)⁺RNA was also isolated from total RNA using the PolyATtract™ mRNA isolation kit (Promega). Up to 5mg total RNA was denatured for 10 minutes at 65°C before the addition of a biotinylated-oligo(dT) probe (500pmoles) and 20xSSC (3M NaCl, 0.3M TriNaCitate buffer, pH 7.2, final concentration 0.48xSSC). While the mixture was cooling to room temperature (up to 30 minutes) the provided Streptavidin Paramagnetic particles (SA-PMPs) were resuspended and washed 3 times with 0.5xSSC using a magnetic rack to capture

particles between washes. The SA-PMPs were then resuspended in 0.5xSSC (0.5ml) and the RNA/biotinylated probe mixture was added. The reaction was incubated at room temperature for 10 minutes to allow annealing of the probe to the SA-PMPs, which were then collected at the side of the tube by the magnetic rack to enable removal of the supernatant. The particles were washed four times as before (resuspension and magnetic capture) with 0.1xSSC, then resuspended gently in RNase-free water to elute the poly(A)⁺RNA. The SA-PMPs were washed with a further 1ml of water to ensure elution and the two eluates precipitated with 2.5 volumes ethanol and 0.1 volume 3M Na-acetate pH 5.2. Poly(A)⁺RNA was collected by centrifugation, washed and resuspended as before; its concentration was measured by spectrophotometry and it was stored in aliquots under liquid nitrogen.

3.3 ANALYSIS OF NUCLEIC ACIDS BY GEL ELECTROPHORESIS

3.3.1 Analysis of RNA by Formaldehyde Gel Electrophoresis

RNA formaldehyde gels were prepared and run by the method of Miller (1987). A 1% HGT agarose gel containing 2.2M formaldehyde and 1xMOPS/EDTA (20mM MOPS, 5mM Na-acetate, 1mM EDTA, pH 7.0) was poured and allowed to set for 30-60 minutes at room temperature, then pre-electrophoresed in 1xMOPS/EDTA for 5-30 minutes at 60V. Small volumes of RNA samples were denatured in 0.8xMOPS/EDTA, 44% formamide and

1.9M formaldehyde for 10 minutes at 70°C then quenched on ice. 1.5ul gel loading buffer (0.7xMOPS/EDTA, 5mg xylene cyanol, 5mg bromocresol green, 400mg sucrose, 1.55M formaldehyde and 35% formamide) was added to each sample before loading. The gel was electrophoresed at 60V for 5-6 hours (until the leading dye front had migrated ca.9cm) with the buffer being mixed every hour. The gel was stained in 5ug/ml ethidium bromide for 5 minutes in the dark and destained for 2 hours or overnight in water in the dark. The gel was visualised and photographed on a 300nm uv trans-illuminator. DNA size markers for RNA gels were ethanol precipitated before use and treated as the RNA samples.

For quick analysis of RNA, 1.5% formaldehyde minigels were poured, the RNA samples treated as above and loaded with 1ul of 1mg/ml ethidium bromide in the well.

3.3.2 Analysis of DNA by Agarose Gel Electrophoresis

3.3.2.1 Restriction and Purification of DNA

Endonuclease restrictions of DNA were carried out as described by Sambrook, Fritsch and Maniatis (1989) in the appropriate buffer supplied with the enzyme. Techniques of phenol/chloroform/iaa (25:24:1) extraction and ethanol precipitation to purify and concentrate DNA are also in Sambrook, Fritsch and Maniatis (1989).

3.3.2.2 Size Analysis of DNA

DNA fragments were separated and analysed by electrophoresis in agarose gels of a range of concentrations (depending upon fragment size) as given in Sambrook, Fritsch and Maniatis (1989). Large gels (18cmx22cm) were cast and run in 1xTAE (40mM Tris-acetate pH 7.7, 1mM EDTA), minigels in 1x TBE (90mM Tris-borate pH 8.0, 2mM EDTA) with both gels and running buffers containing 0.4ug/ml ethidium bromide. Samples were loaded with 2ul gel loading buffer (0.1% orange G, 20% Ficoll 400, 0.1M EDTA pH 8.0). After electrophoresis (at 50-100V) the gels were visualised and photographed as above.

3.3.2.3 Isolation of DNA Fragments by Electroelution

DNA fragments were recovered from 0.6% agarose gels by electroelution for 25 minutes at 50V in sterile dialysis tubing and the buffer used for electrophoresis (Sambrook, Fritsch and Maniatis, 1989). The DNA solution was extracted once with phenol/chloroform/iaa and twice with chloroform/iaa before ethanol precipitation with 0.1 volume 3M Na-acetate pH 5.2 and 2.5 volumes ethanol at -20°C. The DNA was recovered by centrifugation as described previously.

3.3.2.4 Isolation of DNA Fragments using Silica Fines

Silica fines were donated by Dr. R.G. Alexander (now available as 'Finebind' - Amersham International Ltd., Bucks, UK). Fragments excised from a 0.7% agarose gel were

melted at 65°C for 10 minutes with 1ml NaI solution (6.05M NaI, 0.11M Na₂SO₃, filter sterilised and saturated with Na₂SO₃). The mixture was cooled to room temperature for 5 minutes then 5ul of silica fines were added and mixed gently before incubation for 10 minutes at room temperature (with intermittent inversion). The fines and bound DNA were pelleted by centrifugation (15 seconds at 12,000g) and the supernatant removed and replaced with 1ml 70% ethanol. The pellet was resuspended to wash the fines, then collected by another centrifugation. The dried pellet was then resuspended and incubated in 50ul TE buffer pH 8.0 (10mM Tris-HCl pH 8.0, 1mM EDTA) at 37°C for 10 minutes to elute the bound DNA, which was then ready for use without further purification or concentration.

3.3.2.5 Alkaline Agarose Gels

First and second strand cDNA synthesis was confirmed and analysed by alkaline agarose gel electrophoresis as described by Sambrook, Fritsch and Maniatis (1989). A 10ml, 1% alkaline agarose gel made in 1x alkaline buffer (0.3M NaOH, 0.02M EDTA) was set on a small glass slide. The samples were loaded in 2x loading buffer (25mM NaOH, 20% glycerol, 4.6% saturated bromophenol blue) and the gel run in 1x alkaline buffer at 100mA. The gel was then sealed in a plastic bag and exposed to X-ray film to locate the radioactive cDNA.

3.4 ANALYSIS OF NUCLEIC ACIDS BY BLOTTING AND HYBRIDISATION

3.4.1 Northern Blotting of RNA Formaldehyde Gels

Before blotting, RNA formaldehyde gels were washed several times in distilled water (or destained overnight in water) to remove the formaldehyde. The total RNA for Northern analysis was then blotted overnight onto HybondTM-N nylon membrane by the capillary transfer method found in Sambrook, Fritsch and Maniatis (1989) in a buffer of 20xSSC. After the blotting apparatus was dismantled and the position of the wells marked, the nylon membrane was air-dried for 30 minutes, wrapped in cling film and uv-irradiated for 5 minutes on a uv transilluminator (300nm). The cling film was then removed and the blot was baked for 1 hour in an 80°C vacuum oven. Nylon blots were stored in sealed plastic bags.

3.4.2 RNA Dot Blots

RNA was immobilised to nylon membrane using a BRL HybridotTM manifold according to the method of Kafatos *et al* (1979). 5ug samples of RNA were denatured in a solution of 6xSSC, 12% formaldehyde (0.05ug/ul RNA) at 60°C for 15 minutes. The samples were quenched on ice and the solutions diluted with 20xSSC to give a final concentration of 0.02ug/ul RNA and 15xSSC in ca.100ul. The nylon membrane was dampened in distilled water and equilibrated in 15xSSC for 20 minutes; the manifold was assembled in 15xSSC and attached to a water pump. Duplicate samples were then applied to the

sides of the wells, the samples including *E.coli* rRNA for a negative control and dilutions of the relevant probe DNA for a positive control (denatured as the RNA). Each well was rinsed with 100ul 15xSSC; the manifold was then dismantled and the membrane washed in 6xSSC before air-drying for 30 minutes. The nylon blot was then uv-treated, baked and stored as before.

3.4.3 Hybridisation of DNA Probes to Immobilised RNA

3.4.3.1 Preparation of a Radiolabelled DNA Probe

The probe DNA was separated from plasmid DNA by restriction and electrophoresis and isolated by electroelution or silica fines, as described previously. The DNA concentration was measured by spectrophotometry, given that 200D=1mg/ml for DNA at 260nm. Up to 0.25ug was then labelled by random priming by the method of Feinberg and Vogelstein (1983, 1984). The DNA was boiled for 3 minutes to denature then quenched on ice before mixing with 10ul OLB (hexadeoxyribonucleotides at 450D units/ml and dATP, dCTP, dGTP and dTTP (each at 0.096mM) in 240mM Tris-HCl pH 8.0, 24mM MgCl₂, 1M HEPES pH 6.6, 0.35% 2-mercaptoethanol), 0.04ug/ul BSA, 50uCi [α -³²P]dCTP and 2U Klenow enzyme in a final volume of 50ul. The reaction was incubated at room temperature for 2.5 hours or overnight, then stopped by the addition of 5ul 10% SDS. The labelled probe was separated from unincorporated nucleotides by chromatography on a 5ml

Sephadex G50 column equilibrated and run in 50mM Tris-HCl buffer pH 7.5, containing 150mM NaCl, 10mM EDTA and 0.1% SDS. Probe solution was then stored at -20°C until required.

3.4.3.2 Treatment of Blots and Hybridisation of Probes

RNA immobilised to nylon membranes was hybridised to radiolabelled DNA probes by the method found in Sambrook, Fritsch and Maniatis (1989). The blots were prehybridised for 1 hour at 42°C in sealed polythene bags in a shaking water bath. The prehybridisation solution consisted of 50% formamide, 2x Denhardt's solution (0.04% each of Ficoll 400, PVP and BSA), 5xSSPE (0.75M NaCl, 50mM NaH₂PO₄, 5mM EDTA), 0.1% SDS and 100ug/ml herring sperm DNA, the latter denatured by boiling for 7 minutes before being added to the solution at 42°C. The DNA probe was denatured by boiling for 7 minutes, quenched on ice for 2 minutes and then added directly to the pre-hybridisation solution. The blots were hybridised overnight at the same temperature.

To remove unbound DNA, the blots were washed (with agitation) twice at room temperature for 15 minutes in a solution of 2xSSC, 0.1% SDS, then once at 42°C for 10 minutes in 1xSSC, 0.1% SDS. The membranes were then placed on Whatman 3mm filter paper wrapped in cling film, then covered with another wrapping of cling film to prevent moisture loss, and exposed at -80°C to sensitised FUJI RM X-ray film with an intensifying screen. The hybridisation solution was stored at -20°C and could be re-used after heating to 65°C for 10

minutes. The blots could also be re-used after radiography, provided they had remained damp, by stripping the membranes with boiling 0.1% SDS, leaving for 30 minutes, then discarding the SDS solution and repeating with fresh boiling 0.1% SDS. The membranes were then autoradiographed to ensure that the probe was removed and stored wrapped in cling film in sealed plastic bags at 4°C. Results from RNA dot blots and Northern blots were analysed by densitometric scanning (Chapman *et al*, 1983).

3.4.4 Southern Blotting of DNA Agarose Gels

DNA fragments in agarose gels were transferred to nitrocellulose filters by the method of Southern (1975). The gel was denatured for two 30 minute periods in denaturing solution (0.5M NaOH, 1.5M NaCl, 1mM EDTA) then neutralised, again for two 30 minute washes, in 3M NaCl, 0.5M Tris-HCl, 1mM EDTA, pH 7.0. The gel was then rinsed in distilled water and blotted overnight onto a nitrocellulose filter (soaked in water then equilibrated in 20xSSC) by capillary transfer with 20xSSC as the transfer buffer. After blotting, the filter was marked, air-dried for 30 minutes and baked for 1-2 hours *in vacuo* at 80°C.

3.4.5 Hybridisation of a DNA Probe to Immobilised DNA

Southern blots were prehybridised for at least 3 hours at 65°C (shaken) in a prehybridisation solution

containing 5xSSC, 5x Denhardt's solution (0.1% each of Ficoll, polyvinylpyrrolidone and BSA), 0.1% SDS, 100ug/ml herring sperm DNA (boiled to denature before addition). The prehybridisation solution was then replaced with a smaller volume of the same solution, preheated to 65°C, the denatured probe was added and the filters were hybridised overnight at 65°C. For homologous probes, the filters were washed twice (30 minutes at 65°C with shaking) in 2xSSC, 0.5% SDS; for heterologous probes the salt concentration of the wash was raised (to a maximum of 5xSSC) and the SDS lowered (minimum 0.1% SDS). The washing temperature may also be lowered for heterologous probes, although this was not usually necessary. Washed filters were wrapped in cling film and exposed as Northern blots. The hybridisation solution was stored at -20°C and could be reused by heating to 90°C for 10 minutes to denature the DNA.

3.4.6 Colony Hybridisation

DNA from bacterial colonies was immobilised on nitrocellulose filters and hybridised to a DNA probe as described in Sambrook, Fritsch and Maniatis (1989). Putative transformants were transferred to duplicate nitrocellulose filters on YT plates (8g/litre trypticase peptone, 5g/l yeast extract, 5g/l NaCl, 15g/l Bactoagar) incorporating 40mg/l X-gal and 50mg/l ampicillin; colonies were grown up overnight on inverted plates at 37°C. One filter was then treated (colony side up) on Whatman 3mm paper soaked in 10% SDS (3

minutes); denaturing solution (0.5M NaOH, 1.5M NaCl, for 5 minutes); neutralising solution (1.5M NaCl, 0.5M Tris-HCl pH 7.4, for 5 minutes); and 2xSSC, again for 5 minutes, with the filter being blotted on dry 3mm paper between each treatment. The filter was then air-dried for 30 minutes at room temperature and baked for 1-2 hours in an 80°C vacuum oven. Hybridisation to a DNA probe and exposure of the filter were by the same procedure outlined for Southern blots.

3.5 CONSTRUCTION OF A PEA cDNA LIBRARY FROM DESICCATING COTYLEDONS

3.5.1 Construction of a cDNA Library using PUC18 as a Cloning Vector

3.5.1.1 cDNA Synthesis

Double-stranded cDNA was synthesised from poly(A)⁺ RNA isolated from 18DAF and 24,26,28DAF (pooled) cotyledons according to the method of Gubler and Hoffman (1983) using a cDNA synthesis kit (Boehringer Mannheim). 0.2-2ug of poly(A)⁺RNA was incubated in the buffer supplied by the kit with a deoxyribonucleotide mixture (dATP, dCTP, dGTP, dTTP, each at a final concentration of 1mM), primer oligo(dT)₁₅ (2 A₂₆₀/ml), 1uCi [alpha-³²P]dCTP, AMV reverse transcriptase and RNase inhibitor, in a final volume of 20ul. The first strand synthesis reaction was incubated at 42°C for one hour, after which 1ul samples were taken to analyse the incorporation

rate by TCA precipitation. For the second strand synthesis reaction, 0.5ul RNase H solution was added (half the amount suggested in the kit protocol to avoid damage to the DNA) along with a second buffer solution, a further 9uCi [α - 32 P]dCTP and *E.coli* DNA polymerase I (endonuclease-free), the volume being raised to 100ul. This reaction was incubated for 1 hour at 15°C, 1 hour at 22°C and 10 minutes at 65°C. Samples were again taken for analysis, then T4 DNA polymerase was added, and the cDNA solution incubated for 10-15 minutes at 37°C to ensure blunt ends. The reaction was terminated by the addition of 10ul EDTA (0.2M, pH 7.2) and 2ul Sarkosyl solution (10% w/v). The cDNA was then extracted once with an equal volume of phenol/chloroform/iaa and precipitated for at least 1 hour at -20°C with 10ug glycogen, 0.1 volume 3M Na-acetate pH 5.2 and 2.5 volumes ethanol (the glycogen being omitted in the second precipitation). The precipitate was collected by centrifugation at 4°C for 30 minutes at 12,000g, washed with 70% ethanol and vacuum dried before resuspension in sterile distilled water.

The yield of cDNA was estimated using the equation:

$$\text{yield (g)} = \frac{\text{cpm(incorporation)}}{\text{cpm(input)}} \times \text{mol.dCTP} \times 4 \times 330(\text{MW})$$

The cpm incorporated in the first and second strand syntheses were measured by liquid scintillation counting: 1ul samples of cDNA (2 per synthesis reaction) were adsorbed to Whatman GF/C glass-fibre filters (2.4cm diameter); one of each pair was washed free of unincorporated label by TCA precipitation

(Sambrook, Fritsch and Maniatis, 1989) and the filters dried; they were then placed in 1ml Ecoscint scintillation fluid and the radioactivity counted by a Packard 2000CA TriCarb Liquid Scintillation Analyser.

3.5.1.2 Preparation of the PUC18 Vector

PUC18 DNA was restricted for at least 2 hours with the required restriction enzyme and buffer. Where appropriate, the vector was dephosphorylated directly after restriction for 30 minutes at 37°C with 1U alkaline phosphatase (as in Sambrook, Fritsch and Maniatis, 1989). The vector was then extracted with phenol/chloroform/iaa, ethanol precipitated and stored at -20°C. Samples were analysed by agarose gel electrophoresis to confirm complete restriction.

3.5.1.3 Ligation of cDNA to PUC18 using EcoRI/NotI Adaptors

Ligation reactions were carried out according to the method of Sambrook, Fritsch and Maniatis (1989). Up to 2ug 18DAF cDNA was ligated overnight at 15°C to 0.02U (800-1000ng) EcoRI/NotI adaptors in a ligation buffer of 50mM Tris-HCl pH 7.6, 10mM MgCl₂, 10mM DTT, 0.5mM ATP with 2U T4 DNA ligase. The mixture was heated for 10 minutes at 65°C to denature the ligase, then quenched on ice. The adaptors were kinased for 30 minutes at 37°C with 10uCi [γ -³²P]dATP, 0.1mM ATP and 7U T4 polynucleotide kinase (Sambrook, Fritsch and Maniatis, 1989). The kinase was inactivated by heating

at 65°C for 10 minutes; the reaction was then quenched and extracted once with phenol/chloroform/iaa. The adapted cDNA was separated from the free adaptors on a 5ml Sepharose CL-4B column run in 10mM Tris-HCl pH 7.6, 1mM EDTA pH 8.0, 0.1M NaCl. The effluent was collected in 200ul fractions; 1ul samples of each fraction were diluted in 1ml Ecoscint scintillation fluid and the specific activity (cpm) counted by liquid scintillation. The relevant fractions were pooled, ethanol precipitated and collected by centrifugation at 12,000g and 4°C for 30 minutes. The pelleted DNA was washed in 70% ethanol, dried and resuspended in a small volume of sterile distilled water.

The adapted DNA was ligated overnight at 15°C to EcoRI-restricted PUC18 vector (both phosphorylated and non-phosphorylated) in the same conditions as the first ligation, with a molar ratio of either 1:2 (cleaved vector : insert) or 2:1 (dephosphorylated vector : insert). The reaction was terminated by heating at 65°C for 10 minutes, then quenched on ice.

3.5.1.4 Ligation of cDNA to PUC18 using EcoRI Phosphorylated Linkers

cDNA was end-labelled as described by Sambrook, Fritsch and Maniatis (1989) with 1-2U Klenow enzyme in a buffer containing 20uCi [α -³²P]dATP, 10mM each of dCTP, dGTP, dTTP, 50mM Tris-HCl pH 7.5, 10mM MgSO₄ and 0.1mM DTT for 15 minutes at room temperature. The reaction was heated

at 70°C for 5 minutes to denature the enzyme and ethanol precipitated twice. The labelled cDNA was then ligated overnight to 2U (800-1000ng) EcoRI phosphorylated linkers. After heating for 10 minutes at 65°C, the cDNA was restricted with 20U EcoRI (with the addition of the appropriate restriction buffer) for up to 2 hours at 37°C. The linkered cDNA was extracted once with phenol/chloroform/iaa, separated from the free linkers on a Sepharose CL-4B column, as before, and ethanol precipitated. The linkered cDNA was then ligated overnight to either EcoRI-restricted PUC18 (1 vector : 2 insert) or EcoRI-restricted dephosphorylated PUC18 (2 vector : 1 insert).

3.5.1.5 Ligation of cDNA to PUC18 using BamHI 12mer Linkers

End-labelled cDNA was ligated overnight to 2U BamHI 12mer linkers by the aforementioned procedure. After heating to denature the ligase, the linkers were restricted with 10U BamHI; the reaction was then phenol/chloroform/iaa extracted and the linkered cDNA collected as before and ligated overnight to either BamHI-restricted PUC18 or BamHI-restricted dephosphorylated PUC18.

3.5.1.6 Assessment of Linker Activity

To confirm that the EcoRI linkers were performing as specified, a linker check was carried out by the method described in Sambrook, Fritsch and Maniatis (1989). Samples of kinased linkers, ligated kinased linkers, and kinased,

ligated then restricted linkers were analysed on a 10% acrylamide gel cast and run in 0.5xTBE; this was then autoradiographed at -80°C to test the efficiency of the various reactions.

3.5.1.7 Transformation of Bacteria with Plasmid DNA

Transformation was carried out using an adaptation of the protocol provided with DH5alphaME (maximum efficiency) cells (BRL), modified from Hanahan (1983). The ligation mixture was diluted 5-fold and 1ul portions used to transform 20ul aliquots of competent cells DH5alphaME: the DNA was added to the cells on ice and mixed with the pipette tip; the cells were incubated on ice for 30 minutes, heat shocked for 45 seconds at 42°C and returned to ice for 2 minutes; 80ul SOC broth (2% trypticase peptone, 0.5% yeast extract, 10mM NaCl, 2.5mM KCl, 10mM MgCl₂, 10mM MgSO₄, 20mM glucose) was added to each tube; the tubes were shaken for 1 hour at 37°C and 20ul aliquots plated out onto YT plates supplemented with 0.02mg/ml X-gal and 0.05mg/ml ampicillin, which were incubated overnight at 37°C. Transformants were identified by the interruption of the beta-galactosidase gene, giving large white colonies as opposed to small blue untransformed colonies.

For test transformations, DH5alphaME cells were transformed as above with 1-10ng PUC18 DNA, either untreated or restricted and religated.

For subcloning, 50ul aliquots of DH5alphaSE

(subcloning efficiency) cells were used. The method was as above, except that 200ul of SOC broth was added instead of 80ul and the bacterial culture spread in 50ul and 200ul aliquots.

3.5.1.8 Preparation of Competent Cells

As an alternative to DH5alpha cells, TBI1 cells were made competent by the calcium chloride method found in Sambrook, Fritsch and Maniatis (1989). A single colony of TBI1 was taken from a freshly-grown plate and incubated for ca.3 hours (shaken at 37°C) in 100ml LB broth (10g/litre trypticase peptone, 5g/l yeast extract, 5g/l NaCl, pH 7.5) until the culture reached a concentration of ca.10⁷ cells/ml. The cells were transferred to 50ml sterile polypropylene tubes (Falcon 2098) and left on ice for 10 minutes to cool. The cells were collected by centrifugation for 10 minutes at 4,000g and 4°C. After removal of the supernatant and draining, the pellets were resuspended in 10ml 0.1M CaCl₂ at 0°C and stored on ice. The cells were recovered by centrifugation as before, drained and resuspended in 2ml 0.1M CaCl₂ at 0°C. The competent cells could be aliquoted, flash-frozen in liquid nitrogen and stored at this point at -80°C. 200ul aliquots of the TBI1 competent cells were transformed with up to 50ng plasmid DNA. After addition of the DNA and gentle mixing, the culture was incubated for 30 minutes on ice, heat shocked for 90 seconds at 42°C and cooled on ice for 2 minutes; 800ul SOC broth was then added and the culture

incubated at 37°C for 45 minutes with gentle shaking. 200ul aliquots were grown up overnight on YT/X-gal/ampicillin plates as before.

3.5.1.9 Preparation of Plasmid DNA for Restriction Analysis and Subcloning

Plasmids containing DNA inserts were isolated according to the method of Birnboim and Doly (1979). Transformed colonies were grown up overnight at 37°C (shaken) in 10ml aliquots of YT broth (8g/litre trypticase peptone, 5g/l yeast extract, 5g/l NaCl) containing 50ug/ml ampicillin. The bacteria were collected by centrifugation for 10 minutes at 3600g at room temperature and the supernatant discarded. The pellets were resuspended in 200ul of a solution containing 2mg/ml lysozyme, 50mM glucose, 10mM EDTA, 25mM Tris-HCl, pH 8.0 and incubated in sterile eppendorf tubes on ice for 30 minutes. 600ul of a solution of 0.2M NaOH, 1% SDS was added to each tube, mixed by inversion and incubated for 5 minutes on ice; 450ul of 3M Na-acetate pH 4.8 was then added, mixed by several inversions until a white DNA clot formed and incubated at 0°C for 1 hour. The solutions were centrifuged at 12,000g for 10 minutes at 4°C to obtain a clear supernatant, which was precipitated with 0.45 volume of isopropanol at -20°C. The pellet from a further 10 minute centrifugation was resuspended in 200ul 0.05M Tris-HCl buffer pH 6.0 containing 0.1M sodium acetate, then precipitated for 10 minutes with 100% ethanol at -20°C. The last

centrifugation and resuspension was repeated and the resulting solution extracted with an equal volume of phenol/chloroform/iaa before reprecipitation with 4 volumes of 100% ethanol. After centrifugation and vacuum drying, the plasmid DNA was redissolved in sterile water and aliquots were restricted with the appropriate enzyme in the presence of RNase A (final concentration 0.1ug/ul) and analysed by agarose gel electrophoresis.

3.5.1.10 Preparation of Plasmid DNA for Direct Sequencing

The above procedure was modified (Mierendorf and Pfeffer, 1987) to give plasmid DNA ready for sequencing. Transformed bacteria were grown up by shaking overnight at 37°C in 5ml LB broth supplemented with 100ug/ml ampicillin. 1.5ml of each culture was transferred to a 1.5ml sterile Eppendorf tube and spun at 12,000g for 1 minute to pellet the bacteria. The supernatant was discarded, another 1.5ml of the culture was added to the same tube and the centrifugation was repeated. As much as possible of the supernatant was removed and the pellet was resuspended by vortexing in 100ul of a solution of 50mM glucose, 25mM Tris-HCl pH 8.0 and 10mM EDTA, at 0°C. The suspended cells were incubated for 5 minutes at room temperature then 200ul of 0.2M NaOH, 1% SDS was added and mixed by inversion before incubation for 5 minutes at 0°C. 150ul of a solution of potassium acetate pH 4.8 (3M with respect to potassium, 5M with respect to acetate) at 0°C was added to the lysed cells and mixed by

inversion for 10 seconds, then left at 0°C for a further 5 minutes. Chromosomal DNA was pelleted by centrifugation for 5 minutes at 10,000g and 4°C and the supernatant removed and re-centrifuged to ensure that all of the precipitate was removed. It was then incubated with RNase A (final concentration 20ug/ml) for 20 minutes at 37°C. The plasmid DNA was extracted with an equal volume of phenol/chloroform/iaa and the aqueous phase precipitated with 2.5 volumes ethanol at -70°C for 30 minutes. The precipitate was collected by centrifugation (30 minutes at 10,000g, 4°C), rinsed with 70% ethanol (5 minutes at 10,000g) and vacuum dried before resuspension in 16ul sterile distilled water. The DNA was then precipitated with 6.5% PEG (MW 8000) and 0.4M NaCl for 20 minutes at 0°C and collected by centrifugation as before. The resulting dry DNA pellet was resuspended in water and analysed by restriction, gel electrophoresis and sequencing.

3.5.2 Construction of a cDNA Library using Lambda-gt10 as a Cloning Vector

A cDNA library was constructed in lambda-gt10 using the Proclone lambda-gt10 system (Promega), based on the method described by Huynh *et al* (1985) and Sambrook, Fritsch and Maniatis (1989).

3.5.2.1 Growth of Host Cells and Lambda

Single colonies of C600 or C600Hfl cells were grown up (according to the Promega protocol) for 8-16 hours in 50ml LB medium supplemented with 10mM MgSO₄ and 0.2% maltose, in Erlenmeyer flasks shaken at 37°C. The cultures were stored at 4°C until required, with a maximum storage time of 48 hours.

Lambda phage particles were diluted with phage buffer (20mM Tris-HCl pH 7.4, 100mM NaCl, 10mM MgSO₄) and adsorbed to host cells for 30 minutes at 37°C. They were then mixed with molten (45-47°C) TB top agar (10g/litre trypticase peptone, 5g/l NaCl, 8g/l Bactoagar, supplemented with 10mM MgSO₄) - 3ml for small (85mm) plates and 50ml for large (22.5x22.5cm²) plates - mixed gently and poured onto LB plates (LB medium with 15g/l Bactoagar). The hardened plates were then incubated either during the day (8-9 hours) or overnight at 37°C. C600 cells were used for the titration of recombinant lambda-gt10 after cloning, producing turbid plaques from parental phage and clear plaques from recombinant phage. C600Hfl gives only recombinant (clear) plaques and was used for recombinant selection and library amplification; C600Hfl was maintained on LB supplemented with tetracycline.

3.5.2.2 Ligation of cDNA to Lambda-gt10 Arms using EcoRI/NotI Adaptors

Radiolabelled cDNA was synthesised from 24,26,28DAF poly(A)⁺RNA as described previously, up to and including the

phenol/chloroform/iaa extraction. The aqueous phase was then applied to a 5ml Sephadex G50 column in a buffer of 50mM Tris-HCl pH 7.5, 150mM NaCl, 10mM EDTA and 0.1% SDS to remove free nucleotides which may inhibit ligation. After ethanol precipitation, the cDNA was ligated overnight to 0.05U EcoRI/NotI adaptors with 2U ligase. The reaction was terminated by heating at 65°C for 10 minutes, then chilled on ice; the adaptors were then kinased with 5U T4 polynucleotide kinase in the presence of 3mM ATP for 30 minutes at 37°C. The reaction mix was made up to 100ul with TE buffer (10mM Tris-HCl pH 7.4, 1mM EDTA), phenol/chloroform/iaa extracted and run through a 6ml Sepharose CL-4B column in TE containing 0.1M NaCl to remove free adaptors. Fractions containing cDNA were identified by liquid scintillation counting, pooled and ethanol precipitated. Samples from the trailing edge of the peak of cDNA from the column were not included in the precipitation, so as to avoid short, partial cDNAs and ensure avoidance of free linkers.

Ligation to lambda-gt10 and packaging took place according to the Promega protocols. Small scale ligations were packaged and dilutions adsorbed to bacteria and grown overnight to confirm successful cloning and to determine the most efficient ratio of vector:insert DNA. In the large scale reaction, 1ug adapted cDNA was ligated for 3 hours at room temperature to 4ug Protoclone lambda-gt10 DNA with 2U ligase in a buffer of 40mM Tris-HCl pH 7.5, 10mM MgCl₂, 10mM DTT, 1mM ATP, 50ug/ul BSA. The ligation reaction was then

mixed gently with the provided Packagene^R extract (Packagene *in vitro* packaging system, Promega) and incubated at 22°C for 2 hours. The packaged lambda was stored at 4°C in phage buffer with chloroform.

3.5.2.3 Titre, Amplification and Storage of Lambda-cDNA Library

Dilutions of the recombinant phage were adsorbed to C600 bacteria, plated and grown overnight at 37°C to calculate the number of recombinants obtained. The packaging mix was then amplified by adsorption of all the packaged phage to fresh C600Hfl cells which were grown up on an 85mm plate for 8-9 hours. The phage particles were eluted overnight at 4°C in 6ml phage buffer. The resulting solution was diluted and titred on LB plates as before to give the number of plaque-forming units (pfu) per ml, then stored at 4°C with chloroform (0.01 volume). The lambda library was at first stored in 1ml aliquots in sterile Eppendorfs; however, because of the substantial drop in titre noted under these storage conditions, it is now considered advisable to store lambda libraries in glass.

3.6 SCREENING A LAMBDA-cDNA LIBRARY WITH DNA PROBES

3.6.1 Immobilisation of Plated Lambda Library DNA to Nitrocellulose Filters

Nitrocellulose plaque lifts were taken and screened

by hybridisation according to the methods of Sambrook, Fritsch and Maniatis (1989). A freshly-diluted aliquot of the lambda-cDNA library was adsorbed to freshly-grown C600Hfl cells and grown on a large (22.5x22.5cm²) LB plate for 8-16 hours. The plate was stored at 4°C for at least 10 minutes to allow the agar to harden. Two plaque lifts were then taken using 20x20cm² nitrocellulose filters; the first filter was left on the plate for 5 minutes upside-down at 4°C, the second for 15 minutes. Both filters were marked with indelible ink when on the plate. The filters were placed plaque-side up on Whatman 3mm paper soaked in denaturing solution (0.5M NaOH, 1.5M NaCl) for 10 minutes, then blotted on dry 3mm paper and treated similarly with neutralising solution (1M Tris-HCl, 1.5M NaCl, pH 7.2) for 10-15 minutes and 2xSSCP (120mM NaCl, 13mM KH₂PO₄, 15mM TriNaCitrates, 2mM EDTA) for 10 minutes, blotting in between each treatment. They were then air-dried for 30 minutes and baked in a vacuum oven at 80°C for 1.5 to 2 hours.

3.6.2 Hybridisation of DNA Probes to an Immobilised Lambda-cDNA Library

The filters were prehybridised in sealed plastic bags in a 65°C shaken water bath for 3 hours in hybridisation solution as given for Southern blot hybridisation (3.4.4). The solution was then replaced with a smaller volume of hybridisation solution (at 65°C) containing the denatured

probe. For plasmid inserts, 0.25ug of DNA was labelled by random priming, as described in section 3.4.3.1, and used for hybridisation; when using total cDNA as a probe, 1ug was labelled in a scaled-up reaction and hybridised to the filters for 48 hours rather than overnight. Filters were washed and exposed to X-ray film as set out in section 3.4.3.2.

Areas containing plaques which were shown to hybridise strongly to the probe in the primary screen were removed with a sterile cork borer (7mm) and stored at 4°C in phage buffer with chloroform. The phage eluted from the positive plugs was diluted (usually to 10^{-4} or 10^{-5}) and adsorbed to freshly-grown C600Hfl cells, then grown overnight on 85mm LB plates. Duplicate nitrocellulose filters were taken from each plate and treated as given for the primary screen, the radioactive probe being heated to 90°C for 10 minutes before use to denature the DNA. Isolated positive plaques from the secondary screen were removed with a sterile siliconised Pasteur pipette and stored at 4°C in phage buffer with 0.01 volume chloroform.

3.7 PREPARATION OF A SUBTRACTED cDNA PROBE

A subtracted probe consisting of cDNA from desiccating cotyledons (24,26,28DAF) subtracted with 18DAF poly(A)⁺RNA was prepared and radiolabelled to a high specific activity by an adaptation of the method found in Sambrook,

Fritsch and Maniatis (1989).

3.7.1 Synthesis of Poly(A)⁺RNA

Total RNA was synthesised from 18DAF and 24,26,28DAF cotyledons by the hot SDS method (Hall *et al*, 1978, section 3.2.1). Poly(A)⁺RNA from both developmental stages was isolated using the Promega PolyATtract system (3.2.3.2). As a 10- to 30-fold excess of 18DAF:24,26,28DAF nucleic acids was required, extra 18DAF mRNA was isolated on regenerated oligo(dT) cellulose spun columns (section 3.2.3.1).

3.7.2 Synthesis and Treatment of cDNA

Single stranded cDNA was synthesised from 5mg 24,26,28DAF poly(A)⁺RNA as described by Sambrook, Fritsch and Maniatis (1989) in a reaction mix of 0.2ug/ul oligo(dT)₁₂₋₁₈, dATP, dCTP, dTTP, dGTP (each at a final concentration of 1mM), 50uCi [alpha-³²P]dCTP (specific activity 3000Ci/mmmole, 10uCi/ul), 50mM Tris-HCl pH 7.6, 70mM KCl, 10mM MgCl₂, 4mM DTT, 30U RNase inhibitor and 200U AMV reverse transcriptase, in a final volume of 60ul. After gentle mixing and a brief spin, the reaction was incubated for 2 hours at 37°C. 1ul samples were taken for yield estimation by TCA precipitation and the reaction was stopped by the addition of EDTA pH 8.0 and SDS (to 19mM and 0.38% respectively). The RNA was removed by alkaline hydrolysis, the solution being incubated for 30 minutes at 68°C with 0.3N NaOH then neutralised with

0.26M Tris-HCl pH 7.4 and 0.14N HCl. This solution was then extracted with an equal volume of phenol/chloroform/iaa and the aqueous layer passed through a Sephadex G50 spun column equilibrated in 1xTEN (10mM Tris-HCl pH 8.0, 1mM EDTA pH 8.0, 100mM NaCl) by centrifugation at 16,000g until the full volume was recovered (ca.5 minutes). A 20-fold excess by weight of 18DAF poly(A)⁺RNA was added to the labelled cDNA and the nucleic acids precipitated by incubation for 15-30 minutes at 0°C with 2.5 volumes ethanol at -20°C and 0.2 volume 5M ammonium acetate. The nucleic acids were collected by a 5 minute spin at 12,000g, the ethanol removed by aspiration and the pellet dried at room temperature.

3.7.3 Hybridisation of 18DAF Poly(A)⁺RNA to 24,26,28DAF

Radiolabelled cDNA

After resuspension in 6ul RNase-free water, the hybridisation reaction was set up by the addition of 488mM sodium phosphate pH 6.8, 6mM EDTA pH 8.0 and 0.24% SDS, with gentle mixing. The solution was overlaid with mineral oil and placed in boiling water for 5 minutes. The tube was removed to a 68°C water bath and hybridisation took place for up to 24 hours. The mineral oil was then removed and the hybridisation solution diluted with 1ml SS buffer (0.12M sodium phosphate pH 6.8, 0.1% SDS). Single stranded cDNA was then separated from double stranded DNA/RNA hybrids on a hydroxyapatite column.

3.7.4 Hydroxyapatite Chromatography

Hydroxyapatite (HAP) powder (Bio-Gel HTP) was suspended in 0.01M sodium phosphate. A disposable 3ml plastic syringe was assembled with a 2.4cm diameter Whatman GF/C filter covering the inside base and an 18 guage sterile hypodermic needle fitted to the end through a Neoprene gasket. Thin-walled 5mm plastic tubing leading to and from a 60°C water bath was wrapped around the syringe, held in place with masking tape and wrapped in aluminium foil. Water from the bath was pumped through this tube to maintain the temperature of the HAP column at 60°C (method devised by J. Davies, unpublished). The suspended HAP was added carefully to the syringe to give a packed volume of 1ml and was washed several times with 0.01M sodium phosphate pH 6.8 at 60°C. The needle was then sealed with a Neoprene stopper.

The hybridisation reaction was diluted with water to bring the sodium phosphate concentration down to below 0.08M (total volume 1.5ml). The sample was then loaded, the stopper removed and the solution allowed to flow into the column under gravity. The HAP column was washed with 3ml of 0.01M sodium phosphate at 60°C then re-sealed. One column volume (1ml) of 0.15M sodium phosphate (at 60°C) was added and left to soak in under gravity for 5 minutes; the stopper was then removed and the effluent (containing the radio-labelled single stranded cDNA) was collected in 0.5ml aliquots. Two further 1ml volumes of 0.15M sodium phosphate

were added, absorbed and the eluates collected in this way. The procedure was then repeated 3 times with 1ml volumes of 0.5M sodium phosphate at 60°C to elute the double stranded DNA/RNA hybrids.

The fractions containing the radiolabelled single stranded cDNA were pooled and concentrated by several extractions with an equal volume of 2-butanol (vortexing for 20 seconds and centrifugation at 12,000g for 30 seconds at room temperature followed by the removal and disposal of the upper organic phase) until the final aqueous volume was approximately 100ul. The DNA solution was then desalted on a Sephadex G50 spun column equilibrated in TE pH 8.0 and precipitated with 2 volumes ethanol and 0.1 volume 3M Na-acetate pH 5.2 overnight at -20°C.

3.7.5 Radiolabelling of Subtracted cDNA Probe to a High Specific Activity

The subtracted cDNA was labelled to a high specific activity using a scaled-up version of the random priming method (section 3.4.3.1). The resuspended DNA was heated to 60°C for 5 minutes to denature any secondary structures, cooled to room temperature, then incubated with 20ul OLB (to give final dNTP concentrations of 0.02mM each), 0.04ug/ul BSA, 4U Klenow enzyme and 250uCi [α -³²P]dCTP (final volume 100ul) overnight at room temperature. 10ul 10% SDS was added to stop the reaction, which was then phenol/chloroform/iaa extracted and purified on a 5ml G50 Sephadex

column run in 50mM Tris-HCl pH 7.5, 150mM NaCl, 10mM EDTA, 0.1% SDS, the fractions containing the cDNA being identified by liquid scintillation counting and pooled.

3.7.6 Hybridisation of Subtracted cDNA Probe to Immobilised Lambda-cDNA Library

The radiolabelled cDNA was denatured for 7 minutes at 100°C before hybridisation for 48 hours at 65°C (shaken) to duplicate filters of the lambda-cDNA library, prepared and prehybridised as described previously (3.6). After hybridisation, the filters were washed 2 times for 30 minutes at 65°C in 2xSSC, 0.5%SDS and exposed to X-ray film for 4 days, then for a longer exposure of 3 weeks.

3.8 SIMULTANEOUS SCREENING OF ISOLATED LAMBDA-cDNA CLONES

A method of screening several different isolated lambda-cDNA clones on the same filter was used. The cDNA library was screened with a probe of the library itself amplified by PCR and radiolabelled by random priming (see section 3.9.2.2). 49 plaques were isolated, the majority exhibiting high expression but with a few medium and low-expressing clones taken for comparison. A 20x20cm² grid with 49 (7x7) spaces was drawn on the base of a large (22.5x22.5cm²) petri dish. Fresh LB agar was poured into the plate and allowed to dry in a sterile air flow cabinet for 2-3

hours. 1ml of a fresh overnight culture of C600Hfl was then added to molten (45-47°C) TB top agar, poured and left to dry for up to 20 minutes. The phage solutions were diluted in phage buffer and 50ul portions of each lambda clone applied to separate squares of the labelled grid in a 37°C room; after all the liquid was absorbed to the bacterial layer (45-60 minutes) the plates were sealed, inverted and incubated overnight at 37°C. Duplicate nitrocellulose filters were taken from the plate and treated as outlined in section 3.6. The filters were prehybridised overnight and hybridised for 48 hours, one with the original lambda-library-PCR probe, the other with 18DAF total cDNA, radiolabelled by random priming. After exposure, the films were compared to assess differential expression of the chosen clones.

3.9 PREPARATION AND ANALYSIS OF CLONED cDNA INSERTS

3.9.1 Preparation of Recombinant Lambda DNA for Analysis

3.9.1.1 Qiagen Lambda DNA Midipreparation

DNA was prepared from 50ml phage liquid cultures using a Qiagen midipreparation kit, which is based on an adaptation of the method found in Sambrook, Fritsch and Maniatis (1989). Lambda particles from plaques for preparation were eluted overnight in 100ul phage buffer at 4°C. To this was added 300ul of an overnight culture of C600Hfl cells grown in LB medium supplemented with 10mM

MgSO₄, and the mixture was incubated for 20 minutes at 37°C to allow adsorption. The solution was then used to inoculate 50ml prewarmed (37°C) LB medium (containing 10mM MgSO₄) in an Erlenmeyer flask. The culture was shaken at 37°C for 3.5-5 hours until lysis occurred, with the addition of chloroform (final concentration 1%) before the last 15 minutes of the incubation, to aid lysis. The bacterial debris was pelleted by centrifugation in 100ml polypropylene tubes for 10 minutes at 8,000g. The supernatant was removed to a fresh tube and, if necessary, could be stored at 4°C overnight.

The phage suspension was incubated at 37°C for 30 minutes with a solution of 20mg/ml RNase A, 6mg/ml DNase 1 in 100mM Tris-HCl, 300mM NaCl, 10mM EDTA and 0.2mg/ml BSA, pH 7.5, then for 1 hour on ice with a solution of 30% PEG (MW 6000), 3M NaCl. The DNA was pelleted by centrifugation for 10 minutes at 10,000g, dried briefly and resuspended by pipetting in 100mM Tris-HCl, 100mM NaCl, 25mM EDTA, pH 7.5; 4% SDS was then added, mixed gently by inversion and the solution incubated at 70°C for 20 minutes to break the lambda protein coat, before cooling on ice. Next, 2.55M K-acetate pH 4.8 was added, the tube inverted and the solution centrifuged for 30 minutes at 10,000g. The supernatant was then removed with care (avoiding the potassium dodecyl sulphate layer) and applied to a Qiagen-pack 100 column, equilibrated in 750mM NaCl, 50mM MOPS, 15% ethanol, pH 7.0. The column was washed twice with 1M NaCl, 50mM MOPS, 15% ethanol, pH 7.0 and the DNA eluted slowly with 1.2M NaCl,

50mM MOPS, 15% ethanol, pH 8.0. After precipitation at room temperature with 0.8 volume isopropanol, the DNA was centrifuged for 30 minutes as above, washed with 70% ethanol, dried by vacuum and redissolved in TE buffer pH 8.0. The lambda DNA was then analysed by restriction and agarose gel electrophoresis.

3.9.1.2 Small-scale Isolation of Lambda DNA

Small quantities of lambda DNA were prepared for analysis and subcloning by an adaptation of the method found in Sambrook, Fritsch and Maniatis (1989), devised by A. Tommey (unpublished). Dilutions of phage containing the relevant positive plaques were prepared in phage buffer, adsorbed to freshly-grown C600Hfl bacteria and grown overnight at 37°C on LB plates with both bottom and top agar prepared with 1.5% electrophoresis grade agarose instead of Bactoagar (some of whose contaminants inhibit later restriction steps). The plaque density for DNA isolation should be close to confluent lysis. The phage plate lysate was eluted into 4ml SM (20mM Tris-HCl pH 7.4, 100mM NaCl, 10mM MgSO₄, 0.1% gelatin) by shaking at room temperature for at least 2 hours. The suspension was transferred to a glass corex tube, with another 1ml of SM used to rinse the plate. The bacterial debris was then pelleted by centrifugation for 10 minutes at 8,000g and 4°C and the supernatant removed and treated with 1ug/ml each of RNase A and DNase 1 for 30 minutes at 37°C. The phage particles were precipitated by

incubation at 0°C with an equal volume of 20% PEG (MW 8000), 2M NaCl in SM for at least 2 hours. The solution was spun at 10,000g for 20 minutes at 4°C, the supernatant discarded and the drained pellet resuspended overnight with 0.5ml SM.

After transfer to a sterile Eppendorf tube, the phage solution was spun for 2 minutes at 8,000g to remove any remaining debris, then treated again with RNase A for 20 minutes at 37°C. The solution was extracted 3 times with an equal volume of chloroform/iaa (24:1); the aqueous phase was then incubated for 15 minutes at 68°C with 5ul each of 10% SDS and 0.5M EDTA pH 8.0 (final concentrations approximately 0.1% and 5mM respectively) to break the protein coat and release the lambda DNA. The protein was removed by extraction with an equal volume of phenol, then once with phenol/chloroform/iaa and finally with chloroform/iaa, then precipitated overnight at -80°C with an equal volume of isopropanol. After thawing, the lambda DNA was collected by centrifugation for 30 minutes at 4°C, 12,000g, rinsed with 70% ethanol and dried under vacuum before resuspension in 50ul TE buffer pH 8.0.

3.9.1.3 Restriction of Recombinant Lambda DNA

To release cloned cDNA inserts, recombinant lambda DNA was cut with restriction enzymes at the cloning sites (EcoRI/NotI). 5-10ug lambda DNA was restricted in the appropriate buffer with a 3-fold excess of the restriction enzyme in a final volume of 50ul, the incubation taking place

at 37°C for at least 2 hours. If required, the lambda arms were heated at 68°C for 10 minutes to separate the cohesive ends, before analysis by gel electrophoresis. Where other combinations of restriction enzymes were used to isolate cDNA inserts, the restriction buffer most appropriate for the least active endonuclease was used.

3.9.2 Amplification and Isolation of cDNA Clones by PCR Technology

3.9.2.1 Amplification of Inserts from Recombinant Lambda-cDNA Clones by PCR

The cDNA inserts from positive plaques isolated from secondary screens were amplified by the polymerase chain reaction (PCR) according to the methods of Sambrook, Fritsch and Maniatis (1989). DNA from positive plaques was eluted overnight at 4°C in 1ml phage buffer (with 1% chloroform). 40ul of this solution was denatured by boiling for 5 minutes; 20ul was then taken, cooled to room temperature and added to a solution containing all four dNTPs at saturating conditions (200uM each, at pH 7.0), 100pmoles each of primers 288 and 289 (homologous to sites either side of the lambda EcoRI insertion site) and 4.5U Taq polymerase, in a buffer supplied by Promega (standard buffer: 50mM KCl, 10mM Tris-HCl pH 8.3 [at room temperature], 1.5mM MgCl₂, 0.01% gelatin) to give a final volume of 100ul. The samples were overlaid with 100ul mineral oil and amplified in a Pharmacia LKB Gene ATAQ

Controller. The amplification conditions consisted of 29 cycles of 1.5 minutes at 94°C (denaturation), 1 minute at 50°C (annealing) and 2 minutes at 72°C (polymerisation); this was followed by 1 cycle of 1.5 minutes at 94°C, 1 minute at 50°C and 5 minutes at 72°C. The amplified inserts were stored at -20°C before analysis and isolation by agarose gel electrophoresis and electroelution.

3.9.2.2 Amplification of Lambda-cDNA Library for Use as a Probe

20ul of the undiluted lambda library was amplified by PCR as above to give an alternative probe to 24,26,28DAF cDNA. After amplification and removal of the mineral oil, the solution was extracted once with phenol/chloroform/iaa and twice with chloroform/iaa then ethanol precipitated to remove free nucleotides. A small sample of the amplified library was checked by gel electrophoresis and up to 1ug was labelled overnight as described previously (section 3.4.3.1) for use in differential screening experiments.

3.9.3 Subcloning of PCR Products

Efficient subcloning of cDNA inserts amplified by PCR was achieved using a method devised by Jung, Pestka and Pestka (1990). The purified and concentrated inserts were phosphorylated for 45 minutes at 37°C with 7U polynucleotide kinase and 0.6mM ATP in linker-kinase buffer (66mM Tris-HCl pH 7.6, 10mM MgCl₂, 10mM DTT). The kinase was destroyed by

heating for 10 minutes at 65°C; the solution was ethanol precipitated to obtain a small volume for ligation and the fragments were concatamerised by ligation overnight at 15°C with 1U ligase. The ligase was destroyed at 65°C for 10 minutes; the concatamers were then restricted with EcoRI, phenol/chloroform/iaa extracted, ethanol precipitated and ligated overnight to PUC18 cut with EcoRI. The ligation mix was diluted and used to transform DH5alphaME cells.

3.9.3.4 Sequencing of Subclones

DNA from transformants was prepared by the direct sequencing method (section 3.5.1.10) and sequenced on a 373A DNA Sequencer (Applied Biosystems), using the method described in the ABI user manual, adapted from Sanger *et al* (1977) and McBride *et al* (1989).

3.10 PROTEIN EXTRACTION AND ANALYSIS

3.10.1 Extraction of Protein from Pea Cotyledons and Embryonic Axes

3.10.1.1 Extraction in PBS

Samples of frozen embryonic axes were ground in a mortar and pestle in liquid nitrogen, then extracted overnight in 1xPBS (phosphate-buffered saline - 8g/litre NaCl, 0.2g/l KCl, 1.44g/l Na₂HPO₄, 0.24g/l KH₂PO₄), 20ug/ml

leupeptin, 720ug/ml PMSF (in dry ethanol) and 100uM EDTA (100mg tissue/ml) with continuous shaking at 4°C. The extracts were transferred to 15ml corex tubes and spun for 10 minutes at 10,840g. The supernatant was removed and precipitated with TCA (final concentration 10%) for 1 hour at 0°C. The pellets were collected by centrifugation as above, washed in 70% ethanol, dried under vacuum and resuspended in 1x SDS sample buffer (0.1M Tris-HCl pH 6.8, 10% glycerol, 1% SDS, 0.001% bromophenol blue, 0.5% DTT) to a final concentration of 5mg/ml. Samples were stored at -20°C and analysed by polyacrylamide gel electrophoresis (PAGE).

3.10.1.2 Extraction in SDS Sample Buffer

Total protein was extracted from cotyledons or embryonic axes by grinding to a powder (as above) and suspending in 1xSDS sample buffer (omitting the DTT) to 40mg/ml and shaking overnight at 4°C. Debris was pelleted as above and the supernatant removed and stored at -20°C. DTT was added to 0.5% before the samples were analysed by PAGE.

3.10.2 Pulse Labelling and Extraction of Total Protein from Pea Cotyledons

Half cotyledons were isolated aseptically from the testa and embryonic axis and placed in a sterile petri dish, flat surface down, on 20ul drops of a ¹⁴C amino acid mix (Amersham, 50uCi/ml). The cotyledons were incubated in the closed petri dish, with damp filter paper in the lid, at room

temperature for 4 hours. They were then rinsed briefly in sterile water and a thin slice taken from the flat surface of each half cotyledon with a sterile razor blade. The tissue was homogenised in 1xSDS sample buffer (with no reducing agent) using a Polytron; total protein was then extracted by shaking overnight at 4°C (20mg tissue per ml buffer), boiled for 5 minutes and spun at 12,000g for 10 minutes before analysis by gel electrophoresis.

3.10.3 Extraction of Albumins from Pea Cotyledons and Embryonic Axes

Frozen tissue from either pea cotyledons or embryonic axes was ground to a fine powder as above. The albumin fraction was then extracted by shaking the powder overnight at 4°C in 25mM Na-acetate buffer pH 4.8 at a concentration of 40mg tissue/ml. The extracts were spun for 10 minutes at 12,000g and the supernatants stored at -20°C.

3.10.4 Analysis of Proteins by Polyacrylamide Gel Electrophoresis

3.10.4.1 One-dimensional Gel Electrophoresis

Polyacrylamide gels were cast and run according to the methods of Hames (1981) using a minigel apparatus. The main gel, consisting of 12.5% polyacrylamide (30% acrylamide: 1% Bis-acrylamide), 1.5M Tris-HCl pH 8.8, 0.38mg/ml ammonium

persulphate, 0.1% SDS and 0.03% TEMED was set for 30 minutes, then overlaid with stacking gel (4.6% polyacrylamide [30% acrylamide: 0.344% Bis-acrylamide], 0.5M Tris-HCl pH 6.8, 0.51mg/ml ammonium persulphate, 0.1% SDS and 0.1% TEMED), set for 20 minutes. Samples and markers were loaded in 1xSDS sample buffer, being boiled for 3-5 minutes then cooled to room temperature before loading into wells under the running buffer (0.25M Tris-HCl pH 8.3, 1.92M glycine, 1% SDS). Gels were run at 80-100V for 2-4 hours, until the dye edge was at the bottom of the gel. They were then fixed in a solution of 20% methanol and 10% glacial acetic acid until the bromophenol blue turned yellow, rinsed and stained for 4 hours or overnight with gentle shaking in 5 volumes of Coomassie blue stain (0.05% Coomassie blue, 50% methanol, 7% glacial acetic acid). Gels were destained for 4-8 hours (again shaken at room temperature) in a solution containing 50% methanol and 7% acetic acid, then rinsed in water. Gels were dried at room temperature between 2 layers of wet dialysis membrane (cellophane) for 1-2 days and stored at room temperature.

3.10.4.2 Fluorography of ¹⁴C-labelled Protein Separated by Polyacrylamide Gel Electrophoresis

Polyacrylamide gels containing radioactively-labelled proteins were run as above and fixed overnight in 10% TCA, 30% methanol. They were then treated by shaking for two 30 minute washes in DMSO, then for 3 hours in 20% (w/w)

PPO in DMSO and 1 hour in 7% acetic acid (Hames, 1981). The gel was dried for 48 hours in dialysis membrane at room temperature then exposed to X-ray film at -80°C .

3.10.5 Analysis of Proteins by Western Blotting

3.10.5.1 Immobilisation of Proteins on Nitrocellulose

Total proteins separated by electrophoresis on polyacrylamide minigels were blotted onto nitrocellulose filters by the method of Khyse-Anderson (1974). The gel was placed in a stack comprising of: anode plate; 2 layer of 3mm paper soaked in anode buffer 1 (0.3M Tris, 20% (v/v) methanol, pH 10.4); 1 layer of 3mm paper soaked in anode buffer 2 (25mM Tris, 20% (v/v) methanol, pH 10.4); prewetted nitrocellulose filter; 1 layer of 3mm paper soaked in cathode buffer (25mM Tris, 40mM 6-aminohexanoic acid, 20% (v/v) methanol, pH 9.4); one layer of dialysis membrane soaked in distilled water; 2 more layers of 3mm paper soaked in cathode buffer; cathode plate. The gel was blotted at 0.2A for 45 minutes and the filter stored at room temperature.

3.10.5.2 Immunological Detection of Proteins

Specific proteins were identified using an ECL detection system (Amersham, modified by Marion Longstaff). Nitrocellulose filters were blocked for 1 hour at room temperature (or overnight at 4°C) with shaking in 1xPBS, 5% Marvel and 1% Tween 20. The filter was incubated with the

primary antibody, diluted 1 in 5000 or 1 in 10,000 in antisera buffer (1xPBS, 5% Marvel, 0.1% Tween 20) for at least 2 hours (shaken) at room temperature or overnight at 4°C. The filter was then washed three times in antisera buffer before incubation with the secondary antibody (HRP-conjugated anti-rabbit antibody [Bio-Rad]) at a 1 in 2000 dilution for at least 1.5 hours as above. The filter was washed by shaking at room temperature first in PBS, 0.1% Tween 20 for two incubations of 5 minutes and one of 15 minutes, then in water for two 5 minute periods. The filter was then drained quickly before overlaying (protein side uppermost) with a mixture of equal volumes of detection reagents 1 and 2 (Amersham, total volume 0.125ml/cm filter) for 1 minute. After draining, the filter was placed between acetate sheets and exposed to X-ray film for 15 seconds, then one minute, or whatever time needed to obtain a satisfactory autoradiograph. Filters could be stripped and reused if desired.

CHAPTER 4 - CONSTRUCTION AND SCREENING OF A cDNA LIBRARY FROM DESICCATING PEA COTYLEDONS

4.1 RESULTS

The quantitative and qualitative changes in the mRNA population of the seed during late embryogenesis are a necessary preparation for the survival of the seed during extreme desiccation and a preparation for germination. To examine the message population in pea seeds during dehydration it was considered necessary to construct a cDNA library. This would allow an investigation of the expression of messages which are abundant during late embryogenesis, and also provide information on the expression of messages for seed-specific proteins and plant proteins associated with responses to stress in the dehydrating seed. Pea cotyledons were used as an abundant source of poly(A)⁺RNA and an age of 24,26,28DAF was chosen, based on the observation by I.M. Evans and R. Swinhoe (unpublished) of new protein synthesis in pea cotyledons at this point in the dehydration of the mature seed (see section 1.5).

4.1.1 The Construction of a cDNA Library from Desiccating Pea Cotyledons using PUC18 as a Cloning Vector

Initial attempts to construct a cDNA library used the plasmid PUC18 as a cloning vector. To carry out the cloning, the synthesised cDNA was ligated to adaptors or linkers, separated from free oligonucleotides and

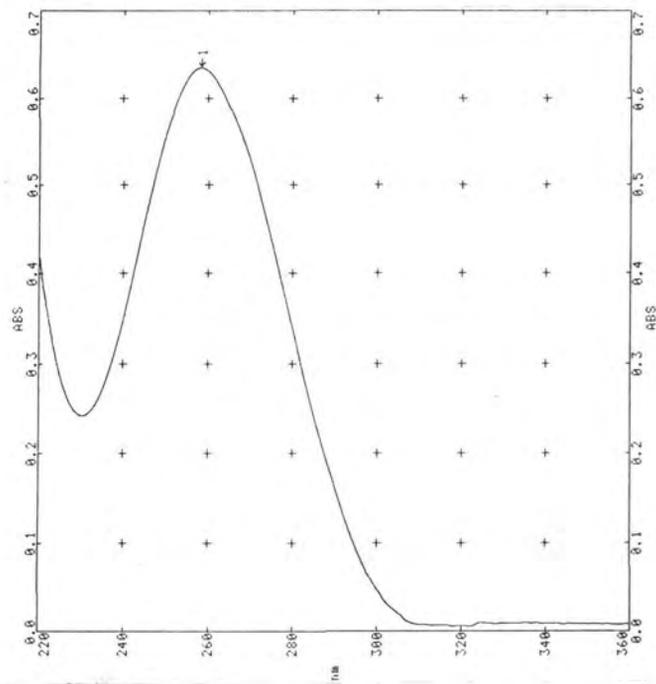
ligated to PUC18 DNA which had been restricted with an appropriate restriction enzyme (as described in section 3.5). Total RNA was extracted from pea cotyledons using a hot SDS method as described in section 3.2.1; Figure 2 shows the quality and quantity of total RNA from 18DAF and 24,26,28DAF cotyledons. Poly(A)⁺RNA was isolated by oligo(dT)-cellulose spun column chromatography, its concentration estimated by spectrophotometry and used for cDNA synthesis (see Fig.2). Typical yields in cDNA synthesis reactions from this mRNA were 2ug cDNA per ug 18DAF poly(A)⁺RNA, and ca.0.45ug cDNA per ug of 24,26,28 DAF poly(A)⁺RNA. As large quantities of poly(A)⁺RNA from desiccating cotyledons were difficult to obtain, cDNA synthesised from 18DAF poly(A)⁺RNA was used to develop an effective cloning procedure. After purification, 18DAF cDNA was ligated to EcoRI/NotI adaptors, which were then kinased with [γ -³²P]-ATP. Analysis of the fractions collected from a Sepharose column used to separate the free kinased adaptors from the adapted cDNA showed two peaks of radioactivity, confirming ligation of the adaptors to the cDNA. However, the amount of radioactivity associated with the cDNA indicated that only a small portion of the cDNA had been adapted. Subsequent ligation of the purified adapted cDNA into EcoRI-restricted PUC18 and transformation of competent cells (DH5alphaME) gave no white transformed colonies.

18DAF cDNA was end-labelled with [α -³²P]-dATP and Klenow enzyme then ligated to phosphorylated EcoRI 8mer

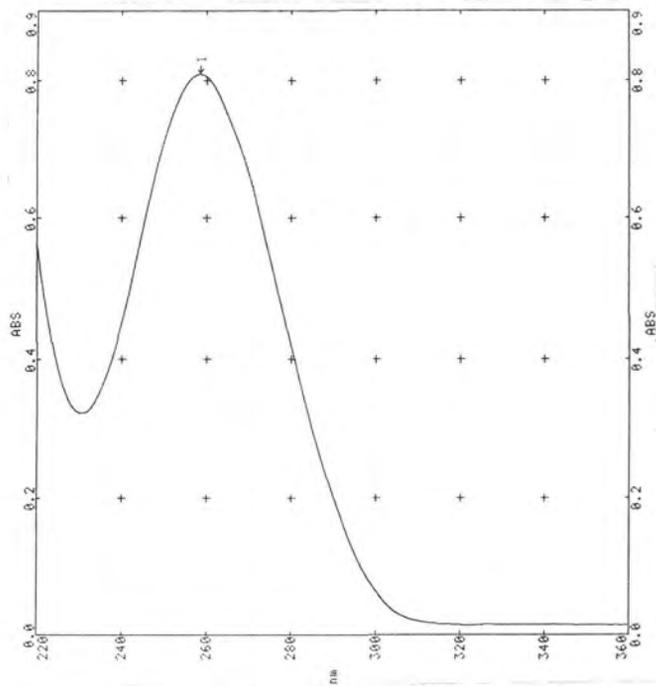
Figure 2. cDNA synthesis from 18DAF and 24,26,28DAF RNA isolated from pea cotyledons for use in the construction of a PUC18 cDNA library.

- a) Spectrophotometric scan of total RNA extracted from 18DAF pea cotyledons; the yield was 1.09mg RNA/g tissue.
- b) Spectrophotometric scan of total RNA extracted from 24,26,28DAF pea cotyledons; the yield was 0.43mg RNA/g tissue.
- c) Formaldehyde gel electrophoresis of 18DAF (track 1) and 24,26,28DAF (track 2) pea cotyledon total RNA.
- d) Spectrophotometric scan of poly(A)⁺RNA isolated from 18DAF pea cotyledon total RNA; the yield was 15ug poly(A)⁺RNA/mg total RNA.
- e) Spectrophotometric scan of poly(A)⁺RNA isolated from 24,26,28DAF pea cotyledon total RNA; the yield was 3.75ug poly(A)⁺RNA/mg total RNA.
- f) Double stranded cDNA synthesised from 18DAF pea cotyledon poly(A)⁺RNA:
Track 1 - EcoRI/HindIII-restricted lambda DNA size marker
(21226, 5148, 4973, 3530, 2027, 1904, 1709, 1375, 947, 831, 564bp)
Track 2 - 18DAF pea cotyledon cDNA (1ug)
Track 3 - 18DAF pea cotyledon cDNA (0.5ug)
- g) Double stranded cDNA synthesised from 24,26,28DAF pea cotyledon poly(A)⁺RNA:
Track 1 - EcoRI/HindIII-restricted lambda DNA size marker
Track 2 - 24,26,28DAF pea cotyledon cDNA (0.5ug)

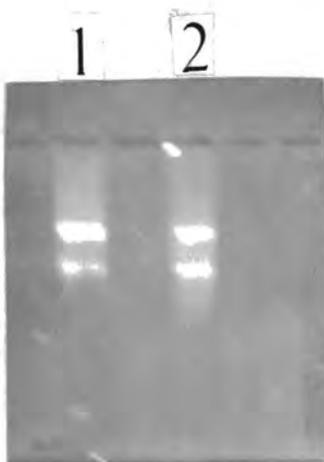
a)



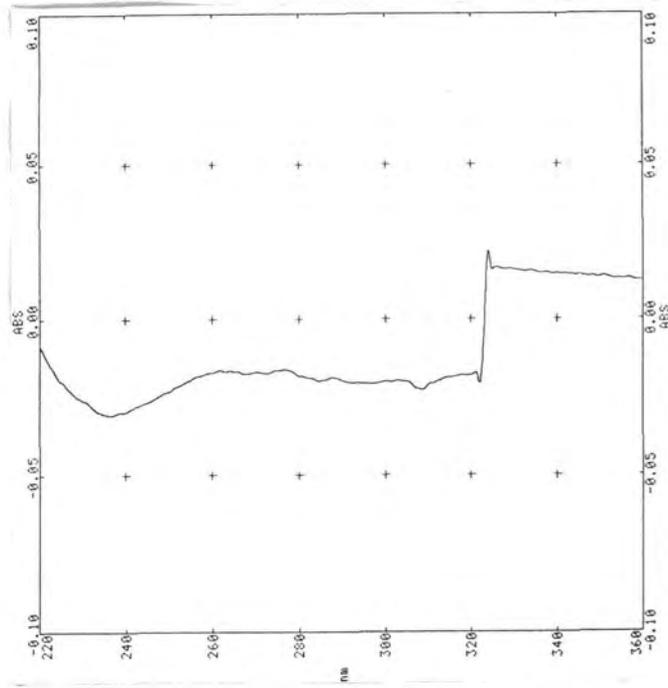
b)



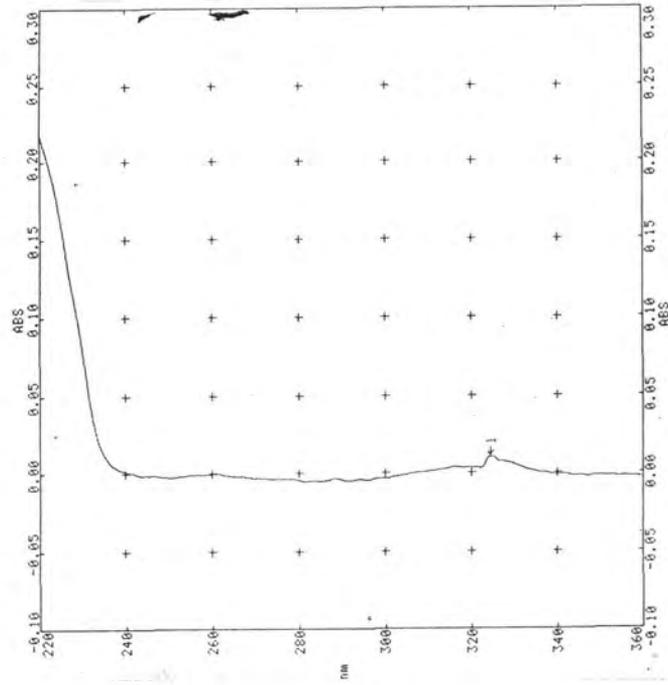
c)



d)



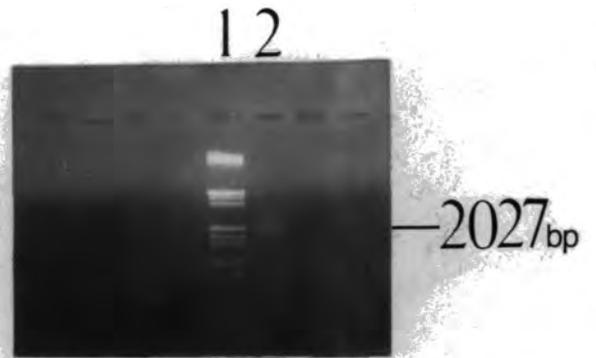
e)



f) 3530



g)



linkers. The constructs were restricted with EcoRI and the linker cDNA was separated from linker fragments on a Sepharose column, then ligated into both EcoRI-restricted PUC18 and EcoRI-restricted-dephosphorylated PUC18 in two separate ligation reactions. Transformation of both DH5alphaME and TBI1 competent cells with both ligation mixes gave no white transformants. Control experiments demonstrated firstly that the EcoRI linkers were functioning as specified and could be ligated to each other, kinased and restricted successfully, and secondly that the transformation of both DH5alphaME and TBI1 competent cells with untreated PUC18 was successful, as was transformation with PUC18 restricted with HincII (to give blunt ends) and then religated, although the efficiency of this transformation was substantially reduced compared to that with untreated PUC18.

The ligation of end-labelled 18DAF cDNA to BamHI 12mer linkers, restriction of the constructs with BamHI and ligation of the isolated and purified linker cDNA to BamHI-restricted PUC18 gave recombinant plasmids which transformed DH5alphaME cells to give a large proportion of white transformed colonies. Preparation and restriction analysis of the recombinant plasmids from a number of these transformed colonies demonstrated the presence of inserts from ca.800 to 2500bp, which were excised from the PUC18 DNA by BamHI.

This cloning procedure was repeated with 24,26,28DAF cDNA (Fig.2), synthesised from poly(A)⁺RNA

isolated by spun column chromatography, using the same BamHI 12mer linkers and BamHI-restricted PUC18. When the final ligation mix was used to transform DH5alphaME cells, no white transformed colonies were obtained. Repetition of this cloning experiment with further batches of newly synthesised 24,26,28DAF cDNA from this and a further preparation of 24,26,28DAF poly(A)⁺RNA gave no recombinant plasmids with the ability to transform the competent cells, and cloning experiments using EcoRI linkers were also unsuccessful. At this point it was considered unproductive to continue attempts to construct a cDNA library from 24,26,28DAF cDNA in PUC18.

4.1.2 The Construction of a cDNA Library from Desiccating Pea Cotyledons using Lambda-gt10 as a Cloning Vector

A cDNA library was constructed in a lambda-gt10 cloning vector using the Proclone kit (Promega). Total RNA was extracted from 24,26,28DAF cotyledons and 24,26,28DAF cDNA was synthesised from poly(A)⁺RNA isolated by spun column chromatography (see Fig.3). The cDNA was separated from free nucleotides on a Sephadex column, purified and ligated to EcoRI/NotI adaptors. The adaptors were kinased with [³²P]-ATP to allow the separation of adapted cDNA from free kinased adaptors on a Sepharose column. After purification, portions of the adapted cDNA were ligated to 0.5ug aliquots of the provided EcoRI-restricted lambda arms; these ligation mixes were packaged using the Packagene system (Promega),

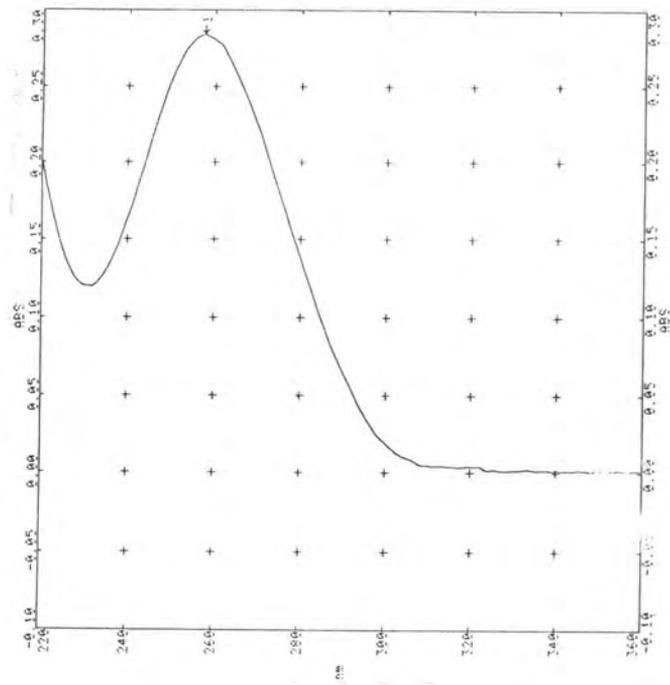
Figure 3. cDNA synthesis from 24,26,28DAF RNA isolated from pea cotyledons for use in the construction of the first lambda cDNA library.

- a) Spectrophotometric scan of total RNA extracted from 24,26,28DAF pea cotyledons; the yield was 0.5mg RNA/g tissue.
- b) Spectrophotometric scan of poly(A)⁺RNA isolated from 24,26,28DAF pea cotyledon total RNA; the yield was 4ug/mg total RNA.
- c) Formaldehyde gel electrophoresis of 24,26,28DAF pea cotyledon total RNA.
- d) Double stranded cDNA synthesised from 24,26,28DAF poly(A)⁺ RNA:

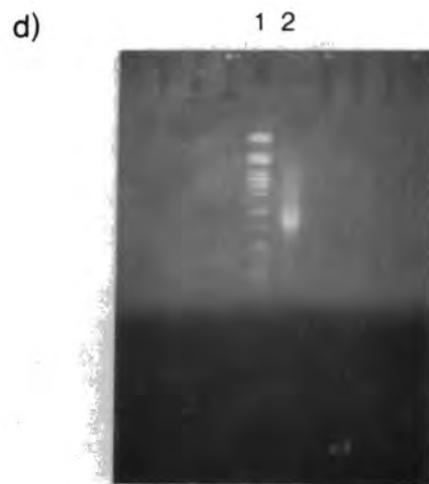
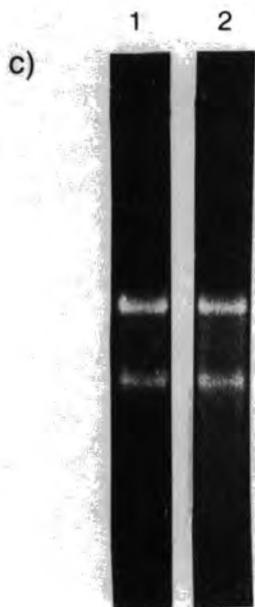
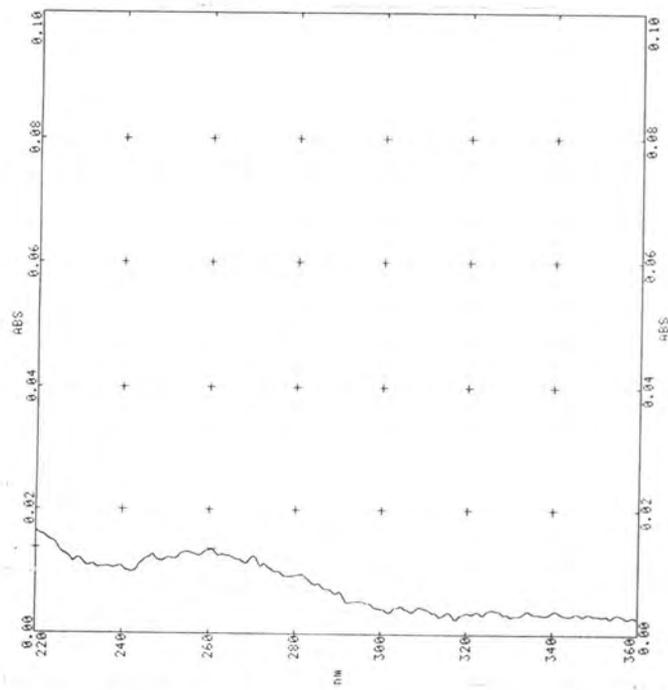
Track 1 - PstI-restricted lambda DNA size marker (11501, 5077, 4749, 4507, 2838, 2556, 2459, 2443, 2140, 1986, 1700, 1159, 1093, 805, 514, 468, 448, 339, 264, 247, 216, 211, 200, 164, 150, 94, 87, 72, 15bp)

Track 2 - 24,26,28DAF cDNA (1ug)

a)



b)

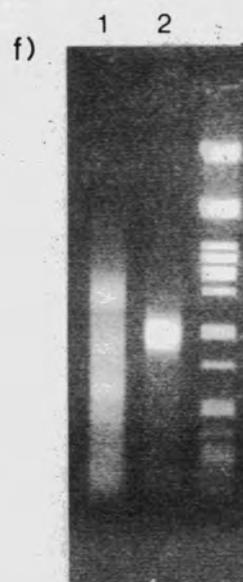
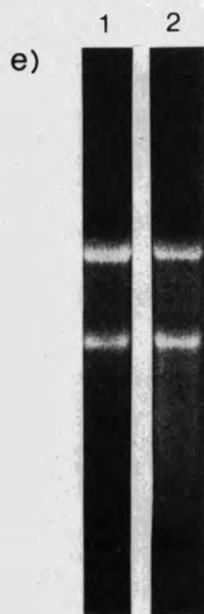
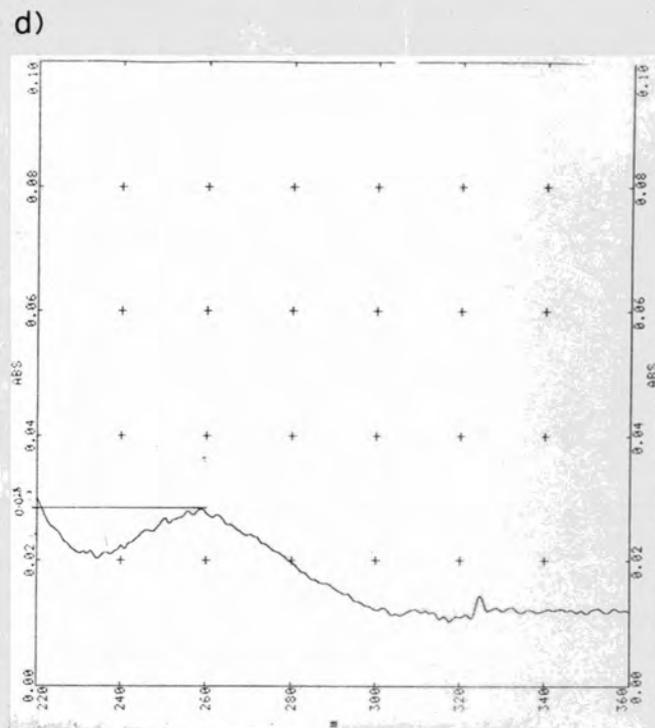
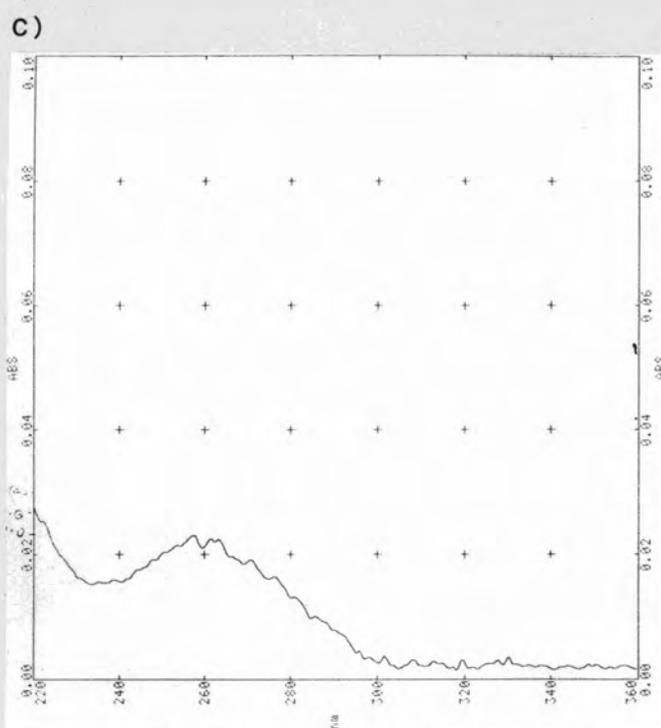
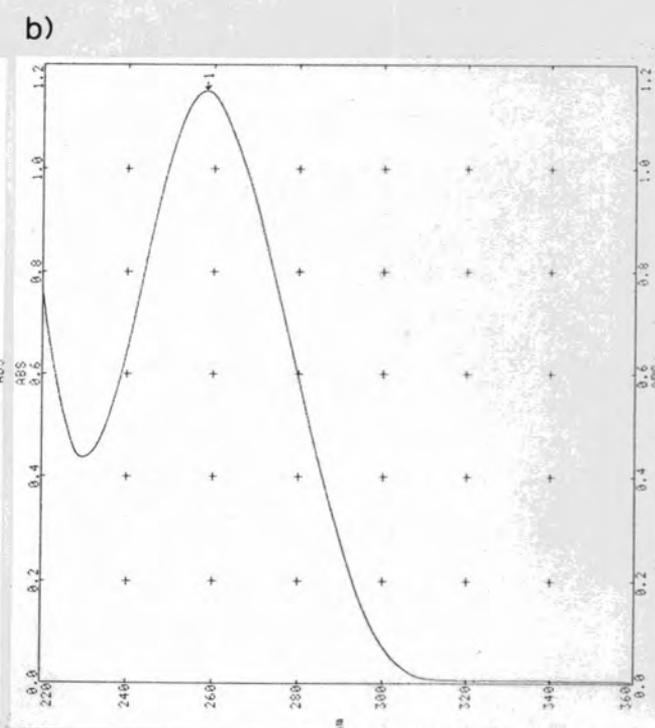
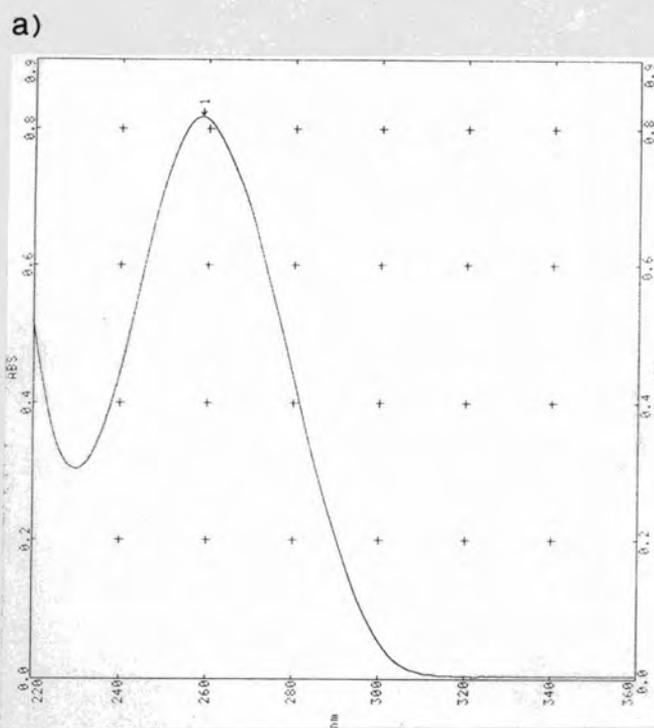


along with a control ligation of vector DNA with no cDNA. Dilutions of the packaged phage were adsorbed to C600 bacteria which were plated and grown overnight to determine both the optimum ratio of insert:vector DNA for cloning and the level of background ligation, indicated by the control. The remaining vector DNA was then ligated to the adapted cDNA, the recombinant lambda DNA was packaged and dilutions adsorbed to C600 cells, plated and grown overnight to estimate the size of the library. The results from these plates indicated that a library of approximately 23,500 clones had been obtained; this figure was lower than could be expected in an ordinary library construction. The kit protocol suggests that the number of transformed plaques (clear) should be 100-1000 fold greater than the number of untransformed (turbid) plaques; in fact the yield of transformed plaques was only 10-30 times greater than untransformed plaques. Nevertheless this library was considered to be sufficient for the investigation of fairly abundant messages. The library was amplified and stored at a concentration of 3.4×10^{11} pfu/ml in aliquots at 4°C with chloroform and at -80°C.

Upon the advent of a new method of poly(A)⁺RNA isolation using magnetic oligospheres, the construction of a second larger cDNA library from desiccating cotyledons was attempted using 24,26,28DAF cDNA synthesised from poly(A)⁺ RNA isolated using the PolyATtract mRNA isolation kit (Promega; see Fig.4). The 24,26,28DAF poly(A)⁺RNA isolated

Figure 4. First strand cDNA synthesis from 18DAF and 24,26,28DAF RNA from pea cotyledons for use in the differential screening of cDNA libraries and for the synthesis of double stranded cDNA for the construction of the second 24,26,28DAF lambda cDNA library.

- a) Spectrophotometric scan of total RNA extracted from 18DAF pea cotyledons; the yield was 1.13mg RNA/g tissue.
- b) Spectrophotometric scan of total RNA extracted from 24,26,28DAF pea cotyledons; the yield was 0.53mg RNA/g tissue.
- c) Formaldehyde gel electrophoresis of 18DAF (track 1) and 24,26,28DAF (track 2) pea cotyledon total RNA.
- d) Spectrophotometric scan of poly(A)⁺RNA isolated from 18DAF pea cotyledon total RNA; the yield was 13.75ug poly(A)⁺RNA /mg total RNA, using the PolyATtract mRNA isolation kit.
- e) Spectrophotometric scan of poly(A)⁺RNA isolated from 24,26,28DAF pea cotyledon total RNA; the yield was 6.25ug poly(A)⁺RNA/mg total RNA, using the PolyATtract mRNA isolation kit.
- f) Single stranded cDNA synthesised from pea cotyledon poly(A)⁺RNA:
Track 1 - 18DAF pea cotyledon cDNA (2ug)
Track 2 - 24,26,28DAF pea cotyledon cDNA (1.5ug)
Track 3 - PstI-restricted lambda DNA size marker



using this kit gave typical yields of 1ug cDNA per ug 24,26, 28DAF poly(A)⁺RNA. The purified cDNA was adapted and cloned into lambda vectors as above using a second lambda gt-10 Protoclone kit. Growth and analysis of the recombinant phage particles from small scale ligations of the adapted cDNA to prepared lambda arms demonstrated successful ligation, the optimum ratio giving approximately 15,500 clones from 0.5ug cDNA. In the packaging of these small-scale ligations the addition of chloroform to the packaged phage particles before titring (as indicated by the Packagene protocol) was omitted. Chloroform was added in the large scale reaction and adsorption and growth of phage particles from this reaction gave no plaques, although a bacterial lawn was obtained. A repeat of this experiment using the remaining prepared lambda arms produced the same result. Consequently a fairly small library was obtained by pooling and packaging the recombinant DNA in the remaining test ligations, giving approximately 26,000 clones. The library was amplified and stored in aliquots at 4°C (with fresh chloroform) and at -80°C, at a concentration of 2.2×10^9 pfu/ml.

4.1.3 Screening of the Lambda-cDNA Libraries with a Probe for *Leg J*

To assess the quality of the 24,26,28DAF cDNA libraries, they were screened with a subclone of the *Leg J* gene (PJC5.2) which is known to be present in pea cotyledons

at this developmental stage (see section 1.4.1).

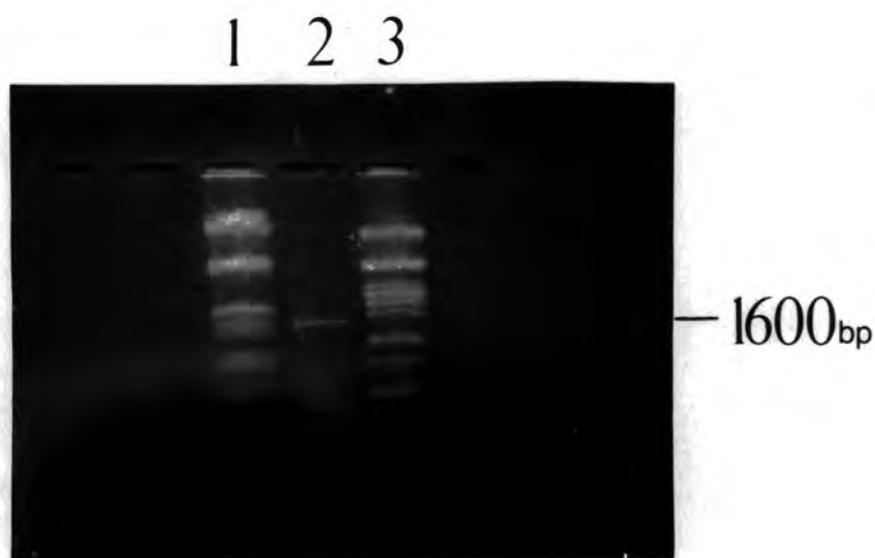
Hybridisation of the radiolabelled PJC5.2 insert to the first lambda-cDNA library immobilised on nitrocellulose filters was at 65°C with washing conditions of 2xSSC, 0.5%SDS at 65°C, as described in section 3.6. In a primary screen of ca.53,000 plaques (2.26x library) the probe hybridised to 52 plaques, giving signals of varying strengths. Agar plugs corresponding to some of the areas of strongest hybridisation were isolated and the phage particles eluted; four positive plaques were isolated from a secondary screen of these phage solutions. The recombinant phage particles from these plaques were eluted, grown up in liquid lysates and their DNA prepared by the Qiagen lambda midiprep system (see section 3.9.1.1). Although this method demonstrated the presence of inserts of over 1kb, very small yields were obtained. To obtain larger quantities of cloned DNA the cDNA inserts were amplified by PCR technology using primers (288 and 289) matching the lambda arms on either side of the cloning site (see section 2.5), thus obtaining cloned cDNA for analysis and subcloning without the surrounding lambda DNA. Using this technique, amplified inserts of ca.1750, 1600 (see Fig.5), 800 and 550bp were obtained and Southern blot and hybridisation analysis demonstrated that all four inserts hybridised strongly to the *Leg J* probe. When the 800bp insert was subcloned into PUC18 using the method of Jung, Pestka and Pestka (1990, section 3.9.3) and sequenced by direct plasmid sequencing, the nucleotide sequence showed

Figure 5. PCR amplification product of a clone isolated from a screen of the first 24,26,28DAF cDNA library with a probe for *Leg J* (PJC5.2).

Track 1: EcoRI/HindIII-restricted lambda DNA size marker*

Track 2: PCR amplified *Leg J* cDNA insert

Track 3: PstI-restricted lambda DNA size marker**.



* 12216, 5148, 4973, 4268, 3530, 2027, 1904, 1709, 1375, 947, 831, 564, 140bp.

** 11501, 5077, 4749, 4507, 2838, 2556, 2459, 2443, 2140, 1986, 1700, 1159, 1093, 805, 514, 468, 448, 339, 264, 247, 216, 211, 200, 164, 150, 94, 87, 72, 15bp.

99.4% homology to the published sequence of a *Leg J* genomic clone Lambda JC5 (Gatehouse *et al*, 1988); the clone ran from position +1123 in the genomic clone to the poly(A) tail, including a spliced intron site. The presence of an intact EcoRI/NotI adaptor at the part of the subclone homologous to +1123 indicates that this clone is a partial cDNA, but the 1750bp fragment isolated from the same screen is the right size to correspond to the full length *Leg J* cDNA.

The second larger 24,26,28DAF cDNA library was screened with the PJC5.2 insert as above to isolate cDNA clones with homology to the *Leg J* sequence. A clear primary screen of ca.85,500 immobilised plaques (3.29x library) was obtained, with the probe hybridising strongly to 20 plaques; this indicated that the *Leg J*, or *Leg J* subfamily, messages formed 0.007% of the total messages in this library, compared to an apparent frequency of 0.04% in the first smaller cDNA library, suggesting that a more extensive library had been obtained. However, isolation of some of the positive plaques, amplification of the cDNA inserts by PCR and analysis by Southern blotting and hybridisation failed to identify any full length cDNAs, the largest isolated insert showing homology to the *Leg J* probe being only 900bp long, with the majority of clones being under 500bp in length.

4.1.4 Screening of Lambda-cDNA Libraries with Probes for Seed and Putative Stress Protein Messages

To attempt to assess the expression of stress

proteins in desiccating pea cotyledons the 24,26,28DAF lambda-cDNA libraries were screened with subclones of the messages from several proteins which have been observed to respond to stress in the whole plant and in seeds of other plant species, to ascertain whether any such proteins were specifically expressed during late embryogenesis for the protection of the seed against water stress and environmental damage.

4.1.4.1 Cowpea Trypsin Inhibitor Protein

The 24,26,28DAF lambda-cDNA libraries were screened for the presence of a message showing homology to the cowpea trypsin inhibitor protein using a subclone of the *CpTI* gene (PASD1) as a probe. The radiolabelled PASD1 insert failed to hybridise to the immobilised first lambda-cDNA library under hybridisation conditions of 65°C and washing conditions of 2xSSC, 0.5% SDS at 65°C. A further hybridisation experiment where the filters were washed at a lower stringency with 5xSSC and 0.1% SDS at 55°C also gave no hybridisation of the probe to the cloned cDNA from desiccating pea cotyledons. A screen of the second larger 24,26,28DAF cDNA library, using a low stringency wash as above, also gave no evidence of homology between the *CpTI* subclone and any of the library cDNA clones.

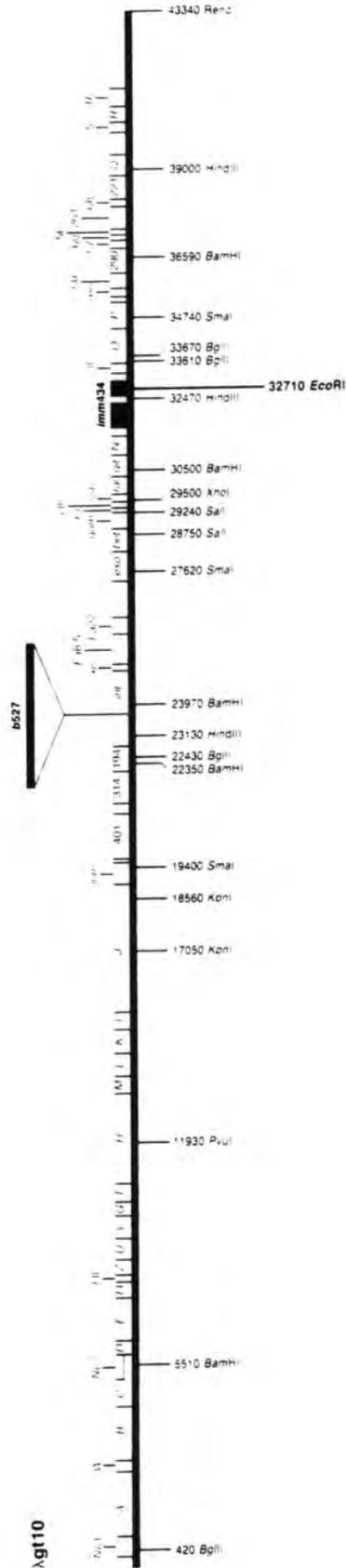
4.1.4.2. *PsMT_A* - A Pea Putative Metallothionein

The second 24,26,28DAF cDNA library was screened with a probe for a pea message that has sequence similarities to class I metallothionein genes and whose protein is known to possess the ability to bind metals, to determine whether the message for this gene was expressed during seed desiccation. With hybridisation occurring at 65°C and washing conditions of 2xSSC, 0.5%SDS at 65°C, no strong hybridisation signals were observed on the library filters after several days exposure; however, a large amount of either background hybridisation or very weak hybridisation was observed. Plaques corresponding to several of these signals were isolated, their phage particles eluted and the DNA prepared by the lambda miniprep procedure (see section 3.9.1.2). Restriction of the resultant DNA and analysis by Southern blotting and hybridisation showed no homology between the isolated clones and the *PsMT_A* probe. A later screen of the first smaller cDNA library proved similarly negative.

4.1.4.3. Wheat Em Late Embryogenesis Abundant Polypeptide

A subclone of the wheat Em cDNA sequence (PMuge Kla) was used to screen the lambda cDNA libraries made from desiccating pea cotyledons to attempt to isolate a pea cDNA clone showing homology to the Em message, which codes for an amino acid sequence that is highly conserved between group 1

Figure 6. Restriction map of lambda-gt10. (taken from Sambrook, Fritsch and Maniatis, 1989).



LEA proteins isolated from a number of species (see section 1.3). Hybridisation of the radiolabelled probe to the second immobilised cDNA library was at 65°C, with washing conditions of fairly low stringency (5xSSC, 0.1% SDS at 65°C) to allow for the heterologous nature of the probe. Hybridisation to several points on the filter was observed, some of which retained the bound probe after the filters were rewashed with 2xSSC, 0.5% SDS at 65°C. Four agar plugs corresponding to the areas of strongest hybridisation were taken, the phage particles eluted and screened again using the same conditions to isolate single positive plaques, whose phage particles were again eluted and the recombinant DNA prepared by the lambda miniprep procedure.

Despite the use of EcoRI/NotI adaptors in the cloning procedure, restriction of the recombinant lambda DNA and analysis by Southern blotting and hybridisation to the PMuge K1a insert probe showed that EcoRI failed to excise the cloned insert from the lambda DNA and that only one EcoRI site appeared to be functional in the recombinant DNA, releasing the 10.63kb lambda arm and leaving the cDNA insert attached to the 32.71kb arm. The insert was excised by restriction of the DNA with EcoRI and HindIII, giving a restriction fragment of the cDNA insert attached to 240bp of lambda DNA (see Fig.6), and subcloned into PUC18. The Em probe showed clear hybridisation to the transformed colonies and to restrictions of the prepared recombinant plasmid DNA, with all hybridisation taking place at 65°C and washing

conditions of 5xSSC, 0.1% SDS. All of the cloned inserts were small, from 300 to 350bp, including the 240bp of lambda DNA. Sequencing of a 340bp insert by direct plasmid sequencing showed that the insert consisted of the 240bp lambda sequence, the first G residue of the EcoRI/NotI adaptor and then a long poly(A) tail of 104bp; sequencing of the other subclones isolated from this library gave the same results.

The smaller 24,26,28DAF cDNA library was also screened with the probe for the wheat Em message, by the same method as before, and also gave clear primary and secondary library screens. Restriction and hybridisation analysis of the resulting lambda DNA minipreps showed that the samples contained inserts of ca.120bp which were excised by EcoRI and which hybridised strongly to the Em probe on a Southern blot washed with 5xSSC, 0.1%SDS at 65°C. Attempts to subclone these EcoRI restriction fragments into PUC18 failed. However, the subcloning of a BglII/HindIII restriction fragment from the lambda minipreps containing 240bp lambda DNA, ca.120bp cDNA insert and another 900bp lambda DNA was successful for both of the two lambda minipreps which showed strong hybridisation to the Em probe. The subcloned DNA from the two recombinant plasmids was sequenced by direct plasmid sequencing and showed in both cases the 240bp lambda DNA, an intact EcoRI/NotI site and a long poly(A) tail of over 80bp, beyond which the sequencer would not read. The reverse sequence identified the 900bp fragment of lambda DNA which

contained no restriction sites. As it was clear that all of the sequences showing strong hybridisation to the Em probe consisted mainly, if not entirely, of poly(A) and the length of the inserts indicated that no coding sequence could be present, the search in the 24,26,28DAF cDNA libraries for a pea clone showing coding sequence homology to the wheat Em probe was not taken further.

4.1.4.4 Pea Leaf Polyubiquitin

The screening of the second 24,26,28DAF cDNA library with a clone of a pea leaf polyubiquitin cDNA (PCU1) gave very clear hybridisation to DNA from several immobilised plaques, four of which were isolated by a secondary screen and their DNA prepared by the lambda miniprep technique. As with the putative Em clones isolated from this library, restriction and hybridisation analysis of the recombinant DNA from the polyubiquitin screen demonstrated that the inserts could not be excised with EcoRI. Two of the lambda minipreps showed no restriction with EcoRI; restriction of the other two minipreps showed the presence of one EcoRI site, but Southern blot and hybridisation analysis indicated that the site was within the cDNA insert, releasing the short 10.63kb lambda arm and a small portion of the insert, with the bulk of the insert remaining with the longer 32.71kb arm. Further restriction analysis with EcoRI+HindIII and BglII+HindIII showed that the inserts which hybridised to the PCU1 probe were all small - less than 200bp - and attempts to subclone

both EcoRI/HindIII fragments containing 240bp lambda DNA attached to the insert DNA and BglII/HindIII fragments containing the insert within 240 and 900bp lambda DNA failed to produce any sequence with homology to the polyubiquitin coding sequence or the surrounding lambda primers.

The first cDNA library was also screened with PCU1 and strong hybridisation to several plaques was observed. Analysis of the resultant recombinant lambda DNA gave four clones showing strong hybridisation to the PCU1 probe with apparently intact EcoRI cloning sites (Fig.7). The DNA was restricted and subcloned into PUC18 using both EcoRI and BglII+HindIII, as the full polyubiquitin sequence contains both EcoRI and BglII restriction sites. Five subclones with inserts showing strong hybridisation to the PCU1 probe were sequenced and shown to contain cDNA inserts of 221, 450, 458, 513 and 549bp, each containing part of the ubiquitin coding sequence. These sequences are analysed below.

4.1.5 An Investigation of Pea Leaf and Cotyledon

Ubiquitin Messages

4.1.5.1 Sequence of a Second Polyubiquitin cDNA from Pea Leaves

As well as the fully-sequenced PCU1 polyubiquitin cDNA subclone, F.Z. Watts also provided a second polyubiquitin cDNA isolated from a pea leaf cDNA library, termed PCU2. This subclone, with an insert of 1.05kb, had

Figure 7. Restriction and hybridisation analysis of recombinant lambda DNA from clones isolated from a screen of the first 24,26,28DAF cDNA library with a probe for pea polyubiquitin (PCU1).

a) Lambda DNA minipreps restricted with BglII+HindIII

Track 1: PU1

Track 2: PU2

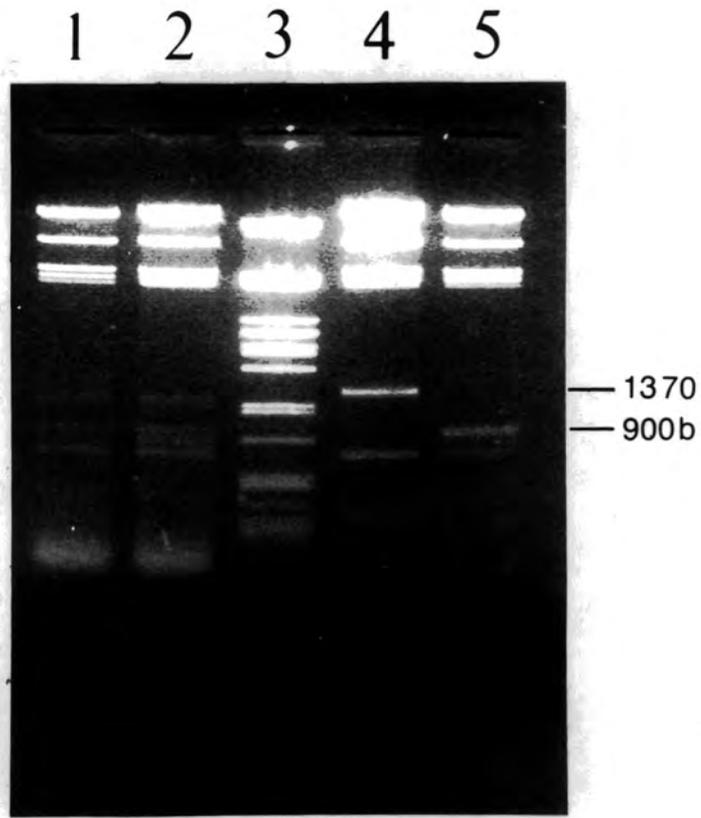
Track 3: PstI-restricted lambda DNA

Track 4: PU3

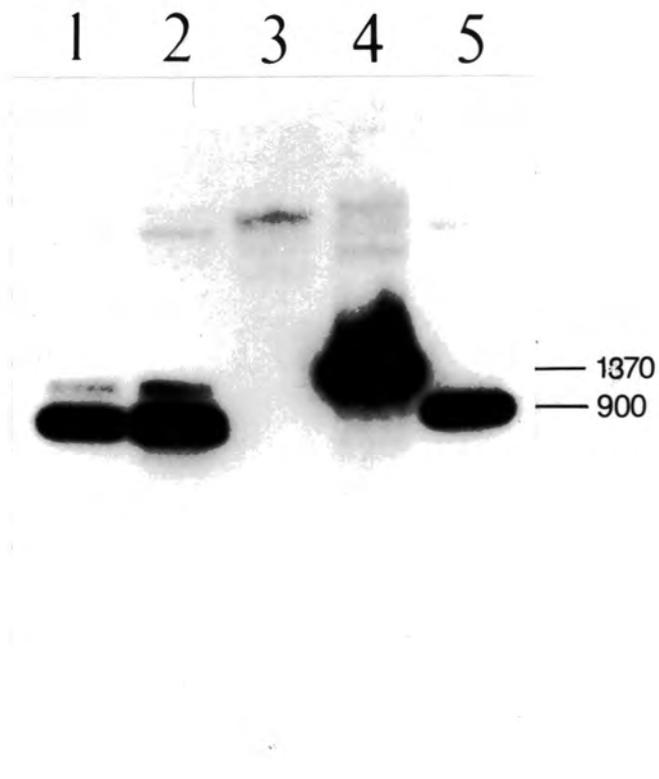
Track 5: PU4;

b) Southern blot analysis of gel a) probed with radiolabelled PCU1 to isolate restriction fragments containing cDNA inserts with homology to pea polyubiquitin.

a)



b)



been partially sequenced and shown to have a slightly different restriction map to PCU1. The entire PCU2 sequence was obtained and confirmed by several subclonings utilising two internal SacI sites (Fig.8). The published sequence of PCU1 has 5 entire tandem repeats, each of 76 amino acids, with 19bp 5' untranslated sequence and 76bp 3' untranslated sequence (Watts and Moore, 1989); the 1050bp sequence of PCU2 consists of 816bp coding sequence, which starts in the middle of a ubiquitin repeat and has 3 further complete repeats, ending in an extra phenylalanine (Phe) residue, as PCU1. It also has 157bp 3' untranslated sequence containing a polyadenylation site AATAAA and a 77bp poly(A) tail. If the two DNA clones are arranged so that their final ubiquitin repeat with the terminal Phe residue are aligned (as in Fig. 8) they show 100% amino acid homology and 81.7% nucleotide homology in the coding sequence. However, the 3' untranslated sequences share only 26.3% homology.

Figure 8. A comparison of the sequences of two polyubiquitin cDNAs isolated from pea leaves (AA = amino acid)

```

PCU1  GAATTCCGATAATTATAAGATGCAAATTTTCGTTAAGACCCTCACCGGCA  50
AA           M  Q  I  F  V  K  T  L  T  G

PCU1  AGACCATCACTCTTGAAGTTGAGAGTTCAGATACCATAGACAATGTTAAG  100
AA  K  T  I  T  L  E  V  E  S  S  D  T  I  D  N  V  K

PCU1  GCCAAGATTCAAGACAAGGAAGGGATCCCTCCTGACCAGCAGCGTCTCAT  150
AA  A  K  I  Q  D  K  E  G  I  P  P  D  Q  Q  R  L  I

PCU1  TTTTGCTGGCAAGCAACTCGAGGATGGGCGTACTCTTGCTGATTACAACA  200
AA  F  A  G  K  Q  L  E  D  G  R  T  L  A  D  Y  N

```

PCU1	TACAGAAGGAGTCCACCCTCCATCTTGTGTTGCGACTTCGTGGTGGGATG	250
AA	I Q K E S T L H L V L R L R G G M	
PCU1	CAAATCTTTGTGAAGACTCTCACTGGAAAGACCATCACCCCTTGAGGTGGA	300
AA	Q I F V K T L T G K T I T L E V E	
PCU2		> 2
PCU1	AGCTCTGATACCATTGACAATGTCAAGGCCAAGATCCAGGACAAAGAGGG	350
AA	S S D T I D M V K A K I Q D K E G	
PCU2	A G G C T G C C A G A T G	52
PCU1	CATTCCCCAGACCAACAAAGGCTCATCTTTGCTGGCAAACAGCTCGAAG	400
AA	I P P D Q Q R L I F A G K Q L E	
PCU2	C C T C C T C G T	102
PCU1	ATGGCAGAACTTTGGCCGATTACAACATCCAGAAAGAATCCACCCTCCAT	450
AA	D G R T L A D Y N I Q K E S T L H	
PCU2	C C T T C A C T C	152
PCU1	CTTGTGTTGCGACTTCGTGGTGGTATGCAGATTTTTGTGAAGACCCTTAC	500
AA	L V L R L R G G M Q I F V K T L T	
PCU2	A C C C C A C T C	202
PCU1	CGGCAAGACAATTACTTTGGAGGTTGAGAGCTCTGACACAATTGACAATG	550
AA	G K T I T L E V E S S D T I D N	
PCU2	A T C A A A C A G T G	252
PCU1	TGAAGGCCAAGATACAGGACAAGGAGGGTATTCCTCCAGACCAGCAAAGG	600
AA	V K A K I Q D K E G I P P D Q Q R	
PCU2	C C G A A G C C C T	302
PCU1	TTGATTTTTGCTGGAAAGCAGCTCGAAGATGGCCGAACTTTGGCCGACTA	650
AA	L I F A G K Q L E D G R T L A D Y	
PCU2	C G T C G CTG	352
PCU1	CAACATTCAGAAAGAATCAACCCTCCATCTTGTTTTTGAGACTTCGTGGTG	700
AA	N I Q K E S T L H L V L R L R G	
PCU2	A C C G T T T T A T C C T	402
PCU1	GTATGCAGATTTTTGTGAAAACCCTCACCGGCAAGACAATTACTTTGGAG	750
AA	G M Q I F V K T L T G K T I T L E	
PCU2	G C C C A	452
PCU1	GTAGAGAGCTCAGACACAATTGACAATGTGAAGGCCAAGATACAGGACAA	800
AA	V E S S D T I D N V K A K I Q D K	
PCU2	A A A T G C C T T	502
PCU1	GGAGGGTATCCCTCCAGACCAGCAAAGGTTGATTTTTGCTGGAAAGCAGC	850
AA	E G I P P D Q Q R L I F A G K Q	
PCU2	G A A G C T G T C T	552
PCU1	TTGAAGATGGCCGAACTTTGGCTGATTACAACATTCAGAAGGAGTCCACC	900
AA	L E D G R T L A D Y N I Q K E S T	

PCU2	C T G G T C T T A C	602
PCU1	CTTCACCTTGTCTTCCTTCGTCTGAGGGGAGGTATGCAGATCTTTGTGAAGAC	950
AA	L H L V L R L R G G M Q I F V K T	
PCU2	C T T C C C G G G C T T T T	652
PCU1	TTTGACAGGAAAGACTATTACCTTAGAGGTAGAAAGTTCAGACACAATCG	1000
AA	L T G K T I T L E V E S S D T I	
PCU2	A G T A T G	702
PCU1	ATAATGTGAAGGCCAAAAATACAGGACAAGGAGGGAATCCCACCAGATCAG	1050
AA	D N V K A K I Q D K E G I P P D Q	
PCU2	C G C	752
PCU1	CAGAGGTTGATCTTTGCTGGGAAGCAATTGGAGGATGGAAGGACTCTTGC	1100
AA	Q R L I F A G K Q L E D G R T L A	
PCU2	AG C GC C	802
PCU1	TGACTACAACATTCAGAAGGAGTCTACTCTTCACCTTGTTTTGCGTCTTC	1150
AA	D Y N I Q K E S T L H L V L R L	
PCU2	G GCTGTTTGTGTTATGTTGTGTTTGTAAATCTTGGTTA	852
PCU1	GTGGTGGATTTTAAGAATTGTATTGGGACCGAAACCGTAAACCGTTTATA	1200
AA	R G G F *	
PCU2	TTTTGAAATGGAAGTTGTGAACTGATGATCTTTTTGGGTCCTATATTGAG	902
PCU1	GTCTTTGGCTTTTATGTTTCAGAACTTGTATAATGGATCC	1240
PCU2	GCCCCCTAGAATAAATTCGATTTTATGTACTGTTATCATATGGTTGTGAA	952
PCU2	TAATTAGTATTTTGCCTTTGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1002
PCU2	AA	1050

4.1.5.2 Sequences of Ubiquitin cDNA Clones Isolated from Desiccating Pea Cotyledons

When analysed, four of the sequences obtained from the five subclones isolated from the first cDNA library (derived from lambda minipreps PU1 and PU3, Fig.7) could be combined to produce a confirmed 550bp sequence (PCU3, Fig.9) of a single partial ubiquitin coding sequence (without the first two codons, which were lost through the use of BglII

in the subcloning procedure), containing 222bp of coding sequence (74 amino acids). This ubiquitin repeat was followed by a C-terminal extension of 52 amino acids before a stop site, 141bp 3' untranslated sequence and an 18b poly(A) tail. The fifth 583bp insert (derived from lambda miniprep PU4, Fig.7) also contained 222bp of typical ubiquitin coding sequence, beginning and ending at the same points as PCU3, again followed by a C-terminal extension of 79 amino acids (showing no significant homology to the extension sequence in PCU3) then a stop site, 109bp 3' flanking sequence and a 13b poly(A) tail (PCU4, Fig.9). This sequence analysis identifies the cDNAs isolated from desiccating pea tissue as coding for the two types of ubiquitin extension proteins which have previously been shown in other organisms to be involved in ribosome biogenesis, the 'tail' proteins forming components of the large and small ribosomes (52-53 residue and 76-81 residue tails respectively).

Figure 9. A comparison of the sequences of two ubiquitin extension protein cDNAs isolated from desiccating pea cotyledons.

PCU3	ATCTTCGTGAAAACCCTAACCGGGAAAACCATCACCTCGAGGTTGAGAG	50
PCU4	C G G C TC	
AA	I F V K T L T G K T I T L E V E S	
PCU3	CAGCGACACCATCGACAATGTCAAGGCCAAGATTCAAGACAAGGAAGGTA	100
PCU4	TTC A A C T A	
AA	S D T I D N V K A K I Q D K E G	

PCU3	TTCCTCCGGATCAGCAGCGTTTGATTTTCGCCGGTAAGCAGCTTGAAGAT	150
PCU4	C A C A AC C T C A A C	
AA	I P P D Q Q R L I F A G K Q L E D	
PCU3	GGACGCACCTTGGCCGATTACAATATCCAGAAGGAATCTACTCTTCATCT	200
PCU4	C TC T C C C C C C	
AA	G R T L A D Y N I Q K E S T L H L	
AA	I I E P S L M A L	
PCU3	TGTTTCTCAGGCTTCGCGGTGGAATCATTGAACCTTCTTTGATGGCTTTGG	250
PCU4	C T C GCCAAGAAGCGTAAGAAGAAGACCTACA	
AA	V L R L R G G A K K R K K K T Y	
AA	A R K Y N Q D K T I C R K C Y A R	
PCU3	CTAGGAAATACAACCAAGACAAAACCATTTGCCGCAAATGCTATGCTCGT	300
PCU4	CCAAGCCGAAGAAGATTAAGCACAAGCATAGGAAGGTGAAGCTAGCTGTT	
AA	T K P K K I K H K H R K V K L A V	
AA	L H P R A V N C R K K K C G H S N	
PCU3	CTTCATCCCAGGGCTGTAACTGCCGCAAGAAGAAGTGTGGACACAGCAA	350
PCU4	CTTCAGTTTTTACAAGGTCGATGATTCTGGGAAGGTTTCAGAGGTTGAGGAA	
AA	L Q F Y K V D D S G K V Q R L R K	
AA	Q L R P K K K I K *	
PCU3	TCAATTGAGGCCAAAGAAGAAGATCAAGTAGATTGTGGTTGATATTGTTT	400
PCU4	GGAGTGCCCGAATGCGGAGTGCGGTGCGGGACTTTTTATGGCGAATCACT	
AA	E C P N A E C G A G T F M A N H	
PCU3	CTCTTTTGCTGTTTGTCTGAGATACTTTTCATTGTTTTGGTTGTTAATT	450
PCU4	TTGATCGTCACTACTGTGGCAAGTGTGGTTTGACCTATGTTTATCAGAAA	
AA	F D R H Y C G K C G L T Y V Y Q K	
PCU3	CTATCATGACTAAAGATTTTGTTTTTCTTAATGTTTAATGTAGTTGGAGT	500
PCU4	GCTGATGCTTGATTTTTGTATGTTAATTTTATAATTTAGTTTTTGAGA	
AA	A D A *	
PCU3	TTAATTTGGTATAATCATTTTGTTATGACATTAAAAAAAAAAAAAAAAAAAAA	550
PCU4	TTGTTTAGATTTTGGTATTTGGTGACATTTGTATTTTCATGCCATGATGGT	
PCU4	TAAATAACCGGTAGTTTCCTTAAAAAAAAAAAAAA	583

The ubiquitin part of these cDNAs show 100% amino acid homology to each other and to the published plant polyubiquitin sequence. Within the ubiquitin coding sequence, PCU3 and PCU4 share 82% nucleotide homology. Like the other ubiquitin extension tail proteins examined (see

Figs.10 and 11), the tails of PCU3 and PCU4 are highly basic, containing ca.30% and 28% lysine and arginine residues respectively. Both tails contain a consensus sequence for nuclear localisation (PCU3: Arg-Pro-Lys-Lys-Lys-Ile-Lys; PCU4 : Lys-Lys-Arg-Lys-Lys-Lys); in both cases the position of this sequence is conserved in the tail sequences from other species. Similarly, there is 100% conservation of the position of the cysteine residues involved in the putative zinc-binding, nucleic acid-binding domain in the 76-81 residue tails, giving a motif of the form Cys-X₄-Cys-X₁₄-Cys-X₂-Cys, and almost 100% in the 52-53 residue tails, of the form Cys-X₂-Cys-X₁₀-Cys-X₄-Cys; the only observed exception is tobacco which appears to have only two cysteine residues in common with the other observed sequences followed by an inserted sequence of 20 apparently unrelated amino acids, which interrupts the area of the consensus tail 'zinc finger' domain (Genschik *et al*, 1990). Tables 1 and 2 give the amino acid percentage homologies of the pea tail proteins with those of several other species.

Figure 10. A comparison of the deduced amino acid sequence of the pea 52 residue ubiquitin extension protein tail with 52 and 53 residue tails from a range of other species:

A.t.= *Arabidopsis thaliana*, ref. Callis, Raasch and Vierstra (1990); tobacco, ref. Genschik *et al* (1990); *C.r.*= *Chlamydomonas reinhardtii*, ref. Pollmann, Kampen and Wetter

(1991); *S.c.*= *Saccharomyces cerevisiae*, ref. Ozkaynak et al (1987); human, ref. Baker and Board (1991); *D.m.*= *Drosophila melanogaster*, ref. Cabrera y Poch, Arribas and Izquierdo (1990); *T.p.*= *Tetrahymena pyriformis*, ref. Neves et al (1991). The proposed nuclear localisation signal is underlined and the cysteine residues thought to be involved in the zinc-binding nucleic acid binding domain are highlighted (Ozkaynak et al, 1987). In the tobacco sequence, * denotes the position of 20 extra amino acids (PRRTPQRTCRVLTWIPQARL).

PCU3	Ile	Ile	Glu	Pro	Ser	Leu	Met	Ala	Leu	Ala	Arg	Lys	12
A.t.								Met					
Tobacco							Lys				Ser		
C.r.							Gln						
S.c.							Lys				Ser		
Human							Arg	Gln			Gln		
D.m.							Arg	Ile			Gln		
T.p.	Gly	Met			Thr	Ile	Ala				Lys		

PCU3	Tyr	Asn	Gln	Asp	Lys	Thr	Ile	Cys	Arg	Lys	Cys	Tyr	24
A.t.						Met							
Tobacco	Phe		Cys			Met							
C.r.				Glu		Met							
S.c.			Cys			Ser	Val						
Human			Cys			Met							
D.m.			Cys			Met							
T.p.			Cys	Glu		Lys	Val			Asp			

PCU3	Ala	Arg	Leu	His	Pro	Arg	Ala	Val	Asn	Cys	Arg	Lys	36
A.t.													
Tobacco	Val		Cys*Pro					Thr					
C.r.								Lys					
S.c.				Pro				Thr					
Human													
D.m.								Thr					
T.p.				Pro		Lys		Thr					

PCU3	Lys	Lys	<u>Cys</u>	Gly	His	Ser	Asn	Gln	Leu	<u>Arg</u>	<u>Pro</u>	<u>Lys</u>	48
A.t.								Gly					
Tobacco	Arg					Thr		His	Val				
C.r.		Ser				Thr							
S.c.	Arg					Thr							
Human						Thr		Asn					
D.m.						Thr		Asn					
T.p.	Arg							Ser			<u>Leu</u>		
PCU3	<u>Lys</u>	<u>Lys</u>	<u>Ile</u>	<u>Lys</u>									52
A.t.													
Tobacco			<u>Leu</u>										
C.r.			<u>Leu</u>										
S.c.			<u>Leu</u>										
Human			<u>Val</u>										
D.m.			<u>Leu</u>										
T.p.			<u>Pro</u>		Glu								53

Table 1. Percentage amino acid sequence homologies between the deduced pea 52 amino acid ubiquitin extension protein tail sequence and 52 and 53 residue tail sequences from other plants, animals, yeast and micro-organisms.

Abbreviations and references as in Fig.10; percentages are calculated using the number of amino acids in the tail sequences of the species listed vertically and are rounded up to 0 decimal places.

	PCU3	A.t.	C.r.	S.c.	Human	D.m.	T.p.
PCU3 (plant)	--	94	87	81	85	83	65
A.t. (plant)	94	--	85	77	87	85	63
C.r. (alga)	87	85	--	81	83	85	63
S.c. (yeast)	81	77	81	--	81	85	73
Human (animal)	85	87	83	81	--	94	63
D.m. (animal)	83	85	85	85	94	--	65
T.p. (protozoa)	64	62	62	72	62	64	--

Figure 11. A comparison of the deduced amino acid sequence of the pea 79 residue ubiquitin extension protein tail with 80-81 residue tails from other plants and a 76 residue tail from yeast:

A.t. = *Arabidopsis thaliana*, ref. Callis, Raasch and Vierstra (1990); tomato, ref. Hoffman *et al* (1991); potato, ref. Garbarino, Rockhold and Belknap (1992); *S.c.* = *Saccharomyces cerevisiae*, ref. Ozkaynak *et al* (1987). Markings as in Fig. 10; < denotes end of coding sequence.

PCU4	Ala	<u>Lys</u>	<u>Lys</u>	<u>Arg</u>	<u>Lys</u>	<u>Lys</u>	<u>Lys</u>	Thr	Tyr	Thr	Lys	Pro	12
A.t.													
Tomato													
Potato													
<i>S.c.</i>	Gly							Val			Thr		
PCU4	Lys	Lys	Ile	Lys	His	Lys	His	Arg	Lys	Val	Lys	Leu	24
A.t.								Lys					
Tomato								Lys	Lys				
Potato								Lys	Lys				
<i>S.c.</i>								Lys					
PCU4	Ala	Val	Leu	Gln	Phe	Tyr	Lys	Val	Asp	Asp	Ser	Gly	36
A.t.										Gly			
Tomato											Thr		
Potato											Thr		
<i>S.c.</i>				Ser	Tyr					Ala	Glu		
PCU4	Lys	Val	Gln	Arg	Leu	Arg	Lys	Glu	Cys	Pro	Asn	Ala	48
A.t.													
Tomato													
Potato													
<i>S.c.</i>			Thr	Lys			Arg			Ser		Pro	
PCU4	Glu	Cys	Gly	Ala	Gly	Thr	Phe	Met	Ala	Asn	His	Phe	60
A.t.	Thr									Ser			
Tomato													
Potato													
<i>S.c.</i>	Thr					Val		Leu				Lys	

PCU4	Asp	Arg	His	Tyr	Cys	Gly	Lys	Cys	Gly	Leu	Thr	Tyr	72
A.t.													
Tomato													
Potato													
S.c.			Leu						His	Ser	Val		
PCU4	Val	Tyr	Gln	Lys	Ala	Asp	Ala	<					79
A.t.					Gln	Gly		Gln	Glu	<			81
Tomato			Asn			Gly	Gly	Asp	<				80
Potato			Asn			Gly	Gly	Asp	<				80
S.c.	Lys	Val	Asn	Ala	<								76

Table 2. Percentage amino acid sequence homologies between the deduced pea 79 amino acid ubiquitin extension protein tail sequence and 76-81 residue tail sequences from other plants and yeast. Abbreviations and references as in Fig.11; percentages are calculated using the number of amino acids in the tail sequences of the species listed vertically and are rounded up to 0 decimal places.

	PCU4	A.t.	Tomato	Potato	S.c.
PCU4 (plant)	--	92	92	92	65
A.t. (plant)	90	--	89	89	64
Tomato (plant)	91	88	--	100	65
Potato (plant)	91	88	100	--	65
S.c. (yeast)	67	68	68	68	--

4.1.6 Differential Screening of the Lambda-cDNA Library

To identify late embryogenesis abundant clones which may have a role in protecting the seed from desiccation, it was decided to screen the 24,26,28DAF cDNA library with 18DAF cDNA (at which stage the seed is at maturity but has not begun to desiccate) to isolate those clones which are newly expressed or whose expression is increased upon desiccation of the seed.

4.1.6.1 24,26,28DAF cDNA Library Screened with 18DAF cDNA

Four filters taken from the same plate of the first cDNA library (ca.53,000 clones, 2.26x the library) were screened with 0.7ug radiolabelled total 18DAF cDNA. Although some hybridisation was observed, the autoradiographs were very faint and the plaques were too close together to allow identification of 'negative' plaques - i.e. those not hybridising or hybridising to a very small proportion of the probe. The experiment was repeated using two duplicate filters of a slightly less dense plate (ca.47,000 clones, 2x the library) and 1.4ug cDNA radiolabelled overnight to a high specific activity. The probe hybridised successfully and some small negative areas were observed on both filters. However, the stretching of nitrocellulose filters during treatment and hybridisation, together with the fact that the vast majority of plaques hybridised to the cDNA probe, made it impossible to identify matching spaces on the filters with

any accuracy, or to extend this to the master plate.

4.1.6.2 24,26,28DAF cDNA Library Screened with a Subtracted Probe

In order to identify clones which are differentially expressed between 18 and 24,26,28DAF in pea cotyledons without interference from abundant messages common to both ages a probe of 24,26,28DAF cDNA subtracted with 18DAF poly(A)⁺RNA was constructed. This should allow the identification of library clones whose expression increases or is initiated during seed desiccation.

First strand cDNA synthesised from 24,26,28DAF poly(A)⁺RNA isolated by the Promega PolyATtract system was subtracted by one round of hybridisation with a 30-fold excess (by weight) of 18DAF poly(A)⁺RNA, as described in section 3.7. Measurement of radioactivity in the fractions from the final column, separating the subtracted cDNA from free nucleotides after labelling to high specific activity, demonstrated the presence of a peak of subtracted cDNA (total ca.0.86ug) labelled to high specific activity, which was pooled and used to probe duplicate filters of the lambda-cDNA library (ca.75,000 plaques per filter). After 4 days then 3 weeks of exposure, no clear duplicate points of hybridisation to the probe could be seen.

4.1.6.3 Differential Screening of Selected Clones

Duplicate filters of the lambda-cDNA library were

screened with a radiolabelled probe of the 24,26,28DAF cDNA library inserts amplified by PCR. Autoradiographs were developed after very brief exposures (2 hours and 5 hours) and 50 positive plaques were isolated; the majority of the selected plaques exhibited strong hybridisation to the probe, with a few plaques showing medium and weak hybridisation taken for comparison. The inserts from the recombinant phage isolated from the secondary screens were amplified by PCR. When the PCR products were analysed on gel the results were uncharacteristic of a normal PCR amplification, exhibiting in most cases a long smear of DNA and a diffuse band of fragments of between 200 and 300bp. Amplification of the stock solutions of the selected clones gave more concentrated solutions of phage particles and the PCR amplification procedure was repeated, but gave a similar PCR product (Fig. 12). 80ul of each PCR product from the second amplification was run on gel with size markers and 18DAF cDNA for a positive control (Figs.13a and 13b), and the DNA transferred to nitrocellulose filters by Southern blotting. The blots were probed with radiolabelled 18DAF cDNA (Figs.13c and 13d). Although it is hard to compare the clones with each other, due to large variations in the concentrations of the amplified inserts, when comparing the gels and the autoradiographs it is clear that several clones are hybridising very weakly to the 18DAF cDNA, despite having a high concentration of DNA on the gel. The clones showing the most obvious differential expression are 6, 8, 10, 21,

Figure 12. PCR amplification of stored phage particle solutions.

Track 1: clone 1 isolated from the first 24,26,28DAF cDNA library by screening with a PCR-amplified library probe

Track 2: PstI-restricted lambda DNA size marker

Track 3: PCR amplification of the *Leg J* cDNA clone shown in Fig.5 after 4-5 months storage of the phage solution.

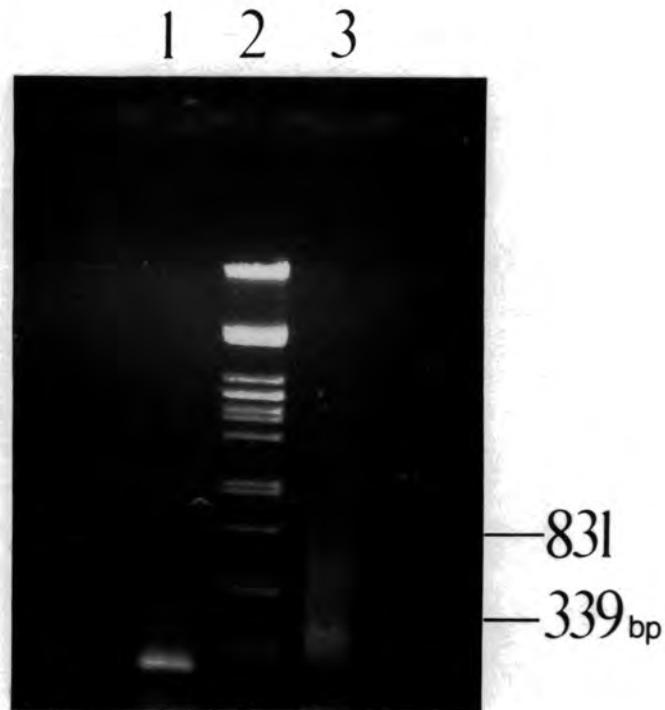


Figure 13. Differential screening by Southern blot hybridisation of 50 selected clones from the first 24,26,28DAF cDNA library amplified by PCR.

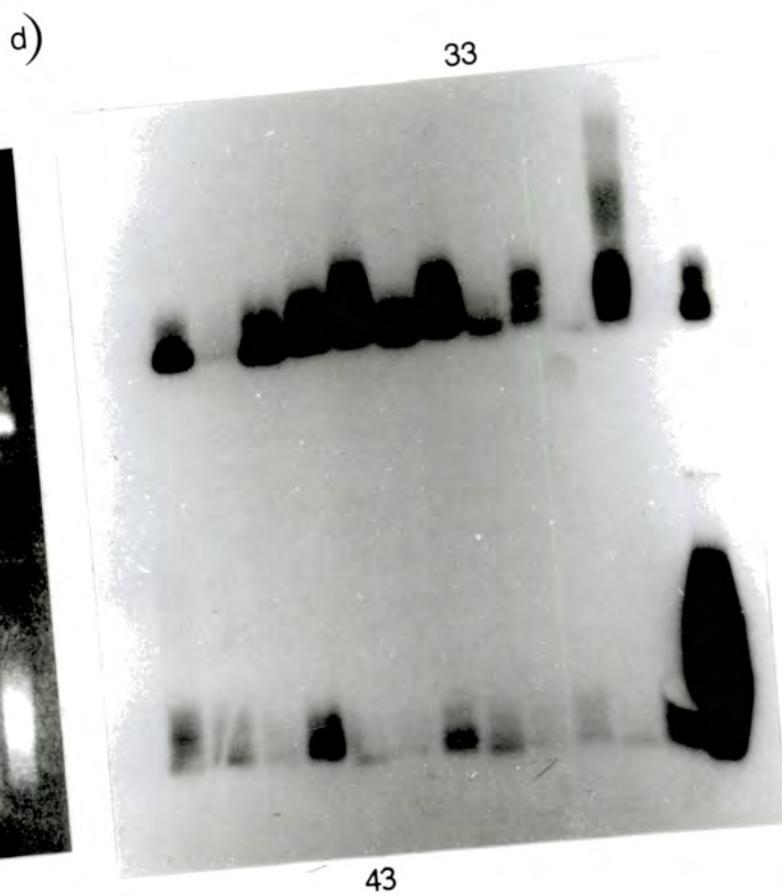
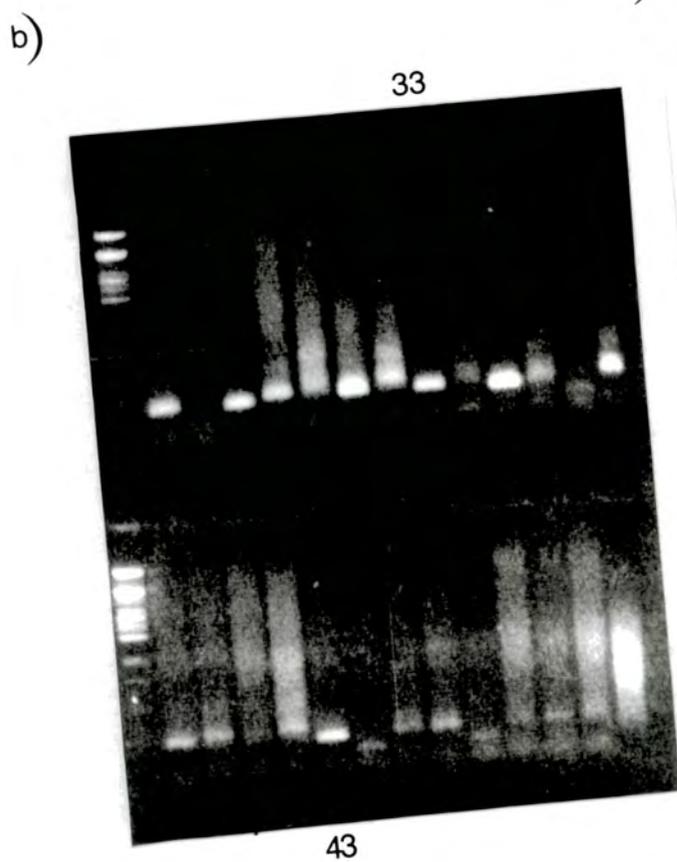
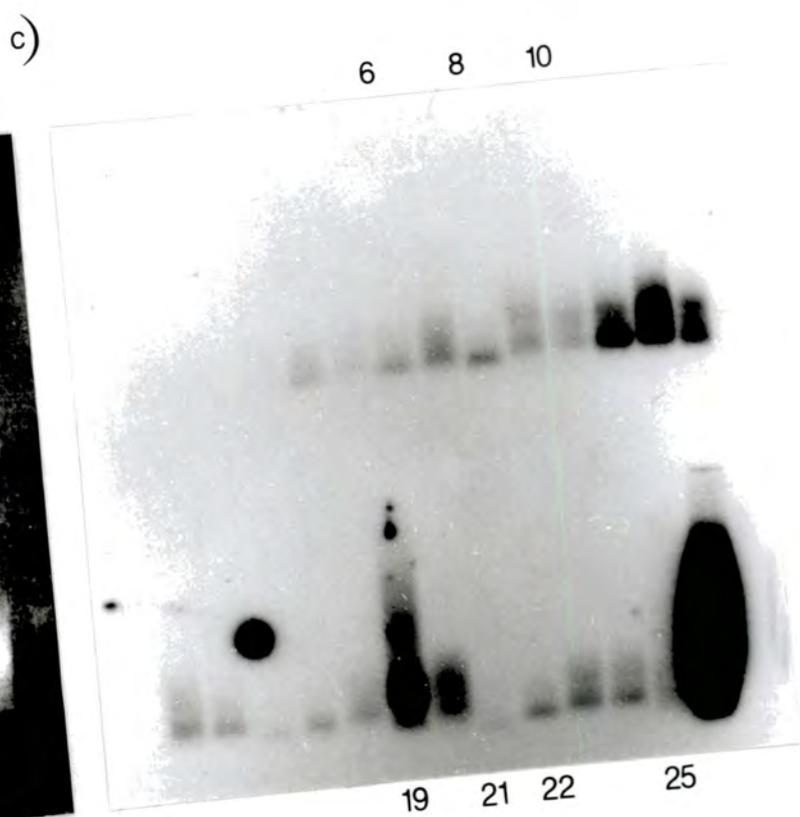
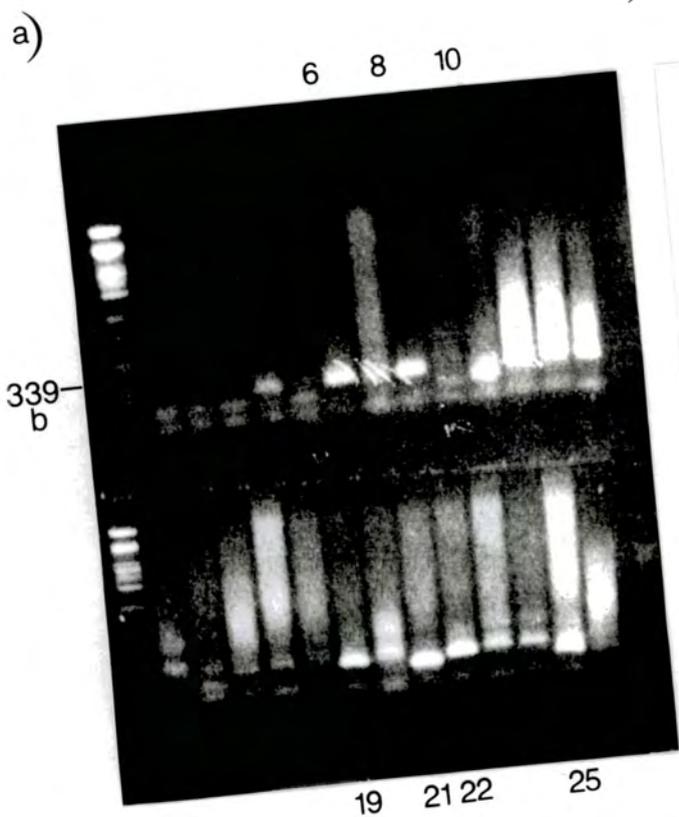
a) and b) Gel electrophoresis of PCR products (80ul of PCR product loaded per track)

Tracks 1 and 14: PstI-restricted lambda DNA size marker

Tracks 2-13 and 15-26: PCR products from clones 1-25 (gel a) and clones 26-50 (gel b)

Track 28: 18DAF first strand cDNA (positive control);

c) and d) Southern blot analysis of gels a) and b) respectively, probed with radiolabelled first strand 18DAF cDNA.



22, 25, 33, 35 and 43; the majority of these had shown strong hybridisation to the 24,26,28DAF probe, the exceptions being 43 (showing weak hybridisation) and 6 and 25 (showing medium hybridisation). Five of these were chosen for subcloning: 6, 8, 21, 33 and 35; 19, which hybridised strongly to the 18DAF cDNA was also taken for subcloning as a positive control. The remaining PCR products were run on gel and the area of electrophoresed DNA corresponding to the area of hybridisation on the Southern blot was electroeluted, purified and subcloned into EcoRI-restricted PUC18 using the method of Jung, Pestka and Pestka (1990). White transformed colonies were obtained from each subcloning experiment; however, sequencing of the plasmids demonstrated that all of the insert sequences consisted of an intact EcoRI/NotI adaptor, a poly(A) tail of between 46 and 76bp and a second EcoRI/NotI adaptor, before the PUC18 sequence was resumed. There was no relationship between the intensity of the hybridisation to the 18DAF cDNA probe and the length of the subcloned poly(A) tail - the subclone of the positive control 19 has 76 adenine bases, while the subclone of 8, which gave considerably weaker hybridisation to the cDNA probe, contained 82 adenine bases.

Further attempts to subclone and sequence the PCR products of clones 8 and 35 gave cloning artifacts of PUC18 containing one or more copies of the lambda PCR primers, extended by PCR to the EcoRI sites, including in the case of the subclone of 35 a short (30bp) poly(A) sequence attached

plaques; these were eluted in phage buffer and the 49 solutions were titred to calculate an approximate number of pfu/ml.

A 20x20cm² grid of 49 squares (7x7) was drawn on the base of a large (22.5x22.5cm²) LB agar plate. The agar was overlaid with top agar containing C600Hfl bacteria, as described in section 3.8, and a 50ul aliquot of an appropriate dilution of each phage solution was added to one of the squares (one clone per square) to give a concentrated area of adsorption well separated from the other drops of phage solution. When the liquid was fully adsorbed, the plate was inverted and incubated to allow development of plaques which were then immobilised on duplicate nitrocellulose filters. One filter was screened with the radiolabelled amplified library of 24,26,28DAF cDNA, the second with radiolabelled 18DAF cDNA (Figs 15a and 15b). Although the first filter screened with the 24,26,28DAF cDNA amplified probe gave a much stronger radioactive signal than the second filter probed with the 18DAF cDNA, the screen identified several clones which hybridised strongly to the 24,26,28DAF library PCR probe but not to the 18DAF cDNA probe (3, 5, 6, 10, 12, 14-17, 19-25, 29-37, 39, 41, 42, 44-47, 49), others which showed faint hybridisation to the 18DAF cDNA probe (1, 7-9, 18, 28, 40, 43) and a few which showed relatively strong hybridisation to both probes (2, 4, 11, 13, 26, 27, 38, 48).

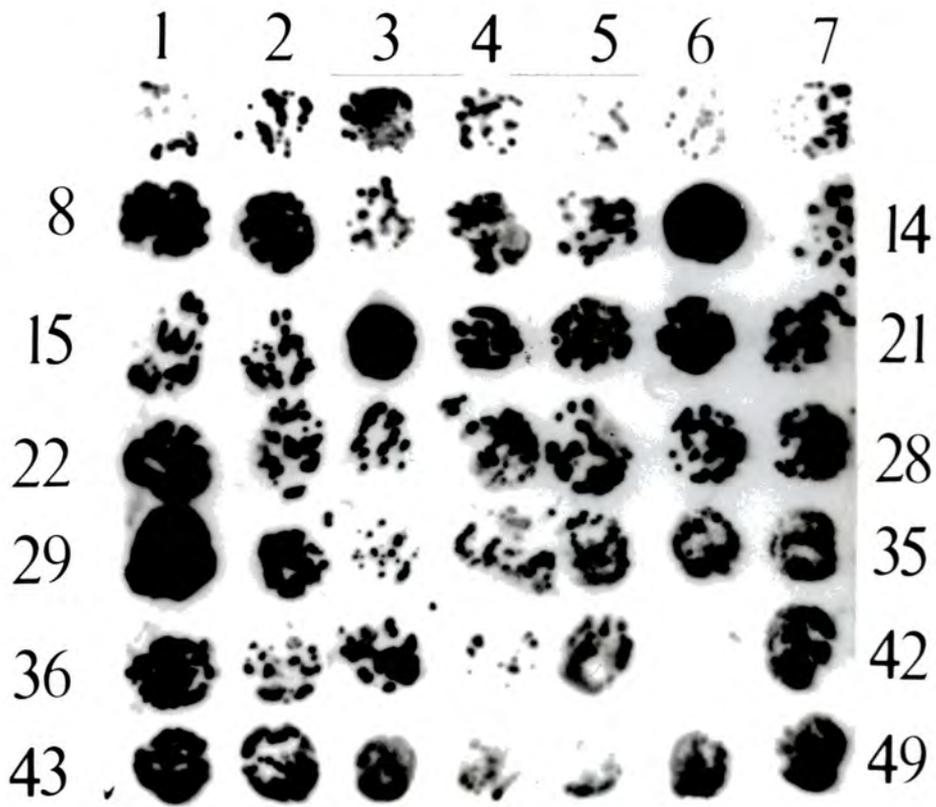
Figure 15. Differential screen of 49 selected clones isolated from the second 24,26,28DAF cDNA library. Clones were immobilised on nitrocellulose filters and screened with:

a) 24,26,28DAF cDNA library amplified by PCR;

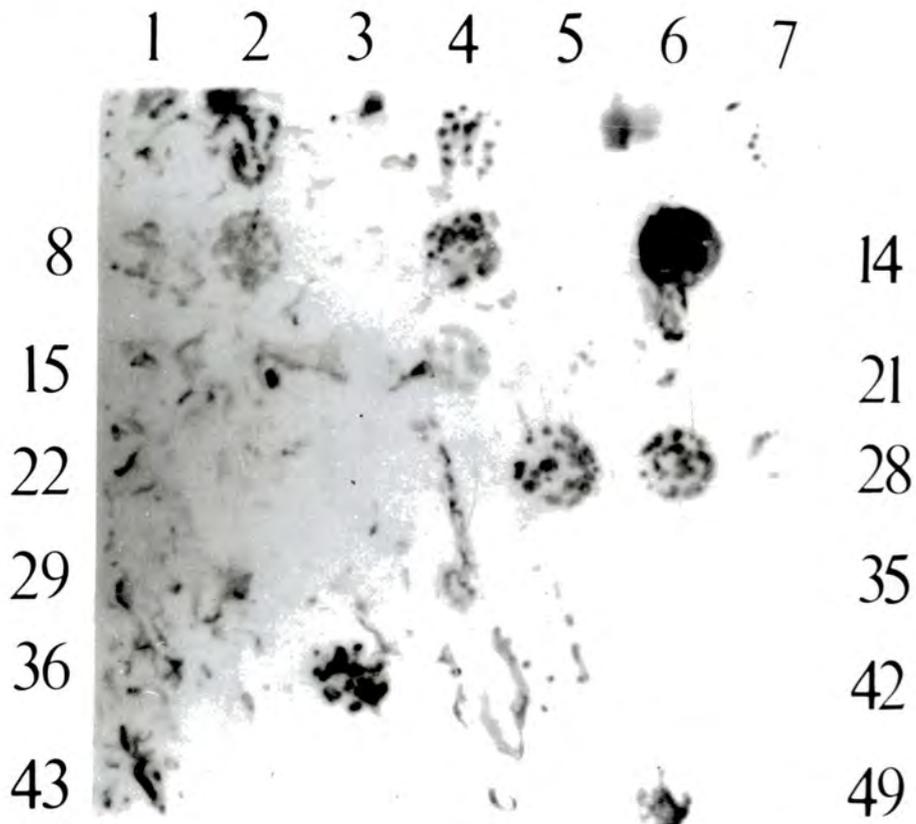
b) 18DAF cDNA.

(Clone numbers run from left to right as indicated)

a)



b)



4.1.6.5 Analysis and Subcloning of Differentially-expressed Clones Isolated from the Second 24,26,28DAF cDNA Library

The phage solutions of a number of clones which exhibited strong hybridisation to the 24,26,28DAF probe and weak or no hybridisation to the 18DAF cDNA probe were amplified to give concentrated solutions of phage particles; recombinant lambda DNA was then amplified and isolated for analysis by the lambda miniprep procedure. Restriction analysis of the resultant DNA showed that the majority of clones had lost one or both EcoRI cloning sites, a result already noted with cDNA clones from this library isolated using the Em probe. Minipreparations were restricted with BglIII+HindIII which should give a restriction fragment containing the cDNA insert within 240 and 900bp of lambda DNA (see Fig.6), giving an insert which is easy to see and isolate on gel. Restriction analysis of the prepared recombinant DNA showed that in the majority of cases the restriction fragment of 1140bp lambda DNA plus the cDNA insert ran just above or level with an 1159bp fragment in the lambda PstI marker; of the 35 clones amplified and prepared only one - 46 - had a substantial cDNA insert of ca.770bp within the appropriate restriction fragment, the others appearing to have no inserts or inserts of less than 50bp (Fig.16). As well as those clones exhibiting weak or no hybridisation to the 18DAF cDNA probe, a lambda miniprep of a

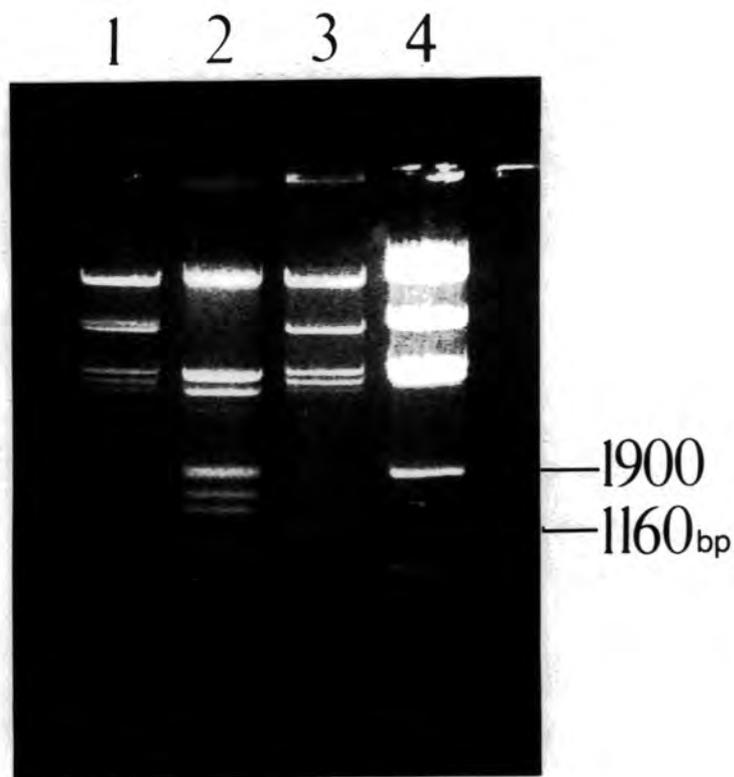
Figure 16. Restriction analysis of recombinant lambda DNA from three clones isolated from a differential screen of the second 24,26,28DAF cDNA library. The clones were restricted with BglIII+HindIII; markers indicate the restriction fragments which contain the cloning site.

Track 1: clone 18

Track 2: EcoRI/HindIII restricted lambda DNA size marker

Track 3: clone 37

Track 4: clone 46.



phage solution whose recombinant plaques showed strong hybridisation to the 18DAF cDNA probe (4) was taken as a positive control; restriction analysis of the recombinant DNA showed a similar lack of insert.

The ca.1900bp BglII/HindIII restriction fragment from the preparation of clone 46 was isolated and subcloned in BamHI/HindIII-restricted PUC18. Partial sequencing and restriction analysis of the recombinant plasmid demonstrated that both EcoRI cloning sites were functional and the cDNA insert was excised from the plasmid and subcloned into EcoRI-restricted PUC18 to give the plasmid PM46. The plasmid was sequenced in both directions to give a 787bp sequence which matched with the first partial sequence (Fig.17). The sequence begins with an intact EcoRI site from the EcoRI/NotI adaptor but only 3bp of the NotI site; this is followed by 770bp of sequence, apparently representing a partial cDNA, with a deduced open reading frame of 248bp including a stop site TAA. When the nucleotide sequence of PM46 was checked against entries in the EMBL Database limited areas of homology were seen to sequences from rice and tobacco chloroplast genome DNA and to a late embryogenesis abundant gene, LEA5, from cotton. However, a comparison of the deduced amino acid sequence with coding sequences in the SwissProt database did not identify any of these or other plant sequences as showing high homology to the isolated clone.

Figure 17. Sequence of the partial cDNA insert from clone PM46, isolated and subcloned from the second 24,26,28DAF cDNA library.

```

PM46  AAAAAACGGGGCGTGGGCAACGAGTCACCATGTCAATGCAGGATGCCGTA  50
AA      K T G R G Q R V T M S M Q D A V

PM46  TTGAACCTTTGCCGCGTGAAATTACGTGACCAGCAGCGTCTCGATAAATT  100
AA      L N L C R V K L R D Q Q R L D K L

PM46  GGGTTATCTGGAAGAATACCCGCAGTATCCGAATGGTACATTTGGTGATG  150
AA      G Y L E E Y P Q Y P N G T F G D

PM46  CAGTTCCCCGCGGTGGTAATGCAGGTGGTGGCGGTCAGCCTGGCTGGATC  200
AA      A V P R G G N A G G G G Q P G W I

PM46  CTGAAATGTAAAGGTCGCTATTTTCAGTAGCGAAAAAAGTAAAATATAAAA  250
AA      L K C K G R Y F S S E K S K I *

PM46  GTCTGACTCGCTATGTGCTGCTTTCGGTGTGCATCTCAACGGTTGTTAAC  300

PM46  CTGCTTACTATCAGCCCGTTGCGTAACAGTGATTCTTTTGTGACAGAGGG  350

PM46  GCAGTGGTTAACGTTTAAATCGATAATTGCATTGCTCATTCTTTGTGGCG  400

PM46  TAGTGTTGGCGATTAATCTGTTTTTCTGCGCTTCTCCAAGCGGTACGCTT  450

PM46  TTCTGGGCAGGCTTTTTTTTGCAGGAAATCGATCTGTTTTTCTCCAGTGAA  500

PM46  AATGCGTTGTGCGACCTTTTTTGC CAAGCCGCTTGGCTTCGGTTATTCAT  550

PM46  ATTGCTGGTTATTGAAGTGATGTGGATTACGCTGGTGTGCGGTATTGGCAA  600

PM46  CGCTTGTAGAATGGCGGATTTGGTTTGAAGCCTATTTTTTACTCTGCTAT  650

PM46  GTACCGTGCTTAATTTACTATTTTTTCTATTGTCGATTCTCTGGCATAA  700

PM46  CGATTTTATGATGGCATGTGACATGTATTTCCGTTGGGGGCATTTTAATA  750

PM46  AGTGAGGAAGTGAT  764

```

The ca.1160bp BglIII/HindIII restriction fragments from DNA preparations of clones 18 and 37 (Fig.16) were also isolated, subcloned into BamHI/HindIII-restricted PUC18 and

partially sequenced. 37 showed medium hybridisation to the 24,26,28DAF probe and no detectable hybridisation to the 18DAF cDNA probe, while 18 showed strong hybridisation to the 24,26,28DAF probe and fairly weak but clear binding to the 18DAF probe. The sequence showed 240bp intact lambda DNA up to the EcoRI cloning site, a slightly damaged EcoRI site, then portions of two NotI sites and another damaged EcoRI site before the lambda DNA sequence resumed (Fig.18).

Figure 18. Sequences of the cloning sites of two clones isolated from the second 24,26,28DAF cDNA library, showing a loss of cDNA inserts.

- a) Cloning site of normal clone, | denoting position of cDNA insert; sequence flanked by lambda DNA (---);
- b) Clone 18, flanked by lambda DNA;
- c) Clone 37, flanked by lambda DNA.

a) -----GAATTCGCGGCCGC|GCGGCCGGAATC-----
 b) -----GAATTGGCGGCCG.....GGCCGAAATC-----
 c) -----GAATTGGCGGCCG.....GGCCGAAATC-----

4.2 DISCUSSION

4.2.1 Construction of a cDNA Library from Desiccating Pea Cotyledons

The problems experienced with the cloning of cDNA from cotyledons into PUC18 seem to rest with the attachment of adaptors or linkers to the cDNA and subsequent restriction of the linkers. The first problem can be explained by the state of the ends of the cDNA. Although the cDNA synthesis procedure includes a 10 minute incubation with T4 DNA polymerase it is possible that this step is not sufficient to give blunt ends on all of the cDNA inserts, inhibiting blunt end ligation to the NotI site of the EcoRI/NotI adaptor. This would explain the relatively small quantities of adapted cDNA observed, despite an excess of adaptors in the ligation reaction. This problem is decreased slightly when using EcoRI and BamHI linkers, as the end labelling of the cDNA with Klenow enzyme may produce a larger proportion of blunt ends. The inability to clone cDNA into PUC18 using the 8mer EcoRI linkers may be explained by inefficient restriction of the EcoRI sites when attached to the cDNA inserts: although the control linker check experiment demonstrated that the linkers were restricted efficiently by EcoRI when concatamerised, other workers have noted that restriction of sites at the extreme ends of DNA is often inhibited (see Jung, Pestka and Pestka, 1990). Restriction of the longer BamHI 12mer linkers appeared to be efficient, however, and

enabled the cloning of 18DAF cDNA inserts in PUC18, the length of selected inserts suggesting that a good proportion of full length cDNAs had been cloned. The inability to clone 24,26,28DAF cDNA in the same manner must therefore be due to the quality of the 24,26,28DAF cDNA. It was noted that 24,26,28DAF total RNA gave a significantly smaller yield of poly(A)⁺RNA per mg total RNA than 18DAF total RNA (Figs.2, 3 and 4). It has been established that levels of total mRNA fall during desiccation; this also presumably includes a certain amount of degradation of mRNA due both to the developmental programme and to damage from desiccation. cDNA synthesis from 24,26,28DAF poly(A)⁺RNA was less efficient than that from 18DAF poly(A)⁺RNA - ca.0.45ug cDNA being produced per ug 24,26,28DAF poly(A)⁺RNA as opposed to 2ug cDNA per ug 18DAF poly(A)⁺RNA. It is possible that the mRNA from desiccating cotyledons has slightly damaged ends due to the apparently high levels of mRNA degradation at this seed age, and, further, the mRNA may experience greater degradation on isolation and/or form secondary structures which produce low yields of cDNA with uneven ends, to which either BamHI and EcoRI linkers will not attach, or which will inhibit linker restriction and cloning into restricted PUC18.

The 24,26,28DAF cDNA used for the library constructed in lambda gt10 was incubated with T4 DNA polymerase for 30 minutes after cDNA synthesis rather than the suggested 10 minutes, to give a higher proportion of blunt ends. cDNA was also separated from free nucleotides

which may inhibit ligation on a Sephadex column before ligation to EcoRI/NotI adaptors. Ligation of the cDNA to the adaptors and to the prepared lambda cloning vector was successful, although the cloning efficiency was at least 10-fold lower than expected, possibly due to a low proportion of adaption of the cDNA. Screening of this first cDNA library with a probe for the *Leg J* gene demonstrated that the messages hybridising to the *Leg J* sequence are present in desiccating cotyledons, confirming the results of other workers (Thompson *et al*, 1989; Thompson *et al*, 1991). The observation of a varying strength of radioactive signal from the plaques hybridising to the probe could be explained by the presence of messages for *Leg L*, whose hybridisation to the probe may be slightly weaker than that to the *Leg J* message, and also by the presence of partial cDNAs of the *Leg J* subfamily.

The use of the Qiagen midiprep system was found to be unreliable; preparations of lambda DNA were obtained and could be restricted but the yields were very small, and as the inserts formed only a small proportion of the total length of lambda DNA, they could not be isolated from midiprep restrictions for subcloning. Other workers have found that the correct density of bacteria before the addition of the phage particles is essential for efficient DNA preparation. PCR amplification utilising primer sites on either side of the lambda-gt10 cloning site gave a better yield of insert DNA isolated from the surrounding lambda DNA.

Amplification of several of the *Leg J* clones from the library screen suggested that the cDNA library contained full length as well as partial cDNAs.

The efficiency of cDNA synthesis from poly(A)⁺RNA isolated using the PolyATtract system was noticeably higher than from that isolated by spun column chromatography, giving ca.1ug cDNA per ug 24,26,28DAF poly(A)⁺RNA. The speed and efficiency of this isolation procedure should give poly(A)⁺ RNA with no further degradation than that already present in the seed. However, the cloning efficiency of the cDNA in the construction of the second cDNA library was no higher than that in the previous library in the test ligation, suggesting that there may still be a problem with the ligation of the adaptors to the cDNA or of the adapted cDNA to the cloning vector. It is possible that the sample of poly(A)⁺RNA contained a high level of RNase, carried over during the isolation process from the presumably high levels directing mRNA degradation in the desiccating seed; this could have acted to give damaged ends or a certain degree of degradation and so explain the failure of the more rapid and efficient mRNA isolation system to produce cDNA with a high cloning efficiency. The level of RNase in poly(A)⁺RNA samples of different developmental ages should be tested by incubation with RNA followed by gel electrophoresis to determine whether this possibility is true, so that preventative measures could be taken in future cloning attempts.

Loss of viable phage particles upon the addition of

chloroform prevented the production of a large library. The chloroform used was a large communal supply; it was concluded that the stock solution must have been contaminated with some other chemical that was deleterious to lambda particles. This effect had been noted in the past by other workers. It was also observed in these experiments that the addition of fresh chloroform reduced the titre of phage particles, although not so completely as the chloroform used in the library construction. Unfortunately the chloroform was not identified as the cause of the problem until all of the lambda arms had been used. Thus the second cDNA library was only slightly larger than the first; nevertheless, due to the use of the PolyATtract isolation system it was assumed that the library would be of good quality and contain a large proportion of full length cDNAs.

The screening of the second lambda cDNA library with PJC5.2 supports the assumption that the library contained a greater diversity of messages, as the messages which hybridised to the *Leg J* sequence formed a smaller percentage of the library. However, isolation of clones with homology to the *Leg J* probe by both PCR amplification and lambda minipreparation gave no inserts above 900bp, suggesting that the library had a greater proportion of partial cDNAs.

4.2.2 The Occurrence of Seed and Putative Stress Protein Messages in Desiccating Pea Cotyledons

4.2.2.1 Cowpea Trypsin Inhibitor Protein

The results indicate that neither of the cDNA libraries constructed from cDNA from 24,26,28DAF desiccating pea cotyledons contained sequences possessing sufficient homology to the *CpTI* probe to allow binding. This does not necessarily rule out the presence of a gene with similarity to the cowpea *CpTI* gene in the pea genome: although in cowpea the protein accumulates during late seed desiccation (J.A. Gatehouse, unpublished) it is possible that a pea cDNA with homology to the *CpTI* sequence occurs at a very low abundance during desiccation and is therefore not present in the library, which was not constructed to include messages of such low abundance; or such a message may be absent during this stage of development but present in the seed or plant tissues at other times. The results also do not preclude the presence in the pea seed of a different protease inhibitor specific to pea which may be expressed during late desiccation but does not exhibit hybridisation to the heterologous cowpea probe. This is discussed further in Chapter 5.

4.2.2.2 Pea *PsMT_A*

The *PsMT_A* transcript has previously been isolated from pea root, leaf and 14DAF cotyledons (Evans *et al*, 1990).

The inability to isolate a clone with homology to the *PsMT_A* transcript from the cDNA libraries constructed from desiccating cotyledons indicates that *PsMT_A* is either rare or absent in the mRNA population of mature dehydrating cotyledons and so cannot be classed as a late embryogenesis abundant gene. The changes in *PsMT_A* transcript levels in the seed are investigated further in Chapter 5.

4.2.2.3 Wheat Em Protein

In all hybridisation experiments to the wheat Em subclone the radioactive probe bound strongly to the apparently positive plaques with a low stringency washing and this strength of hybridisation was retained throughout DNA preparation and subcloning. In all cases the clones which hybridised to the Em probe were short and a long poly(A) tail made up most if not all of their length. If a gene with some amino acid and nucleotide sequence homology to the wheat Em message did exist in the pea genome, it is unlikely that the 3' untranslated sequence would be conserved - no homology is seen between the 3' flanking sequences of the Em genes from the more closely related wheat and maize (Litts *et al*, 1987; Litts *et al*, 1992). It can therefore be concluded that the probe was hybridising to clones containing long poly(A) tails of between 50 and 100bp and not to any specific sequence which bore homology to the Em coding sequence. The plaques isolated from the library screen were those showing the strongest hybridisation to the Em probe; consequently, if any

pea sequence exists which shows homology to the wheat Em LEA message the homology is either very weak and hybridisation between the probe and cDNA can be broken by washing at 5xSSC, 0.1%SDS at 65°C, or the sequence itself is rare or absent at this developmental stage, in contrast to the expression of Em-like genes in other plants. The presence in pea seeds of a message with homology to the wheat Em sequence is discussed further in Chapter 5.

4.2.3 An Investigation of cDNAs Coding for Pea Ubiquitin

4.2.3.1 Attempted Isolation of Ubiquitin Clones from the Second 24,26,28DAF cDNA Library

From the results it is clear that ubiquitin messages are present in the cotyledons during desiccation at 24,26,28DAF, although they do not appear to be particularly abundant at this developmental stage, forming approximately $5 \times 10^{-3}\%$ and $1.4 \times 10^{-3}\%$ of the cDNA content of the first and second cDNA libraries respectively. However, in spite of the presence of lambda-cDNA clones showing strong hybridisation to the polyubiquitin probe, ubiquitin sequences could not be isolated easily from the second cDNA library. As with the results from the Em screen, restriction analysis showed that both EcoRI sites in the cloning adaptors had been lost, which made excision of the clones for subcloning and sequencing difficult. However, a further problem was noted with the continued growth and preparation of insert DNA - inserts were

continually being lost or shortened during further growth, lambda mini-preparations and subcloning. In the primary screen of the first lambda cDNA library the plugs which were eluted contained approximately 10 plaques, including the positive plaque. Yet on secondary screens of filters containing between 100 and 200 plaques, only one or two plaques on each filter showed hybridisation to the polyubiquitin probe. Further, when white colonies obtained by transformation of DH5alphaME cells with restriction fragments containing the insert DNA were grown overnight at 37°C for colony hybridisation, it was noted that over half of the white colonies turned partially blue, although they still gave a strong signal when hybridised to the radiolabelled PCU1 insert. This phenomenon was not observed in the subcloning of other types of cDNA. It is possible that the polyubiquitin cDNA is unstable in the lambda-gt10 clones during storage and propagation - being composed of tandem repeats it would be especially susceptible to recombination if allowed by the host bacterium - and that some or all of the polyubiquitin coding sequence was 'jumping out' of the lambda DNA during growth and replication. The cells provided with the lambda-gt10 Protoclone cloning kit for maintenance and screening of cDNA libraries are not identified in the kit protocol as RecA⁻. A similar lack of stability of polyubiquitin was proposed for polyubiquitin messages from yeast cells (Ozkaynak *et al*, 1987); these workers first isolated a polyubiquitin clone containing 6 ubiquitin repeats

from a lambda-gt11 library, but a later isolation (presumably from the same gene as yeast is thought to have only one polyubiquitin gene) from a plasmid-based library yielded a sequence of 5 repeats. However, they also stated that the difference in repeat number may be a case of natural variation between yeast strains.

4.2.3.2 Isolation of Two Pea Ubiquitin Extension Protein cDNA Clones from the First 24,26,28DAF cDNA Library

Screening of the first cDNA library with a polyubiquitin probe was more successful, producing cDNA clones whose sequences corresponded very well to those of the two identified ubiquitin extension proteins, seen in other species to be involved in ribosome biogenesis, with the tail proteins forming part of the ribosomes and the ubiquitin acting temporarily as a 'chaperone' (Finley, Bartel and Varshavsky, 1989). The ubiquitin amino acid sequences are totally conserved with those of other plant species although, as in most cases, there is some divergence at the nucleotide sequence level. The two tail protein sequences are also very highly conserved with similar tails from several other species. In the small number of sequences analysed in Tables 1 and 2 the highest degree of tail amino acid conservation appears to be within plants (52 amino acid tails, pea PCU3 and *A.thaliana*; 76-81 amino acid tails, pea PCU4, *A.thaliana*, tomato and potato), within animals (52 amino acid tails, human and *Drosophila*) and between plants and animals. A

slightly lower level of conservation is seen between the 52-53 residue tails of plants and animals and those of the alga *Chlamydomonas reinhardtii* or the yeast *Saccharomyces cerevisiae*, with the protozoa *Tetrahymena pyriformis* showing the least conservation of amino acid sequence. In the 76-81 residue tail the yeast sequence shows a considerably lower level of homology to the plant sequences, although over a quarter of the disparate residues are in the last 8 amino acids of the yeast sequence (Fig.11). Despite variation in amino acids, the positions within the two tail types of the putative nuclear localisation site and the proposed cysteine-rich zinc-binding nucleic acid binding domain are conserved in all cases except in a 52 residue-type tail isolated from tobacco, which has an insert of 20 amino acids (reviewed by Pollmann, Kampen and Wetter, 1991). This high conservation between very different organisms has been taken to imply a similarity of function of the ubiquitin extension proteins of different species (Ozkaynak *et al*, 1987; Baker and Board, 1991; Pollmann, Kampen and Wetter, 1991). It may therefore be assumed that the pea ubiquitin extension proteins coded by these cDNAs are involved in ribosome biogenesis, with the ubiquitin possibly performing some further function in the cell after its associated tail protein has been incorporated into a ribosomal subunit. Ribosomal RNA and ribosomes have been shown to be synthesised during seed growth and maturation and stored in the dried seed for use upon germination (reviewed by Payne, 1976). Thus the presence of

messages coding for the ubiquitin extension proteins, assuming they are used in ribosome production, is consistent with the constitutive expression of these extension proteins to build up ribosomes both for use and storage, and also perhaps to form part of the long-lived mRNA fraction to allow rapid synthesis of ribosomal proteins upon germination. The apparently low level of transcripts hybridising to the ubiquitin probe seen in the library screens may be due to a naturally low level of polyubiquitin and ubiquitin extension protein expression: Gausing and Barkardottir (1986) found that higher plants seem to have a low level of ubiquitin, determined by antibody precipitation, while a number of researchers have shown that ubiquitin extension protein messages are only abundant in actively dividing cells or tissue (reviewed by Garbarino, Rockhold and Belknap (1992). Alternatively, it may be due to a general lower level of expression as the seed desiccates and the RNA and protein synthesis mechanisms become less active.

4.2.3.3 Polyubiquitin cDNAs in Pea

Despite the use of a polyubiquitin probe, no polyubiquitin cDNA sequences were isolated from either 24,26,28DAF cDNA library, although their presence has previously been demonstrated in pea leaves. This suggests either that polyubiquitin is absent in desiccating pea seeds, exhibiting differential expression between tissue types or upon seed aging, or is present at low levels. If the latter

case is true, the levels of the polyubiquitin messages were either so low as to avoid detection in the screen used or the cDNAs were difficult to clone, possibly due to the proposed instability of repeated transcripts in this cloning system. It seems probable that free ubiquitin would be required in the developing seed for protein degradation and its other proposed functions. It would also presumably be required in the early stages of dehydration for the same roles until the seed becomes quiescent. However, as it is thought that ubiquitin polypeptides can be recycled it is possible that any requirement for ubiquitin can be filled by the expression of ubiquitin extension proteins and/or of polyubiquitin genes in the young seed, building up enough ubiquitin to allow the seed to function through maturation and desiccation without the need for continued expression of polyubiquitin genes. The expression of messages hybridising to a ubiquitin probe in different seed ages and pea tissues is examined further in Chapter 5.

The full sequencing of the second polyubiquitin partial cDNA from pea leaves (donated by F.Z. Watts) demonstrated that pea contains two different polyubiquitin genes with non-homologous 3' flanking sequences. Thus pea contains a ubiquitin multigene family of at least four members, including two polyubiquitin and two ubiquitin extension protein genes. Unfortunately, although the PCU2 clone was shown to code for three full and one partial ubiquitin repeats, the complete number of ubiquitin repeats

in this message could not be determined, nor was any 5' flanking sequence cloned. Similarly, the 5' flanking sequences of the pea ubiquitin extension protein cDNAs were not isolated; the use of BglII during subcloning restricted the cDNA after the first two codons of the ubiquitin sequence and attempts to subclone the remaining restriction fragment containing the 5' flanking sequence were not successful. Thus the 5' flanking regions of the PCU2 polyubiquitin and PCU3 and 4 ubiquitin extension protein cDNAs could not be compared with the short 5' flanking sequence in the polyubiquitin PCU1 cDNA clone (Watts and Moore, 1989) or with polyubiquitin and ubiquitin extension protein genes and messages from other species. Further research is necessary to determine the full number of ubiquitin coding genes in the pea genome, and to isolate and examine the promoters and flanking regions of these genes to attempt to identify characteristic features or regulatory elements which may aid understanding of the expression of these genes.

4.2.4 Differential Screening of 24,26,28DAF Lambda-cDNA Libraries

4.2.4.1 24,26,28DAF cDNA Library Screened with 18DAF cDNA

From the screening of the first cDNA library with 18DAF total cDNA it appears that the majority of messages present in the 24,26,28DAF library from desiccating

cotyledons are present in the mature seed before desiccation. This could merely be a consequence of the persistence of messages throughout seed development as proposed by Goldberg *et al* (1981); another explanation (which does not exclude the persistence of messages) is that the synthesis of messages present in desiccating cotyledons begins or increases as the seed moisture content falls gradually during seed filling. The presence of LEA messages has been observed in seeds from the mid embryogenesis stage (reviewed by Goldberg, Barker and Perez-Grau, 1989), and the synthesis of messages stored for use upon germination, which, though not necessarily abundantly synthesised before germination, are expected to form part of the total message population in the desiccating seed, has also been noted during late maturation. Abundant seed-specific messages for seed storage proteins, lectins etc. are still present at this developmental stage, although their levels are falling, and messages for constitutive proteins may also be present at decreased levels. Some places of weak or negative hybridisation of the immobilised plaques to the 18DAF cDNA were noted on short exposure of the hybridised filters, some perhaps representing the messages for the proteins that Evans and Swinhoe found to be synthesised upon desiccation (see section 1.5). Unfortunately the overwhelming hybridisation to the majority of plaques combined with the stretching of nitrocellulose filters during processing made isolation of these few plaques impracticable.

4.2.4.2 Screening of the 24,26,28DAF cDNA Library with a Subtracted Probe

The results indicate that a cDNA probe of 24,26,28DAF cDNA subtracted with 18DAF poly(A)⁺RNA was obtained, but that there was insufficient DNA or radioactivity present in the probe to give detectable hybridisation to the small plaques used in the cloning experiment. Sambrook, Fritsch and Maniatis (1989) suggest two rounds of hybridisation to ensure full subtraction; however, it is not only any newly-expressed clones which are relevant to LEA message studies, but also those clones whose expression increases between 18 and 28DAF. Consequently, it was not desirable to subtract all of the messages which were present at 18DAF, but rather to remove that proportion which is there at both developmental stages, and so highlight any differential expression. To this end, only one round of subtractive hybridisation was carried out to avoid the subtraction of those messages which are present at 18DAF but whose levels increase significantly during desiccation. Unfortunately the one round of hybridisation appears to have been sufficient to extract nearly all of the messages which are common to both ages, leaving a probe which, although labelled to high specific activity, was insufficient to screen the cDNA library. Due to the expense of this procedure it was not repeated.

4.2.4.3 Screening of Selected Clones from the First 24,26,28DAF cDNA Library with 18DAF cDNA

To isolate those messages which are abundant during seed desiccation, the first cDNA library was probed with the library itself, amplified by PCR and radiolabelled. The inserts from the selected and isolated clones were then amplified by PCR. Due to the time spent on other differential and specific screens, the lambda cDNA had been stored for several months before this experiment was conducted, and the isolated clones were stored for a further month before PCR amplification. This led to a drop in titre with storage at 4°C and required amplification of the phage particles by overnight growth in plated C600Hfl cells and elution of the phage particles. The PCR amplification of 50 clones after this storage and repeated growth gave no clear bands of DNA, only fragments of several sizes causing a faint smear of DNA down the gel and a concentrated band of fragments of 200-300bp. The use of different primer stocks, Taq Polymerase, dNTP and buffer stocks all gave no difference in the PCR product, while tests of the PCR technique with plasmid DNA gave a typical correct amplification. This implied that the problem was in the lambda DNA itself: possibly the inserts were being lost on storage and growth, leading to amplification of partial inserts; this induced heterogeneity of the clones, with the stored stocks containing phages which had lost different lengths of

inserts, and so giving a range of sizes of amplified fragments; alternatively the lambda primer sites could have sustained some sequence alteration, leading to inefficient annealing of the primers during amplification with the production of different sizes of amplified inserts. This effect of attempted PCR amplification on stored lambda stocks was also seen with clones isolated from the *Leg J* screen. The stock solution that contained a ca.1600bp insert after isolation (Fig.5) only produced an indistinct ca.800bp insert after 3 months storage (Fig.12).

Despite the shortness of the PCR products, a screen of a Southern blot of the amplified inserts with 18DAF cDNA showed differential expression of several clones. In accordance with previous results and the findings of Goldberg *et al* (1981) nearly all of the inserts showed some hybridisation to the 18DAF cDNA, but a proportion (approximately one in five) of the inserts which had exhibited strong hybridisation to the 24,26,28DAF PCR library probe showed comparatively weak hybridisation to the 18DAF cDNA probe. Isolation, subcloning and sequencing of the PCR products of 5 clones which showed weak hybridisation and one clone which showed strong hybridisation to the 18DAF cDNA probe revealed that only poly(A) tails had been subcloned, while further experiments also failed to subclone any cDNA sequence. It is very unlikely that the isolated clones consist solely of poly(A) cloned with EcoRI/ NotI adaptors into lambda-gt10: in the construction of the cDNA library the

cdNA was size-selected to avoid short fragments of cdNA; further, the clear difference seen between the strengths of hybridisation of the selected clones to 18DAF cdNA indicates that some cdNA sequence other than the poly(A) isolated is present, but could not be isolated. Presumably the subclones obtained are artifacts, some consisting of two partially amplified poly(A) sequences with one EcoRI/NotI adaptor each joined in the middle, others containing one poly(A) sequence with an adaptor and primers amplified to the EcoRI site then cut by EcoRI. This indicates that at least some of the primers did anneal to the lambda-cdNA clones and extend, but the short primer fragments and amplified poly(A) fragments are subcloned in preference to any non poly(A) cdNA sequence.

4.2.4.4 Screening of Selected Clones from the Second 24,26,28DAF cdNA Library with 18DAF cdNA

The 24,26,28DAF screen gave a much stronger radioactive signal than the 18DAF cdNA screen. Part of this may be due to the fact that there is a certain amount of background hybridisation of the PCR primers at the ends of the PCR products to the homologous primer sites in the lambda DNA. The 18DAF probe was also used to probe the second filter, which usually gives a somewhat weaker signal. It is also generally agreed that the complexity of abundant messages is greater at 18DAF than at 24,26,28DAF and this may have contributed to the difference in hybridisation levels of the two probes to the isolated clones. However, it is

possible to identify those clones whose expression is noticeably greater at 24,26,28DAF than at 18DAF. Fewer clones appear to be present at both stages from the differential screen of this library, as compared to the first differential screen; however, in both cases it is probable that some of the clones occur more than once in the 50 or 49 selected plaques, the first screen perhaps collecting several copies of proteins which are present at both stages, such as seed storage proteins, the second containing a higher proportion of differentially-expressed clones.

The subcloning and sequencing of the BglIII/HindIII 1160bp fragments of clones 18 and 37 explain the lack of inserts in the majority of lambda DNA minipreparations. It would appear that recombination is occurring between the two NotI sites - which by their sequence appear to be a good target site for crossing-over - and the inserts are being lost. If it was simply a matter of the subcloning of free adaptors, one would have expected them to be intact and not to have shown up on the 24,26,28DAF screen, 23bp of sequence out of 43,340bp not being sufficient for the strength of hybridisation observed. Also, the cDNA used for cloning was separated carefully from free adaptors on a Sepharose column. Further, 18 and 37 showed different levels of hybridisation both to the 24,26,28DAF and the 18DAF probes in the differential screen, with the 18DAF cDNA binding quite clearly to the plaques of clone 18 but showing no noticeable hybridisation to clone 37. It is possible that the degraded

state of the 24,26,28DAF poly(A)⁺RNA gave irregularities at the ends of the cDNA, despite the T4 polymerase step at the end of cDNA synthesis; these damaged ends could form unstable attachments to the NotI site, highlighting this area for recombination. This could also explain the loss of restriction sites, due to mending of the damaged DNA, resulting in the insertion of different bases as the DNA is stored and replicated. Recently, one of the manufacturers of lambda cloning kits, New England Biolabs, mentioned in their catalogue that EcoRI sites are frequently lost during cDNA cloning, although this has not been admitted by any of the other companies.

The inability to isolate more than one cDNA insert from the plaques used in the differential screen calls into question the validity of the apparent differences in message populations between 18 and 24,26,28DAF. However, as with the first differential screen of PCR products, three points support the indicated differential expression of messages between the two developmental stages. Firstly, the filter screened with radiolabelled 18DAF cDNA clearly shows that some clones hybridise and others do not. This could merely indicate that those clones which hybridise to the probe contain cDNA inserts and the clones to which the probe does not bind have lost their inserts; however, the amplification, isolation and restriction analysis of lambda DNA from a clone which showed fairly strong hybridisation to the 18DAF cDNA probe as a positive control (4) was found to contain no

insert of any discernible length, indicating that inserts were being lost from the lambda clones after the screening experiment. Thirdly, clone 46 which was isolated and sequenced to show a cDNA insert exhibited strong hybridisation to the 24,26,28DAF library PCR probe but no detectable hybridisation to the 18DAF cDNA probe, indicating that even if the message for clone 46 is present at 18DAF its levels, either actual (implying up-regulation) or proportional (implying a continuation of message accumulation when the majority of other message levels are falling) are significantly increased during seed desiccation. Analysis of clone 46 itself shows that it is a partial cDNA, the majority of the sequenced length apparently consisting of 3' non-coding sequence. The loss of part of the NotI site at one end of the clone indicates that a portion of the clone may have been lost during cloning and amplification; surprisingly, the clone has no poly(A) tail - this may also have been lost during cloning or growth. Unfortunately, no sequences showing significant homology to the PM46 deduced coding sequence were identified in the nucleotide and amino acid sequence databases. The expression of PM46 in the seed is examined further in Chapter 5.

The loss of both restriction sites and inserts is not necessarily as complete as the results suggest. It must be remembered that the size of the lambda DNA means that any cDNA inserts form only a small percentage of the DNA mass; in restriction analysis of the DNA small quantities of any size

fragment, or even fairly large quantities of DNA fragments under 500bp, are very difficult to see on gel. It is possible that the amplification and maintenance procedures cause some, or even the majority, of lambda particles to lose their inserts; in this case the remaining inserts may well be too faint to see and isolate. Unfortunately it is impossible to separate phage containing cDNA inserts from those with no inserts as the remnants of the EcoRI/NotI adaptors act as inserts and allow plaque formation in C600Hfl (which do not allow plaque formation of lambda particles with no inserts). It appears that the main faults with the procedures used in this thesis were the use of the C600HF1 cells provided with the kit for growth and maintenance of the lambda clones, which clearly allowed recombination; the use of EcoRI/NotI adaptors, which are used commonly mainly because of the low frequency of NotI sites in animal systems and which seem to provide a prime site for recombination, especially if attached to incomplete or damaged ends; and the length of time used to test different methods of lambda DNA preparation, subcloning of PCR products and differential screening, as the long storage led to significant drops in titre which then required further overnight growth of lambda particles to amplify the DNA, giving further opportunity for recombination and loss of inserts. The use of Eppendorfs for storage at 4°C with chloroform is now also thought to be detrimental to the titre of phage particles and the quality of lambda DNA, as the chloroform is thought to leach

plasticisers from the tubes. In retrospect it is recommended to use a RecA⁻ strain of bacteria for growth of clones, storing the library for as short periods as possible in glass with fresh chloroform. As the problem with loss of restriction sites appears to be fairly common, the choice of a lambda cloning vector with alternative restriction sites fairly close to the cDNA insert position would also be recommended.

CHAPTER 5 - CHANGES IN THE LEVELS OF SPECIFIC MESSAGES AND PROTEINS IN DEVELOPING AND TREATED PEA SEEDS

Seeds have been identified as specialised organs with a pre-set developmental programme which determines major changes in patterns of mRNA and protein levels. Although this programme must be controlled to the extent that all events necessary to the production of a viable seed occur (e.g. the synthesis of sufficient storage reserves and mRNA to be stored for germination), and abnormal development (i.e. precocious germination) is inhibited, it must also contain a certain 'plasticity' to allow the seed to adapt to environmental conditions while remaining viable (reviewed by Gatehouse *et al*, 1986). A major part of the seed's maturation to allow germination and growth upon imbibition is a period of desiccation, which is thought to be essential for successful germination and growth in seeds of many species (Bewley, 1979). As dehydration is known to cause a variety of injuries in plant cells (see section 1.1; reviewed by Levitt, 1980) it is important for the seed to have the ability to protect essential structures and proteins from the effects of dehydration and rehydration, and/or to repair or resynthesise damaged structures and macromolecules. To remain viable, the seed must also protect itself during development and quiescence from attack by invertebrates; further, upon germination the seed may need not only to be able to absorb and bind essential metals from the environment

for cellular processes and enzyme structure but also to possess a defence against toxic heavy metals which may also be absorbed during imbibition. To examine the seed's ability to remain viable despite potential internal and environmental damage the presence of messages or proteins from several 'stress-response' proteins was examined in pea seed: a putative pea metallothionein message, *PsMT_A*, thought to be involved in binding of essential metals and possibly short-term detoxification of heavy metals; cowpea trypsin inhibitor protein, which protects its native plant from insect attack; pea ubiquitin, a protein which is thought to be universal and is associated with a variety of cellular processes, notably the degradation of damaged or unwanted proteins and, in the form of a ubiquitin extension protein, ribosome biogenesis; a molecular chaperone, chaperonin 60, which has also been implicated in the degradation and reassembly of damaged proteins and macromolecular structures, although its best-characterised role is in normal cellular metabolism for protein assembly and transport within the cell. Another putative stress protein message examined was that of the wheat Em polypeptide - a typical LEA protein thought to be involved in desiccation protection which shows significant amino acid homology to similar polypeptides found in both mono- and dicotyledons. Hybridisation of the pea partial cDNA clone PM46, isolated from a differential screen of the 24,26,28DAF lambda cDNA library with 18DAF cDNA, to total RNA from developing peas was also investigated. As well as those

messages and proteins thought to be associated with stress response, the expression of components of one of the major storage proteins of pea were examined - *Leg A* and *Leg J*, messages for a major and a minor polypeptide (respectively) of the heteropolymer legumin which, together with vicilin, forms up to 50% of the pea seed's total protein content. The normal expression of these messages and proteins from mid-cotyledon expansion (12DAF) to late desiccation (30DAF) was compared with the effect of premature desiccation and the application of exogenous ABA, thought to be somehow involved in osmotic and other stress transduction, to try and gain an understanding of how the pattern of gene expression in the developing and desiccating seed relates to the production of a viable, protected organ with the ability to germinate and grow upon imbibition.

The expression of these messages was examined using Northern blots of total RNA from developing and treated cotyledons extracted by the hot SDS method of Hall *et al* (1978) and from developing embryonic axes extracted by the guanidinium hydrochloride method of Logemann, Schell and Willmitzer (1987). Northern blots were commonly loaded at 10ug per track and the RNA visualised before blotting to ensure that the RNA weights were equal. The comparison of the strengths of hybridisation to RNA samples of different ages should give some indication of how the message levels of the proteins under study vary, if at all, during development and after treatment; however it must be noted that the

results are only from single batches of peas and are not quantitated, and should be taken only as preliminary indications of message level. Dot blots of total RNA were also screened to give a more concentrated area of hybridisation and to allow comparison between a larger number of RNA samples on the same blot. The amount of hybridisation of each dot in these blots was not measured quantitatively, as in most cases the length of exposure required was over two weeks and made counting of the radioactivity by liquid scintillation analysis difficult and unreliable, and the series of tests and controls required to allow accurate quantification of the level of hybridisation by densitometry were not attempted. However, densitometric scans of the two or more dot blots used for each probe, averaged and corrected to take into account the diameter of the spot of hybridisation, gave a rough unquantitative pattern of message expression during development and between tissues.

5.1 RESULTS

5.1.1 The Effect of Cotyledon Age on Specific Message Levels

The changes in message levels of selected probes for seed-specific messages and messages associated with stress response as the seed ages and desiccates were examined using cotyledons from peas of ages from 12 to 20DAF. Each RNA sample was composed of RNA from cotyledons from several

Pods.

5.1.1.1 *PsMT_A*

mRNA species detected by the *PsMT_A* gene exhibit a considerable decrease in level with seed age (Fig.19a). The probe hybridises quite strongly to the total RNA fraction from 12DAF cotyledons, detecting a single band of approximately 600b. However, at 14DAF the message level had decreased significantly and continued to fall until it was undetectable, even after a 3 week exposure, at 26DAF. This fall in transcript level is confirmed by a screen of total RNA on a dot blot with the *PsMT_A* probe (Figs.20a and 21a); this demonstrates the rapid decrease in message concentration after 14DAF with the level of probe hybridisation to seed RNA appearing to fall below the level of that to the negative control (*E.coli* rRNA) by 22DAF.

5.1.1.2 Pea Ubiquitin

An unexpectedly complicated pattern of expression is seen from total RNA samples of varying cotyledon ages probed with a pea polyubiquitin cDNA isolated from leaves (PCU1). The transcripts hybridising to the ubiquitin coding sequence appear to exist at fairly low levels within the seed - 40ug of total RNA had to be loaded onto a Northern blot exposed for over two weeks before clear hybridisation could be seen (Fig.19b). The changes in levels of pea ubiquitin-coding messages are also shown on a dot blot (Fig.20b).

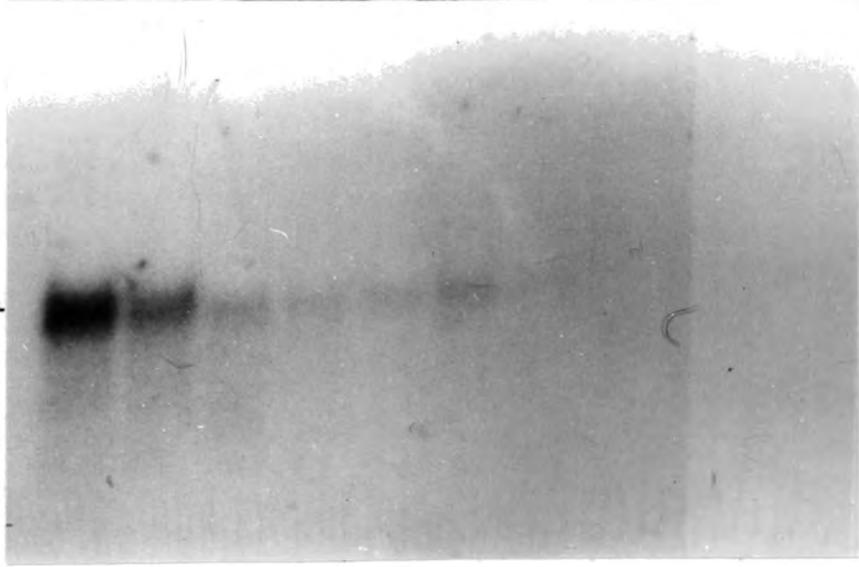
Figure 19. Northern blot analysis of total RNA from developing pea cotyledons, probed with a) *PsMT_A*; b) PCU1 (pea polyubiquitin); c) PJC5.2 (*Leg J*); d) PRC3.1 (*Leg A*). All blots contain 10ug RNA per track except blot b) which contains 40ug RNA per track. Average size of hybridising bands is indicated.

Tracks 1-10: total RNA from 12, 14, 16, 18, 20, 22, 24, 26, 28 and 30DAF cotyledons respectively.

a)

1 2 3 4 5 6 7 8 9 10

600

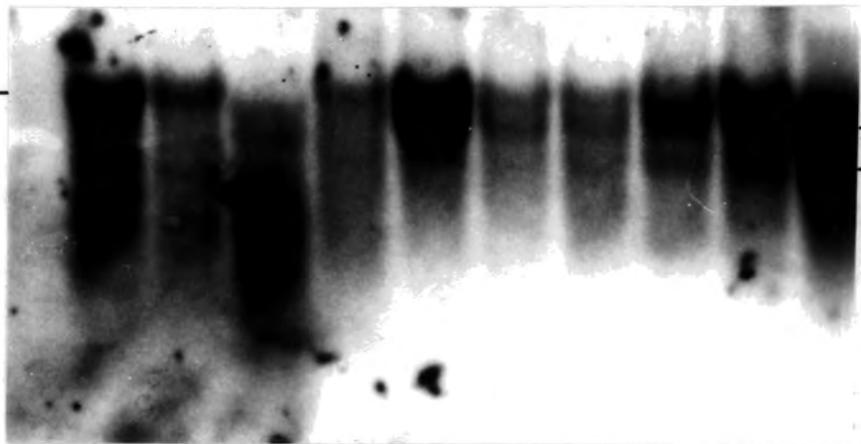


b)

1 2 3 4 5 6 7 8 9 10

1800

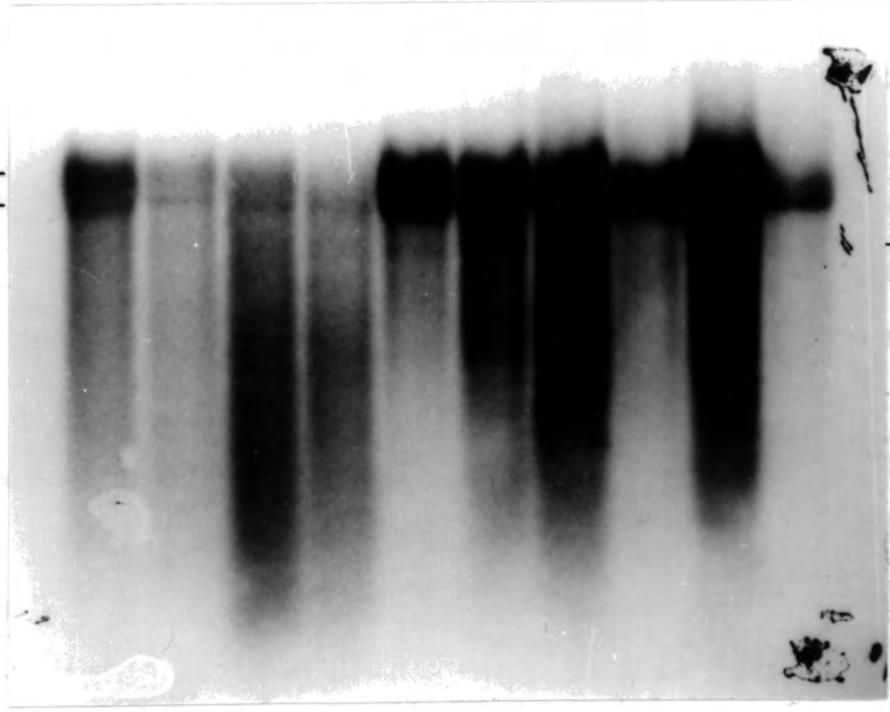
1400
1300_b



c)

1 2 3 4 5 6 7 8 9 10

1880 —
1680_b —



—1350

d)

1 2 3 4 5 6 7 8 9 10

18s —

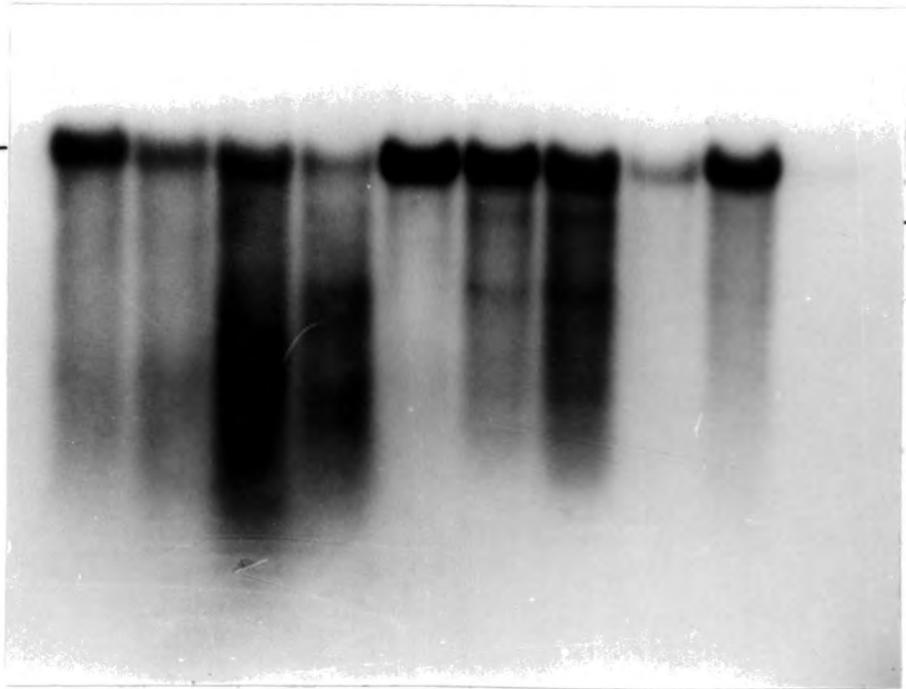
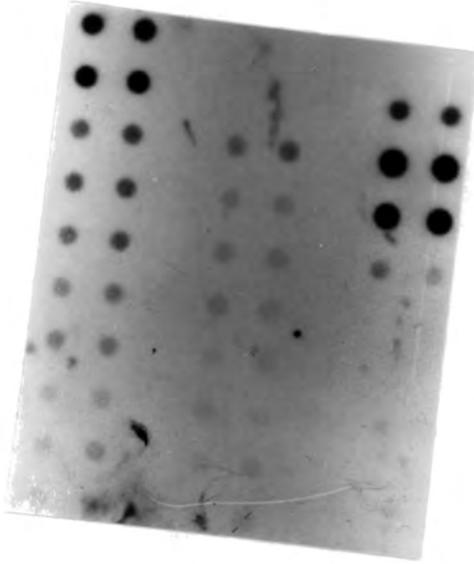


Figure 20. Dot blot analysis of total RNA from developing pea cotyledons, embryonic axes and non-seed tissues, probed with a) *PsMT_A*; b) PCU1 (pea polyubiquitin); c) PJC5.2 (*Leg J*). 2.5ug RNA loaded per well. Pea leaf, stem and root RNA was kindly donated by Dr. I.M. Evans, and *E.coli* rRNA by Dr. D. Bown.

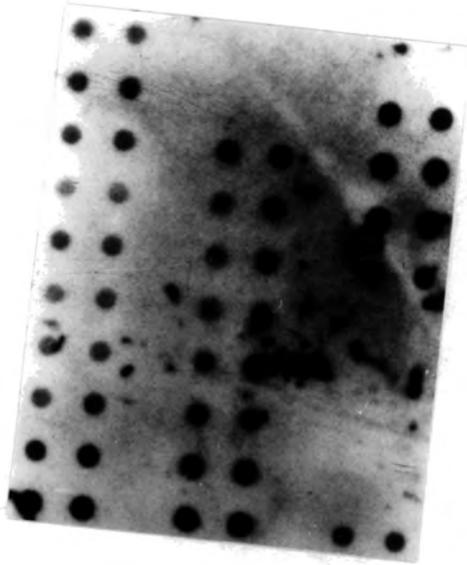
Columns 1 and 2: total RNA from untreated developing cotyledons from 12 to 30DAF (rows 1-10);
Columns 3 and 4: total RNA from developing embryonic axes from 16 to 30DAF (rows 3-10);
Columns 5 and 6: total RNA from non-seed tissues (rows 2-4); *E.coli* rRNA (*Er* - negative control, row 5); 0.1-1ng plasmid insert DNA (*P* -positive control, row 10).

Column	1	2	3	4	5	6
Row	12	12	--	--	--	--
	14	14	--	--	leaf	leaf
	16	16	16	16	stem	stem
	18	18	18	18	root	root
	20	20	20	20	<i>Er</i>	<i>Er</i>
	22	22	22	22	--	--
	24	24	24	24	--	--
	26	26	26	26	--	--
	28	28	28	28	--	--
	30	30	30	30	<i>P</i>	<i>P</i>

a)



b)



c)

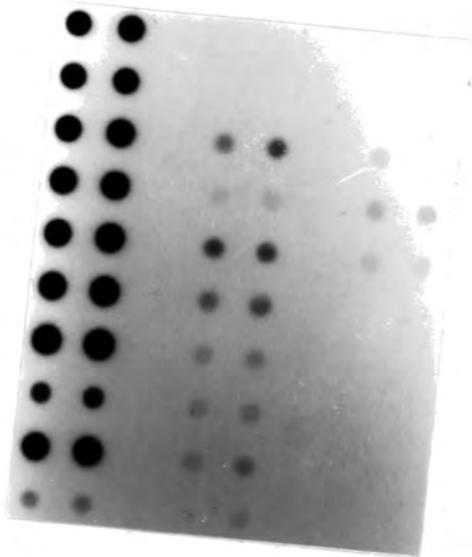
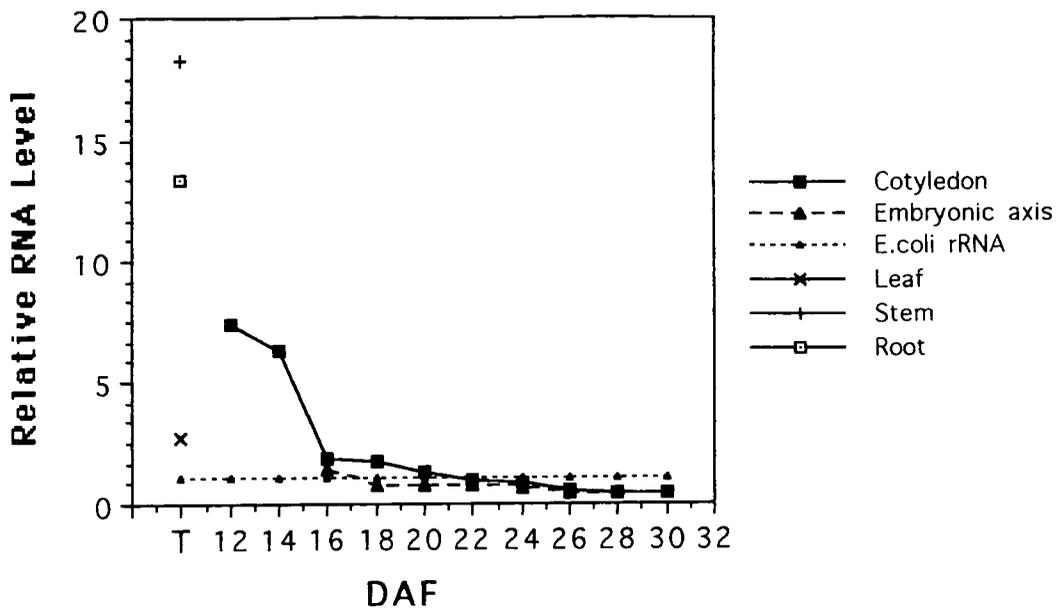
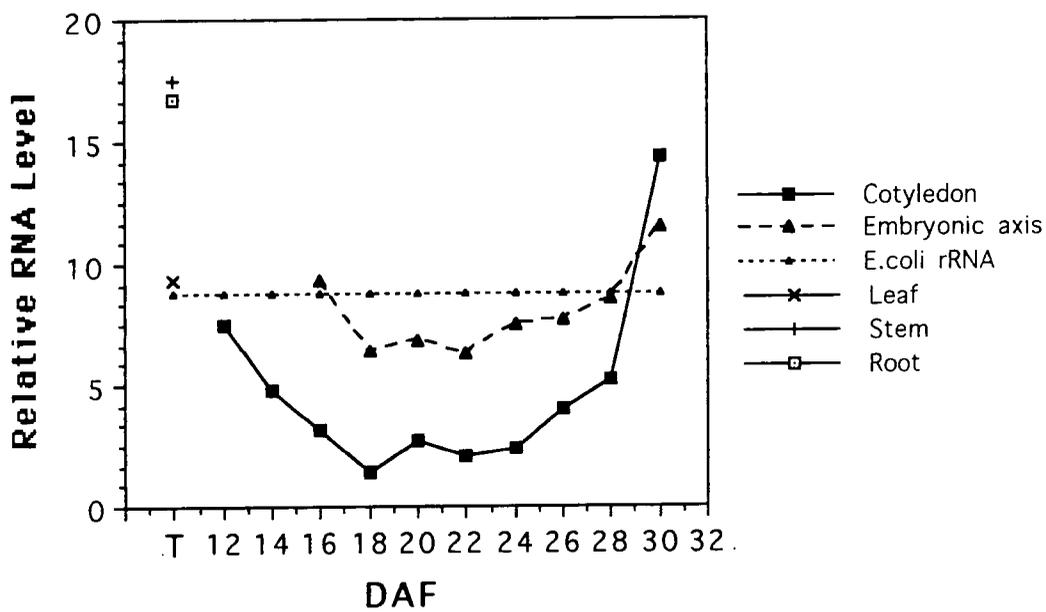


Figure 21. A tentative, unquantitative analysis of the pattern of expression of selected pea protein messages estimated by densitometric analysis of dot blots of total RNA from developing pea cotyledons, embryonic axes and non-seed tissues probed with a) *PsMT_A*; b) PCU1 (pea polyubiquitin); c) PJC5.2 (*Leg J*). T indicates the level of hybridisation of the probes to non-seed tissue total RNA.

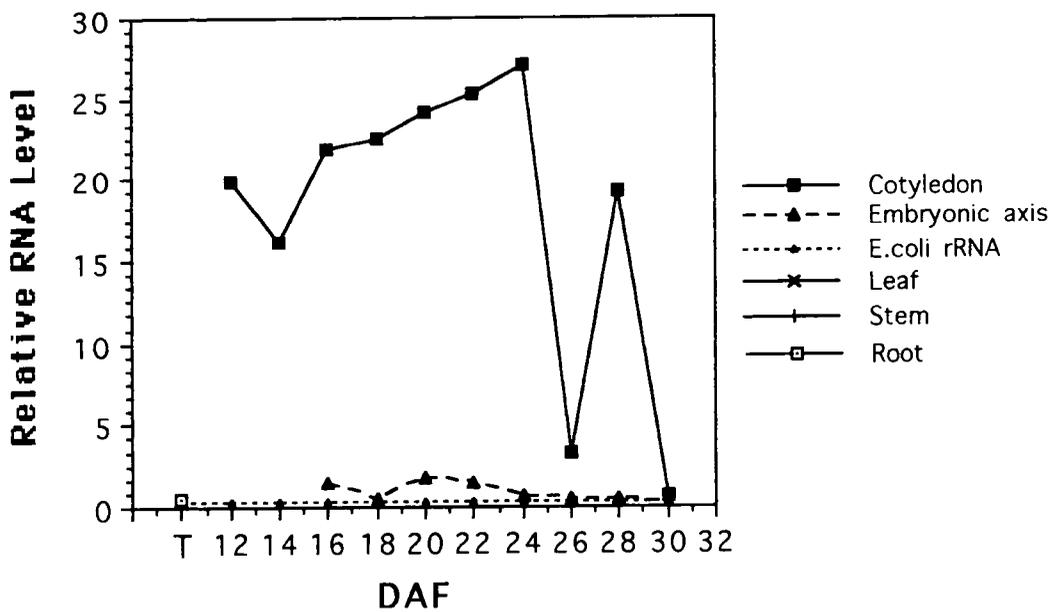
a)



b)



c)



Because of the observed fairly low level of ubiquitin coding messages in the pea seed this blot was exposed for over two weeks and the background hybridisation makes analysis difficult. However, a densitometric analysis of this and other dot blots gives a general indication of the pattern of expression (Fig.21b). It is difficult to know how far to interpret these dot blot results, as most of them fall below the level of hybridisation to the negative control. However, other work (Thompson, 1989) has shown that the *E.coli* rRNA hybridises to a range of pea messages and so it should be taken as an indication of the background level of hybridisation rather than an absolute measure. The densitometric analysis does not agree with the displayed dot blot results at 12DAF - all other Northern and dot blots showed that the level of hybridisation of the ubiquitin probe to 12DAF seed RNA was noticeably greater than that to 14DAF RNA and so the estimated pattern of expression in Fig.21b has been adjusted in accordance with these other readings; the lower level of hybridisation to 12DAF RNA in Fig.20b was presumably caused by underloading or the use of a not fully thawed sample.

The dot and Northern blots suggest three peaks of expression within the time period studied, the first occurring at or before 12DAF in the developing seed. The proportional level of transcripts hybridising to the ubiquitin coding sequence then falls to a trough at 18DAF, before a peak at 20DAF (at the onset of seed desiccation in

most pea batches); although the dot blot results indicate that this peak is fairly low, the Northern blot (which is assumed to be more reliable) suggests a very strong peak of ubiquitin message level at this time. The level of hybridisation to the ubiquitin probe then falls again before a rise to its apparently highest level at 30DAF. As well as the pattern of total ubiquitin-coding message expression in the developing cotyledons, the Northern blot in Fig.19b also shows the changes in expression of several different ubiquitin transcripts. Unfortunately, a degree of degradation in the samples, especially at 16DAF, confuses the results somewhat. However, it is clear from this and other Northern blots that the hybridising band at ca.1.8kb which is present throughout this period of cotyledon development is joined at 20DAF by a second band of hybridisation at ca. 1.4kb. The two bands appear to be equally strongly expressed at 20DAF and are present throughout the rest of the developmental period investigated (which is undergoing desiccation at and after 20DAF), following the pattern of a brief fall then a rapid rise in both bands of hybridisation to a peak at 28 and 30DAF, although it seems that the 1.4kb band achieves prominence at 30DAF. This rise during late development is also seen in a third message, ca.1.3kb, which becomes apparent at 24DAF, although it may be present at low levels at 22DAF. There also appears to be a very faint band of hybridisation at around 1.3kb in the 14DAF sample, although this is not present at 12 or 18DAF and may be due to

degradation; another apparent band at ca.1.35kb in the 12DAF sample was seen on the autoradiograph to be caused by background hybridisation.

5.1.1.3 *Leg J*

An analysis of the pattern of expression of the *Leg J* message shows that it is abundant during late embryogenesis, but the increase in transcript concentration with age is not even. The results obtained from the hybridisation of the *Leg J* probe to a Northern blot of total RNA from developing cotyledons, run very slowly to give good separation of RNA, (Fig.19c) give a detailed picture of *Leg J* subfamily expression, lighting up two or possibly three transcripts which exhibit differential expression. At 12DAF the two main upper bands, at ca.1880 and 1680b, are of approximately equal intensity; both bands fall to an equal level at 14DAF, then increase in intensity at 16DAF. At 18DAF the level of both messages has fallen, the level of hybridisation to the message at 1880b appearing to be lower than that to the band at 1680b, although this may be caused by degradation of the RNA sample. Both bands reach a sustained peak of message level at 20, 22 and 24DAF; on shorter exposures of this blot it can be seen that the lower band becomes more prominent after 20DAF, the upper band appearing to decrease in intensity and to be absent by 28 and 30DAF. Hybridisation of the *Leg J* probe to total RNA at 28DAF indicates a rapid peak in message concentration - this

appears to include both upper and lower bands, although the strength of the signal makes separation of the bands difficult. At 20DAF (and possibly 18DAF) there is the possible expression of a third band at ca.1350b; this signal reaches a maximum at 24DAF (seen on a lower exposure) then seems to disappear.

The screening of dot blots of total cotyledonary RNA with PJC5.2 gives similar results (Figs.20c and 21c), although no noticeable fall in *Leg J* subfamily message level is seen at 18DAF from the dot blot results, which also suggest that the level of hybridisation to the *Leg J* probe reaches a minimum at 30DAF, rather than at 14DAF as suggested by Northern blots.

5.1.1.4 *Leg A*

The temporal pattern of accumulation of the *Leg A* transcript is shown by Northern blot (Fig.19d). As with the *Leg J* results there is a fall in message concentration between 12 and 14DAF, the hybridisation of the probe to the total RNA then increasing noticeably at 16DAF and falling again at 18DAF. Hybridisation then indicates a peak of message level at 20, 22 and 24DAF, which is followed by a rapid drop in transcript level at 26DAF, a second isolated peak at 28DAF and then the minimum, with barely detectable hybridisation of the *Leg A* probe, at 30DAF. Comparing the results from *Leg J* and *Leg A* (Fig.19c and d) it appears that the noticeable increase in the level of hybridisation to the

probe between the earlier (12-18DAF) and later (20-30DAF) RNA samples is of a smaller magnitude when using the *Leg A* probe than when using the *Leg J* probe. Like the results obtained by hybridisation with a probe for *Leg J*, it appears that more than one transcript can be seen at some of the cotyledon ages: one band is apparent throughout development at ca.1800b, but below this at 22 and 24DAF (and possibly 18 and 20DAF) there appear three shorter bands at ca.1400, 1050 and 570 nucleotides. As these do not correspond to the lengths of other transcripts in the *Leg A* subfamily it is assumed that these are specific degradation products, or the result of cross-hybridisation to some other unidentified transcripts.

The results obtained from a similar screen of a dot blot of total cotyledonary RNA are similar to those of the Northern, except for the lack of a noticeable fall in message level at 18DAF, as noted in the *Leg J* screen (Fig.23d).

5.1.2 The Effect of Plant Tissue - Specific Message Levels in the Developing Embryonic Axis and in Non-seed Tissues

The changing levels of the selected messages was examined in total RNA extracted from embryonic axes from the same pea seeds used for cotyledonary RNA extraction. The level of expression in the seed was also compared to that in non-seed tissues of pea.

5.1.2.1 *PsMT_A*

From the dot blot results it appears that the level of the *PsMT_A* transcript in the embryonic axis is only just above hybridisation to the negative control (*E.coli* rRNA) at 16DAF and below that thereafter (Figs.20a and 21a); it follows the pattern of expression found in the cotyledon, but at lower levels, with the messages in both tissues reaching a plateau at 26DAF. Hybridisation of the *PsMT_A* probe to a Northern blot containing 10ug samples of total RNA from embryonic axes was not detectable, presumably due to the low level of the transcript. The level of expression of the *PsMT_A* transcript in non-seed tissues is also indicated in Figs.20a and 21a. Levels of the *PsMT_A* message in the leaf are fairly low, the hybridisation to leaf total RNA appearing to be less than that to 14DAF cotyledonary RNA. However, the message level in the root (from which the gene was originally isolated) and the stem are significantly higher than in the seed tissue at all ages examined.

5.1.2.2 Pea Ubiquitin

The screening of a Northern blot of 10ug samples of embryonic RNA with a polyubiquitin probe gave no clear hybridisation. However, the probe did hybridise to embryonic RNA on a dot blot (Fig.20b), although as with the hybridisation to the total cotyledonary RNA much of the hybridisation is apparently below the level of hybridisation to the negative control. The amount of hybridisation of the

ubiquitin coding sequence to the embryonic RNA seems to be greater than that to the cotyledons at most developmental stages. The expression pattern seen in the embryonic axis, of a fall in ubiquitin coding message level from 16DAF to a slight peak at 20DAF, another fall then a rise to a final peak at 30DAF, seems to correspond well to the pattern seen in the cotyledons, although the peak at 20DAF is much less pronounced and the final level of hybridisation at 30DAF does not seem to be quite as high in the embryonic axis as in the cotyledons.

The hybridisation of the probe to total RNA from non-seed tissues indicates that the maximum level of ubiquitin-coding messages is found in the stem, closely followed by that in the root; both of these organs appear to have higher levels of hybridisation to the ubiquitin probe than that seen at the point of peak expression in the seed (at 30DAF) and considerably higher levels than are seen during most of axis and cotyledon development. The level of hybridisation of the ubiquitin probe to leaf total RNA appears to be lower than that seen in stem and root; further, although the leaf ubiquitin-coding transcript levels seem to be higher than those found in the majority of seed ages examined, appearing to be approximately equal to the level seen at 16DAF in embryonic axes, they are apparently lower than the seed's maximum level of ubiquitin hybridisation to seed tissues in both embryonic axes and cotyledons at 30DAF, when the seed is under considerable water stress.

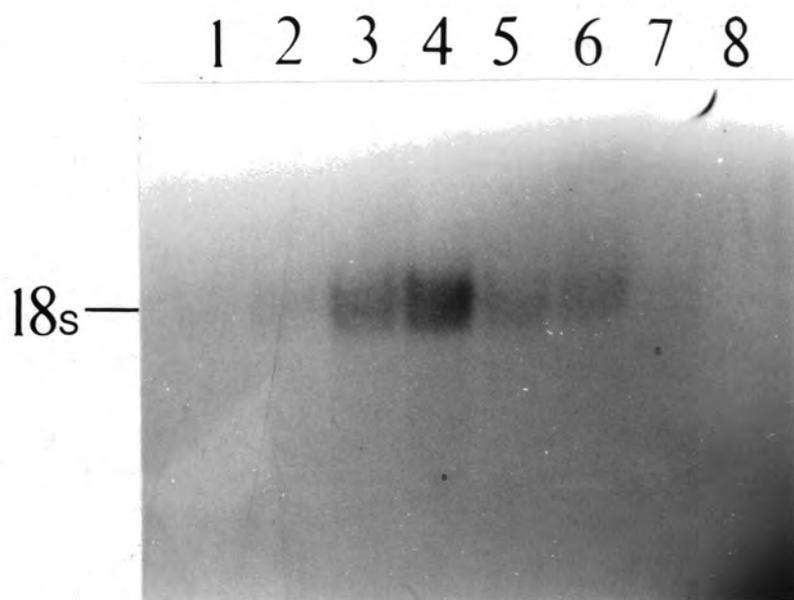
5.1.2.3 *Leg J*

The levels of messages hybridising to the *Leg J* probe in total RNA from embryonic axes are shown in Figs.20c and 22. In both cases the *Leg J* subfamily transcript levels are substantially lower in the axis than in the cotyledons but, as in the cotyledons, they show a degree of differential expression as the axis ages. The Northern blot analysis (Fig.22) indicates that the level of *Leg J* subfamily transcripts increases in the embryonic axes from a low, virtually undetectable level at 16DAF to a maximum at 22DAF, then falls again, being undetectable after 26DAF. The separation of the two 1880b and 1680b bands is not obvious, but the width of the hybridising band suggests that both transcript sizes are present, although this requires confirmation by the use of specific probes. In contrast, the results obtained from hybridisation of the *Leg J* probe to embryonic RNA on a dot blot suggest a peak of expression at 20DAF rather than 22DAF and also indicate a higher level of hybridisation to the 16DAF sample than was indicated in the Northern blot. However, noticeable hybridisation of the PJC5.2 probe to 16DAF embryonic total RNA on other similar dot blots (not shown) was not observed; this discrepancy was possibly caused by unintentional overloading of the 16DAF sample due to inadequate thawing.

The inclusion of non-seed tissues in the dot blot (Fig.20c) demonstrates that the hybridisation of the *Leg J*

Figure 22. Northern blot analysis of total RNA extracted from developing pea embryonic axes probed with PJC5.2 (*Leg J*). 10ug RNA loaded per track.

Tracks 1-8: RNA from 16, 18, 20, 22, 24, 26, 28 and 30DAF embryonic axes, respectively.



message to non-seed tissues is either very low or absent. The hybridisation to both stem and leaf appears to be less than that to the negative control (*E.coli* rRNA). The signal from root RNA seems slightly stronger than that from the negative control and the lower levels found in the embryonic axis; however, it is lower than the levels found in the cotyledons at all developmental ages examined.

5.1.2.4 *Leg A*

As with *Leg J*, the level of *Leg A* message in the embryonic axis is very much lower than in the cotyledons and a clear result from the hybridisation of the *Leg A* probe (PRC3.1) to a Northern blot of embryonic axis total RNA (10ug/track) was not obtained. Dot blot analysis gave a very similar pattern of hybridisation of total RNA to the *Leg A* message as that to the *Leg J* probe, with an overall low level and a slight maximum at 20DAF, although unlike the *Leg J* results this was not continued to 22DAF; there was no appreciable hybridisation of the probe to the 16DAF sample. The hybridisation of the *Leg A* probe to total RNA from leaf, stem and root tissue in all cases seemed to be less than the background hybridisation to the negative control (*E.coli* rRNA) and the hybridisation levels to all cotyledon ages examined.

5.1.3 The Effect of Premature Desiccation on Specific Message Levels in Immature Cotyledons

Immature pea pods were excised from the plant and exposed to a drying treatment to try to separate the effects of desiccation on the developing seed from the message level changes observed in the normal developmental programme, and so investigate the effects of drying on the seed. Two methods of premature desiccation of 14DAF cotyledons were examined: air drying for 24, 48 and 72 hours and drying over silica gel (silica drying), presumed to give more rapid desiccation, for 48 and 72 hours and 8 days. The main problem with these treatments arose from the controls. To ensure that the presence of water was the only difference between treated and control seeds, control pods were taken from the plant and placed with their peduncles in sterile water for 24, 48 and 72 hours, to parallel the air drying treatments. However, all of the total RNA extracted from such controls was substantially degraded. It is possible that the isolation from the plant nutrients and hormones and the exposure to water induces the seed to try to germinate; or that the shock of removal from the plant induces degradation processes which require water and so are not seen in the desiccation treatments. Whatever the cause, an accurate comparison of the different levels of transcript message between treatments and controls was difficult to obtain. On most Northern blots the control was loaded in larger amounts so that the intensity of the ribosomal bands

was comparable to that of the treated samples. This resulted in a larger degree of hybridisation to the degraded samples than could be expected, confusing the results. When an equal weight of control RNA was used, the degradation resulted in a smeared signal which is hard to compare with the treated samples. It was considered more sensible, therefore, to compare the message level found in treated samples to that in cotyledons of the same age harvested straight from the plant, assuming that the slight difference in temporal expression found between most batches will not affect the results substantially.

5.1.3.1 *PsMT_A*

Premature desiccation of 14DAF cotyledons appears to have a negative effect on the level of the *PsMT_A* transcript (see Figs.23a and 24a). After 24 hours of air drying the message level in the treated seed, which has a biological age of 15DAF, is very similar to that of the 16DAF untreated sample; however, the message level drops on further drying and after 72 hours the level appears to have been brought down to that normally seen at around 24DAF on the plant, while the cotyledons are only at 17DAF. The silica drying gives a similar effect to the air drying at 48 and 72 hours, while 8 days of silica drying apparently causes the transcript level to fall to that usually seen around 26DAF when the cotyledons are only at 22DAF.

Figure 23. Dot blot analysis of total RNA extracted from pea cotyledons from untreated developing pods and from pods isolated from the plant and exposed to premature desiccation.

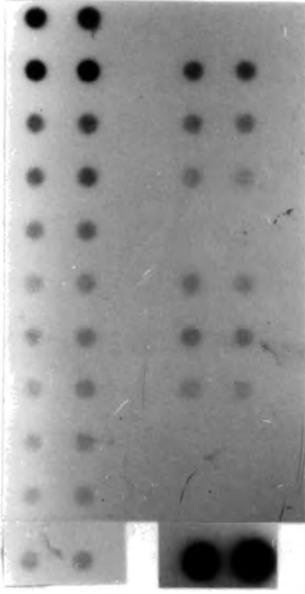
2.5ug loaded per well. Blots are probed with a) *PsMT_A*; b) PCU1 (pea polyubiquitin); c) PJC5.2 (*Leg J*); d) PRC3.1 (*Leg A*).

Columns 1 and 2: total RNA from untreated developing cotyledons from 12 to 30DAF (rows 1-10); *E.coli* rRNA (*Er*, row 11);

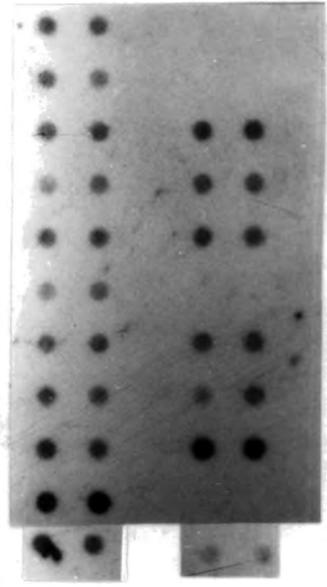
Columns 3 and 4: total RNA from cotyledons from pods taken from the plant at 14DAF either air dried (A, rows 2-4) or silica dried (S, rows 6-8) for 24, 48, 72 hours or 8 days; 0.1-1ng plasmid insert DNA (P, row 11).

Column	1	2	3	4	
Row	1	12	12	---	---
	2	14	14	A24	A24
	3	16	16	A48	A48
	4	18	18	A72	A72
	5	20	20	---	---
	6	22	22	S48	S48
	7	24	24	S72	S72
	8	26	26	S8	S8
	9	28	28	---	---
	10	30	30	---	---
	11	<i>Er</i>	<i>Er</i>	P	P

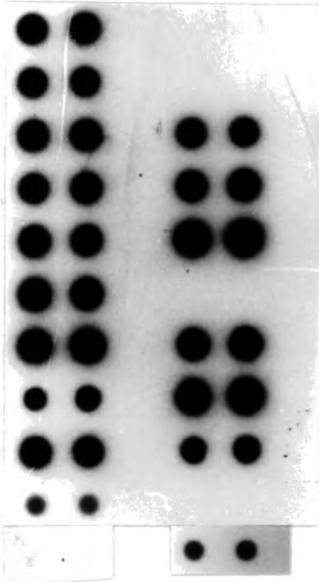
a)



b)



c)



d)

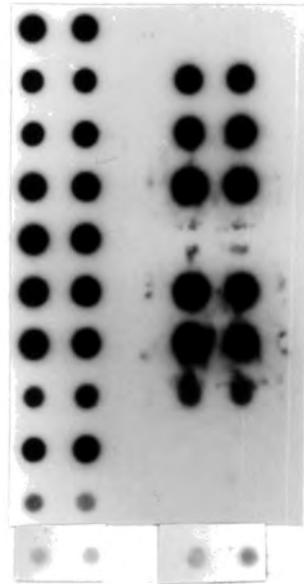
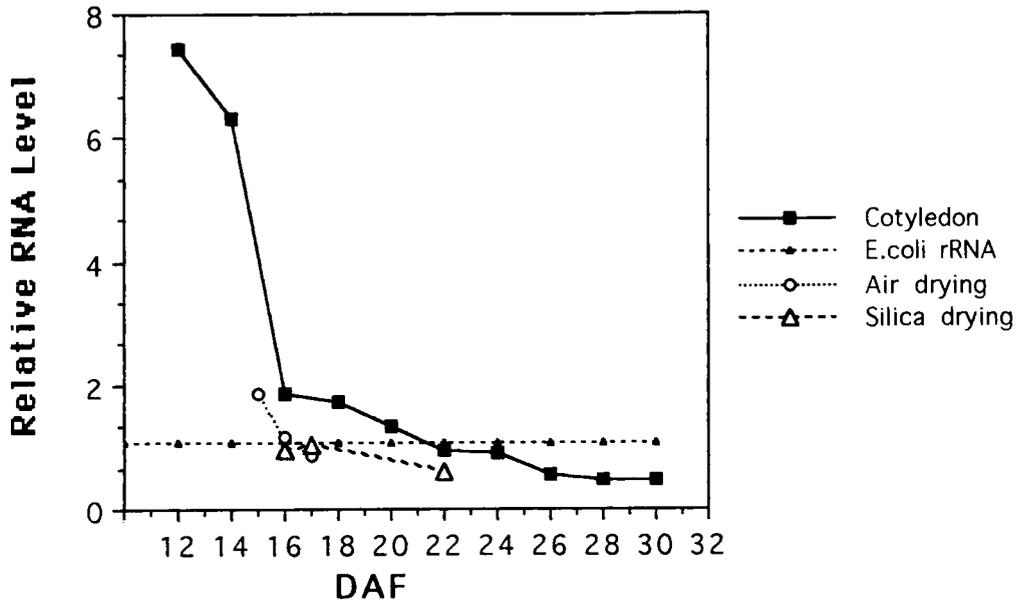
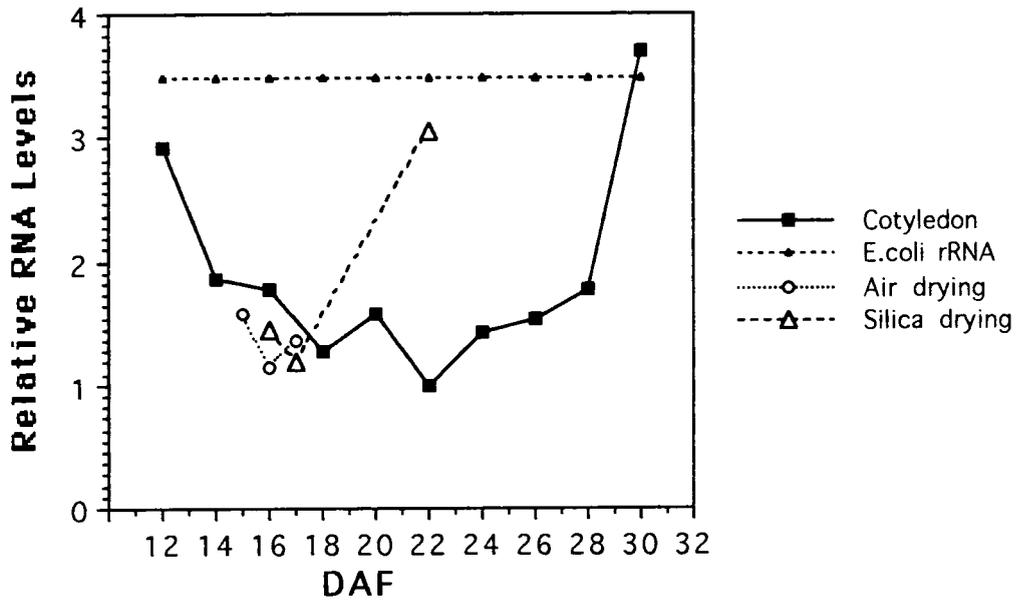


Figure 24. A tentative, unquantitative analysis of the pattern of expression of selected pea protein messages estimated by densitometric analysis of dot blots of total RNA from pea cotyledons from untreated developing pods and from pods (from a separate batch) isolated from the plant and exposed to premature desiccation, probed with a) *PsMT_A*; b) PCU1 (pea polyubiquitin); c) PJC5.2 (*Leg J*); d) PRC3.1 (*Leg A*).

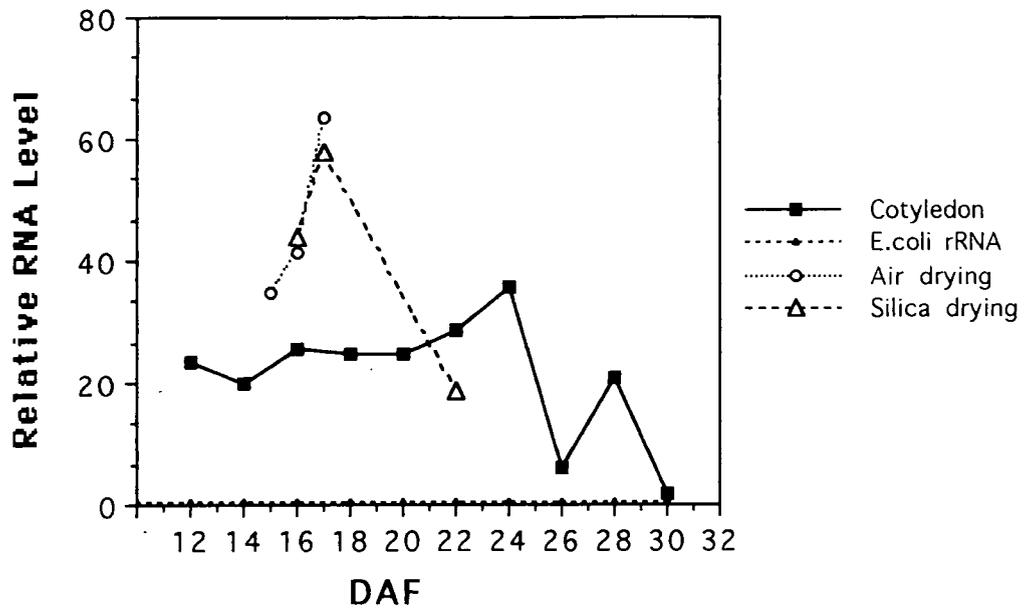
a)



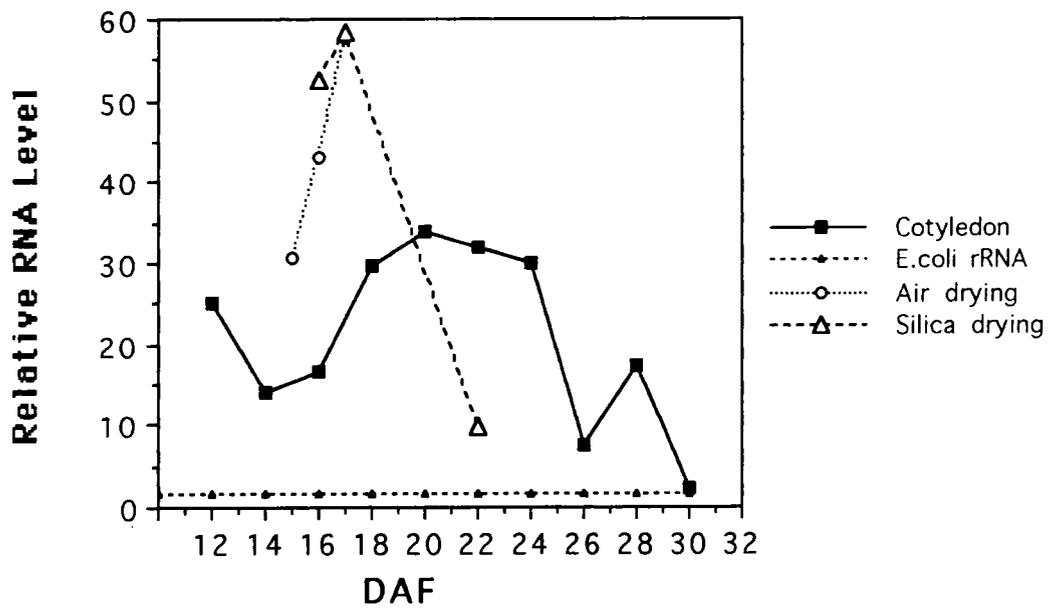
b)



c)



d)



5.1.3.2 Pea Ubiquitin

Hybridisation of the pea polyubiquitin probe was barely detectable on Northern blots of control and desiccation treatments containing 10ug total RNA per sample: very faint hybridisation was seen only to the 8 day silica dried sample, at around 18S; thus the ubiquitin-coding messages do not appear to be induced to a high level by premature desiccation of 12DAF cotyledons. However, hybridisation of the probe to a dot blot (Figs.23b and 24b) indicates that although 24 to 72 hours of desiccation, both by air and silica, gives message levels similar to or slightly below those seen in the untreated plant, there is a noticeable increase in the level of hybridisation to the probe after 8 days silica drying, reaching a level of hybridisation usually seen only between 28 and 30 DAF at a biological age of 22DAF, and significantly higher than the hybridisation to the negative control.

5.1.3.3 *Leg J*

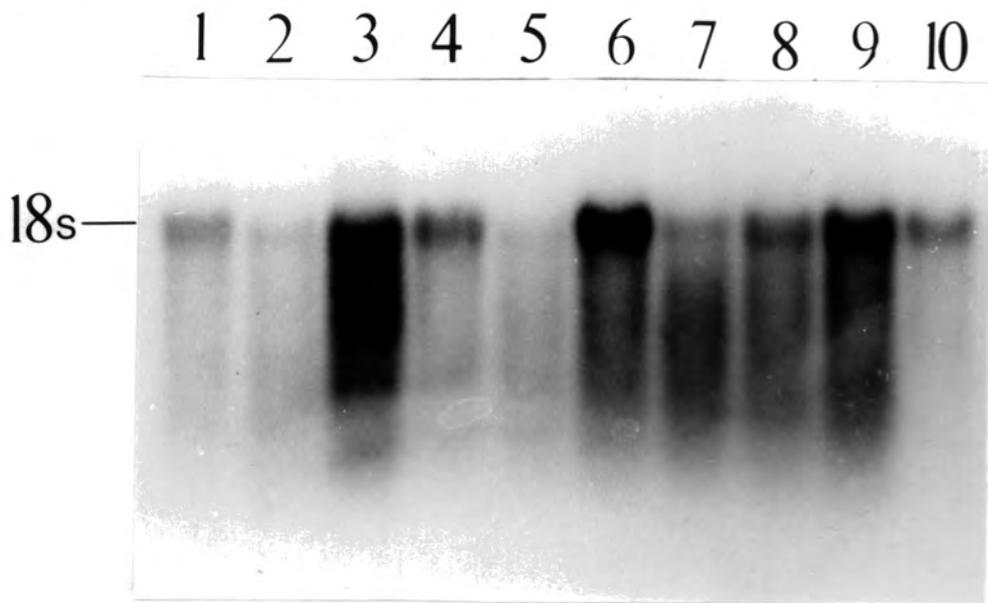
The imposition of premature desiccation on developing pea cotyledons gave a very clear increase in the level of message hybridising to the *Leg J* probe. From the Northern blot (Fig.25a) it appears that the level of hybridisation to the 14DAF sample is greater than that to the 24 hour air-dried sample, although other similar Northern blots did not indicate such a large difference between the two samples. Nevertheless, it is clear that continued air

Figure 25. Northern blot analysis of total RNA extracted from immature pea cotyledons from pods which were either untreated, prematurely dried by air or silica drying, or incubated in water. Blots were probed with: a) PJC5.2 (*Leg J*); b) PRC3.1 (*Leg A*) and contained total RNA from:

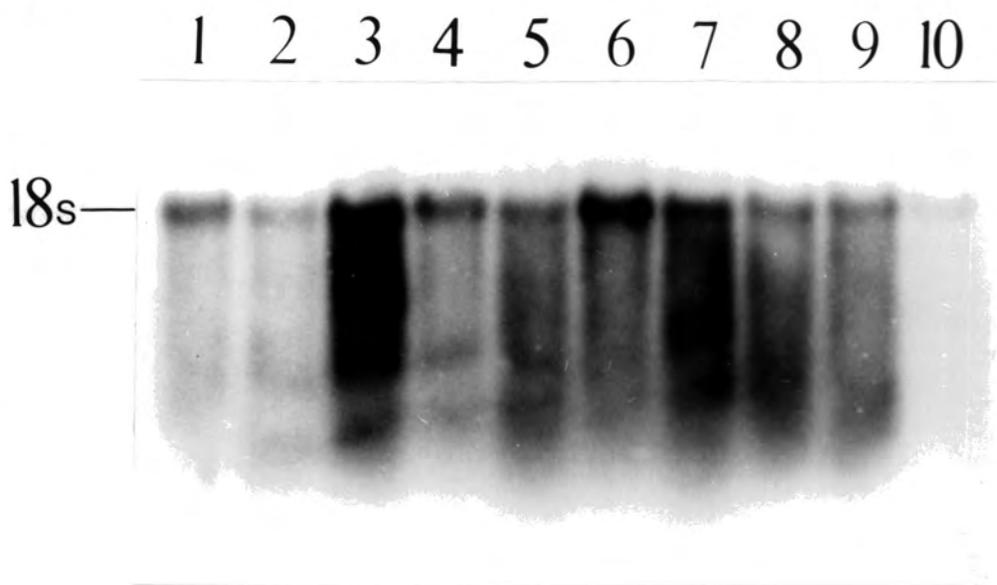
- Track 1: 14DAF cotyledons (untreated)
- Track 2: air dried, 24 hours
- Track 3: water control, 24 hours
- Track 4: air dried, 48 hours
- Track 5: water control, 48 hours
- Track 6: air dried, 72 hours
- Track 7: water control, 72 hours
- Track 8: silica dried, 48 hours
- Track 9: silica dried, 72 hours
- Track 10: silica dried, 8 days

(Untreated and dried samples are loaded at 10ug RNA per track; the water controls were loaded to give 21S and 18S ribosomal bands of approximately equal intensity to the dried samples).

a)



b)



drying gives a noticeable increase in hybridisation to the *Leg J* subfamily probe. This pattern is repeated in the results from 48 and 72 hour silica drying, with the message level then falling substantially after 8 days silica drying. The controls are degraded, but it is clear that, with the exception of the 24 hour air dried sample, the level of hybridisation to the dried samples is considerably greater than that to the water controls when comparing the area at 18S. The hybridisation of PJC5.2 to total RNA from developing and prematurely dried cotyledons on a dot blot (Figs.23c and 24c) appears to give results which agree very well with the Northern blot results.

From the Northern blot in Fig.25a it is difficult to separate the two bands (1880b and 1680b). However, similar Northern blots taken from gels which were electrophoresed at a slower rate indicated that it is the lower 1680 band which responds first to air drying, the 1880 band appearing after 48 hours and the two bands reaching an almost equal intensity after 72 hours air drying. The same pattern is found with silica drying, with the 1880b band showing the greatest decrease in hybridisation to the probe after 8 days silica drying.

5.1.3.4 *Leg A*

A screen of the same Northern blot used in Fig.25a with a probe for *Leg A* shows that premature desiccation leads to an increase in the level of *Leg A* subfamily messages which

is slightly similar to but of a much smaller magnitude than the increase seen with the *Leg J* subfamily messages. Although the level of hybridisation to the PRC3.1 probe increases noticeably from 24 to 72 hours air drying, there is also a noticeable increase in the level of hybridisation to the 48 and 72 hour water controls. Thus although the dot blot results (Fig.23d) indicate significantly higher levels of the *Leg A* subfamily transcripts after premature desiccation, the increase does not seem to be a sole result of desiccation. In contrast, there appears to be no significant increase in the level of *Leg A* subfamily messages between 48 and 72 hours silica drying, while further silica drying gives a fall in message level to a barely detectable level, well below the concentration usually seen at this seed age (22DAF).

5.1.4 The Effect of Exogenous ABA on Specific Message Levels in Immature Cotyledons

Immature pods were excised from the plant and treated with concentrations of ABA from 10^{-4}M to 10^{-6}M to assess the effects of applied ABA on specific message levels in the developing seed.

5.1.4.1 *PsMT_A*

The application of exogenous ABA appeared to have no effect on the expression of the *PsMT_A* message.

Hybridisation to the *PsMT_A* message was seen on a Northern blot of the ABA treatments, but the strength of hybridisation was no greater than that to the 14DAF untreated sample or to the water control and showed no variation with increasing ABA concentration. Further, there was only very faint hybridisation to the samples treated for 72 hours, giving the same intensity as the water and untreated controls; this agrees with the observed low level of *PsMT_A* transcript at this developmental stage and indicates that this range of ABA concentrations neither induces nor reduces levels of the *PsMT_A* transcript.

5.1.4.2 Pea Ubiquitin

ABA appeared to have no noticeable effect on the expression of ubiquitin-coding transcripts when applied exogenously to developing cotyledons. Long exposure of a Northern blot of control and ABA-treated RNA (20ug) did show hybridisation of the polyubiquitin probe to the seed RNA, but, as with *PsMT_A*, the hybridisation of the probe to treated samples was comparable to that to untreated samples and controls.

5.1.4.3 *Leg J*

Treatment of developing cotyledons with exogenous ABA induces a definite increase in the level of *Leg J* transcript, as demonstrated by Northern blot analysis (Fig. 26a). Although the Northern blot does not distinguish

Figure 26. Northern blot analysis of total RNA extracted from pea cotyledons from untreated pods and from pods (from a separate batch) isolated from the plant and incubated with varying concentrations of ABA and water.

a) Northern blot probed with PJC5.2 (*Leg J*);

b) Blot a) stripped and reprobed with PRC3.1 (*Leg A*).

Track 1: 14DAF cotyledons (untreated)

Track 2: water control, 24 hours

Track 3: 10^{-6} M ABA, 24 hours

Track 4: 10^{-5} M ABA, 24 hours

Track 5: 10^{-4} M ABA, 24 hours

Track 6: water control, 72 hours

Track 7: 10^{-6} M ABA, 48 hours

Track 8: 10^{-5} M ABA, 48 hours

Track 9: 10^{-4} M ABA, 48 hours

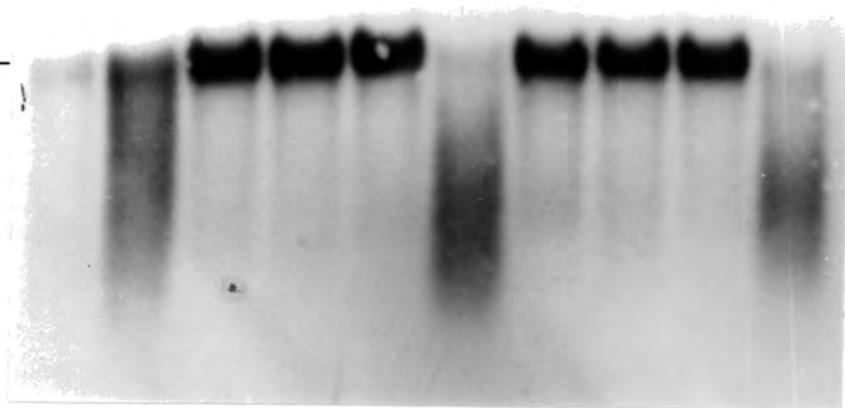
Track 10: 18DAF cotyledons (untreated).

(All samples were loaded at 10ug RNA per track except the water controls, which were loaded to give ribosomal bands of approximately the same intensity as the other samples).

a)

1 2 3 4 5 6 7 8 9 10

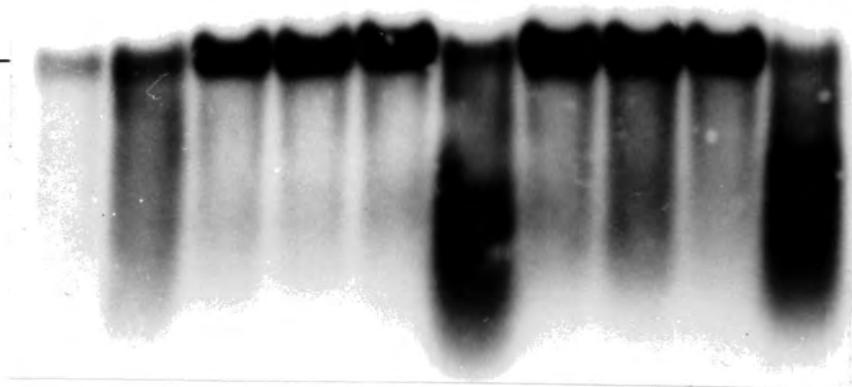
18s—



b)

1 2 3 4 5 6 7 8 9 10

18s—



between any separate bands, it is clear that the level of *Leg J* transcript is significantly higher in cotyledons treated with ABA than in either the water controls or the total RNA samples from cotyledons left on the plant. The water controls, taken from the same batch of plants as the ABA-treated samples, were again substantially degraded. There is a noticeable rise in hybridisation to the *Leg J* probe between the 14DAF sample and the 24 hour water control, but the level of hybridisation to the 24 hour ABA-treated samples is significantly higher than that to the water control in the area at 18S. The levels of hybridisation to the 72 hour water control and the 18DAF samples appear to be approximately equal at 18S and considerably lower than that seen to the 72 hour ABA-treated samples. Given the strength of hybridisation to the ABA-treated samples it is very difficult to ascertain whether there is any significant difference between the three concentrations of ABA and the two incubation times; shorter exposures also did not suggest any significant difference between the six ABA-treated samples, although by eye it appears that the longer exposure to ABA gave slightly lowered levels of hybridisation to the *Leg J* probe.

5.1.4.4 *Leg A*

Hybridisation of the *Leg A* probe to the same Northern blot screened with *Leg J* (Fig.26b) indicates that, like *Leg J*, incubation of pods with ABA leads to an increase

in cotyledonary *Leg A* subfamily message level. However, although the hybridisation of the probe to the ABA-treated samples is much higher than that to the 14DAF sample, the hybridisation to the water controls and the 18DAF sample is also fairly high. When comparing the area at 18S it still appears that ABA treatment gives an increased level of hybridisation to the probe but, as with the premature desiccation results, the apparent increase in *Leg A* subfamily message levels upon exposure to ABA when compared to the controls is less than that seen with the *Leg J* subfamily levels. As before, it is difficult to assess on this overnight exposure whether the three concentrations and two lengths of treatments give any significant differences in the effect on message level; however, unlike the *Leg J* results, it appears by eye that continued exposure to the two lowest levels of ABA (10^{-6} M and 10^{-5} M) gives a slight increase in the level of hybridisation to the probe.

5.1.5 Em Message Levels in Pea

No specific hybridisation was detected when Northern blots of total pea cotyledon RNA were probed with a subclone of the wheat Em cDNA (PMuge K1a), only a degree of non-specific binding to the 18S and 23S ribosomal bands. The same result was seen with Northern blots of prematurely desiccated and ABA-treated cotyledons, indicating that the pea seed does not contain a message showing sufficient homology to the Em cDNA to allow hybridisation.

5.1.6 Cowpea Trypsin Inhibitor Levels in Pea

A probe for the *CpTI* gene failed to bind strongly to total RNA from treated and untreated cotyledons, embryonic axes or non-seed tissues when hybridised at 42°C with formamide and washed with 5xSSC, 0.1%SDS, twice at room temperature and once at 42°C. Densitometric analysis of the faint background hybridisation showed that the low levels of binding of the probe to all treated, untreated and non-seed samples was approximately the same and was no greater than the hybridisation to the negative control (*E.coli* rRNA).

5.1.7 Hybridisation of the PM46 cDNA Insert to Pea RNA

It was very difficult to observe clear hybridisation of the PM46 cDNA insert, isolated from a 24,26,28DAF library, to total RNA from developing pea cotyledons. A signal was detected on a Northern blot containing 30ug RNA per track and exposed for over two weeks (Fig.27). However, the background hybridisation is extremely high and the hybridisation to the RNA in the area of 18S, (with a faint level of hybridisation also seen to the 23S area) did not appear to show any particular pattern of expression. Hybridisation of the probe to total RNA in dot blots produced similar results with a high background, faint hybridisation to the total RNA and no discernible pattern of expression. No hybridisation was detected to a Northern blot

Figure 27. Northern blot analysis of total RNA from developing pea cotyledons, probed with PM46. 30ug total RNA was loaded per track.

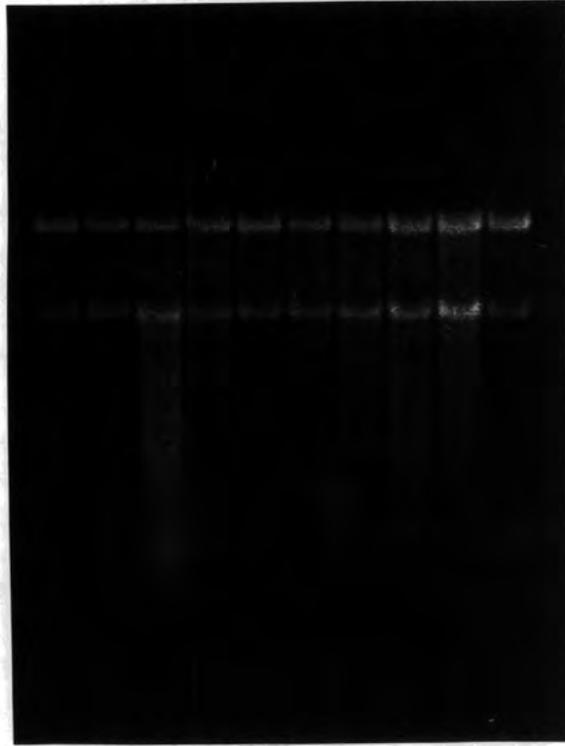
a) Formaldehyde gel electrophoresis of total RNA from developing cotyledons:

Tracks 1-10: total RNA from 12, 14, 16, 18, 20, 22, 24, 26, 28 and 30DAF cotyledons respectively;

b) Northern blot analysis of gel (a) probed with PM46.

a)

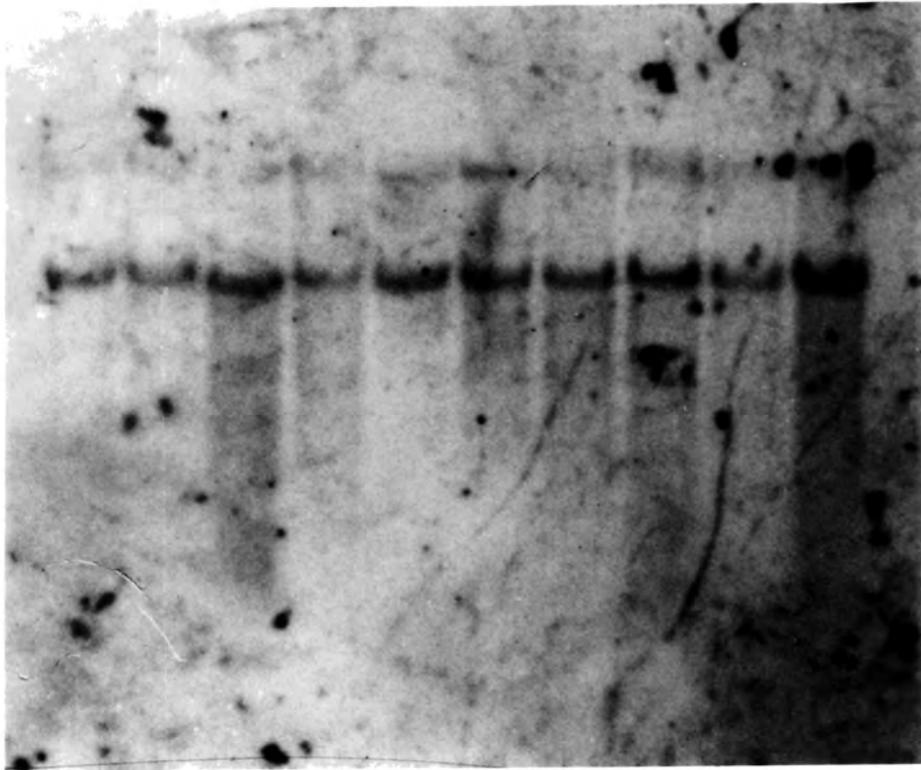
1 2 3 4 5 6 7 8 9 10



b)

25s—

18s—



of 10ug samples of embryonic axis RNA or to a blot of desiccation-treated cotyledons (20ug RNA per track). Screening of a Northern blot of ABA-treated cotyledons and controls (30ug/track) again produced a signal of hybridisation to the 18S area and fainter hybridisation to the 23S area and a high level of background hybridisation.

5.1.8 The Effect of Seed Age and Tissue on the Total and Albumin Protein Fractions in Developing Peas

Figs. 28a and 28b show the total albumin protein fraction from 14-32DAF cotyledons and embryonic axes respectively, each sample containing 40mg fresh weight/ml. In both the cotyledons and the embryonic axes the total albumin levels increase steadily with age (loaded on a fresh weight basis), the major bands occurring at ca.26K and ca.12K in the cotyledons (corresponding to the major and low MW albumins respectively). The protein mass appears to be more evenly distributed between the bands in the axis; the pattern of bands in the albumin extracts from embryonic axes is similar to that in the cotyledons, but the major albumin at 26K is not as noticeable. Further, a band of protein at ca.18K appears in the embryonic axis at 24DAF, remaining fairly clear until the level falls at 32DAF; this band is not conspicuous in the cotyledons or in the embryonic axis at earlier developmental stages and appears to be the only prominent albumin band which does show a sudden increase in density as the seed ages. Other bands which are fairly

Figure 28. Polyacrylamide gel electrophoresis of albumin and total protein fractions from developing and desiccated cotyledons and embryonic axes. Protein was extracted at 40mg/ml and electrophoresed under reducing conditions.

a) Albumin proteins from developing cotyledons

Tracks 1 and 12: SDS-7 size marker*

Tracks 2-11: albumin proteins from 14, 16, 18, 20, 22, 24, 26, 28, 30 and 32DAF cotyledons;

b) Albumin proteins from developing embryonic axes

Tracks 1 and 12: SDS-7 size marker

Tracks 2-11: albumin proteins from 14, 16, 18, 20, 22, 24, 26, 28, 30 and 32DAF embryonic axes;

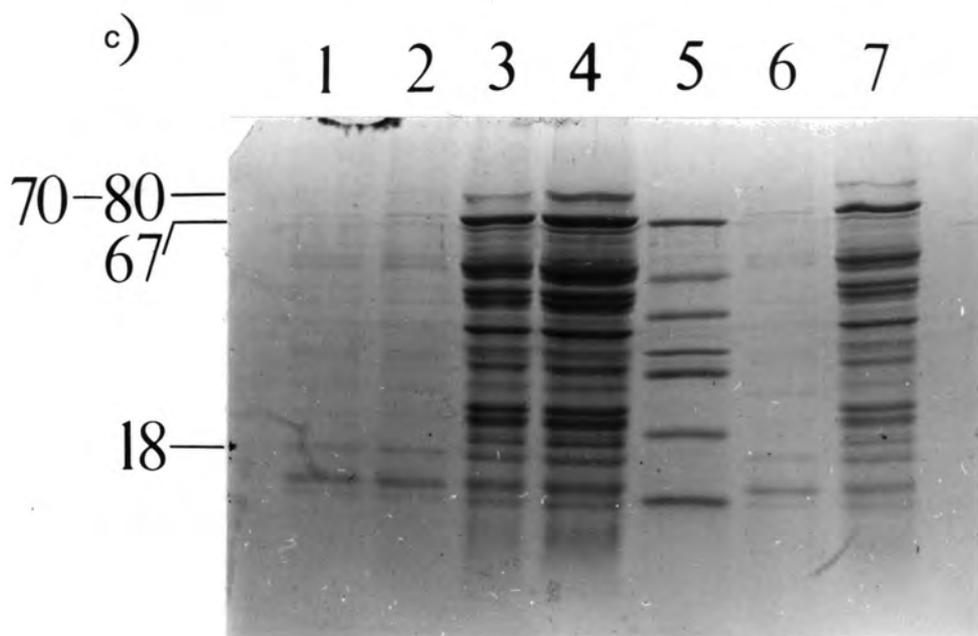
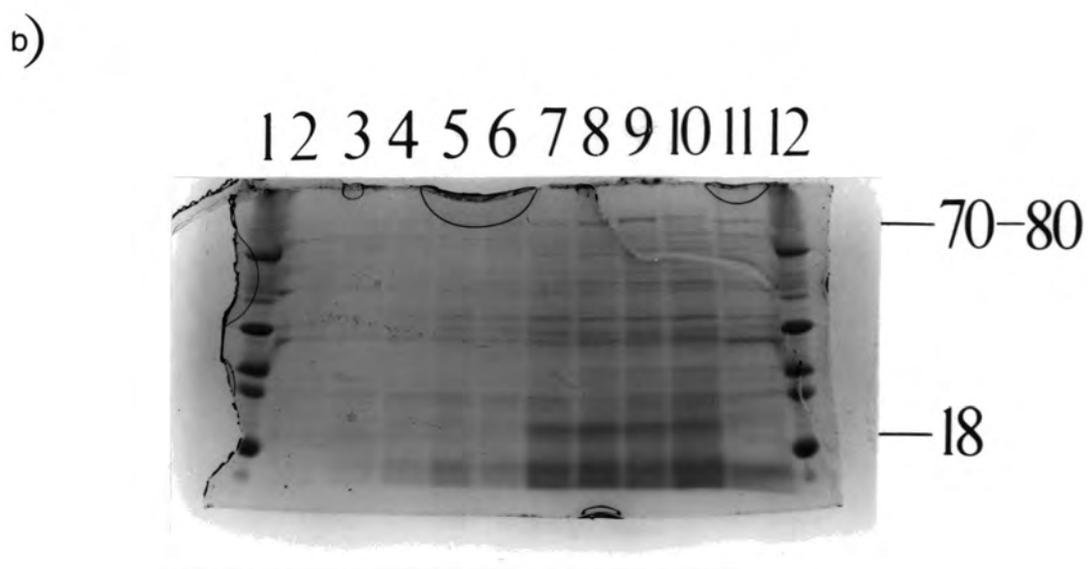
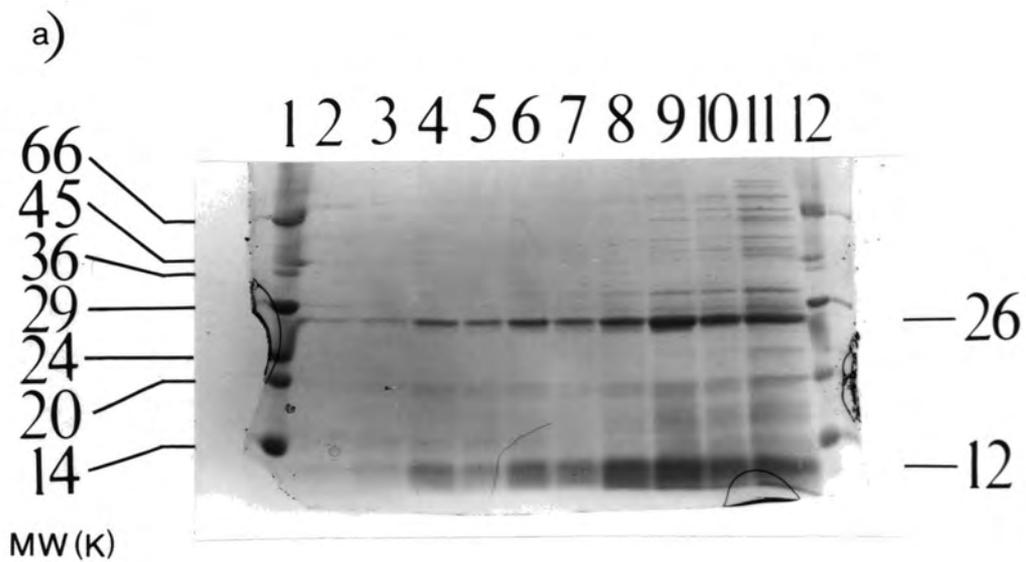
c) Total protein from desiccated cotyledons and embryonic axes

Tracks 1, 2 and 6: total protein from desiccated embryonic axes (10, 20 and 5ul loaded respectively)

Tracks 3, 4 and 7: total protein from desiccated cotyledons (10, 20 and 5ul loaded respectively)

Track 5: SDS-7 marker.

*SDS-7 size marker: 66K, 45K, 36K, 29K, 24K, 20.1K, 14.2K
(Marker and protein band sizes shown are approximate)



concentrated in the axis are at ca.70-80K, whose level appears to increase to 28DAF then fall and which is not obvious in the albumin samples from cotyledons; and at ca.32K, the pea cotyledons containing a band in a similar area.

The ca.26K major albumin observed in cotyledons and embryonic axes is also visible in Fig.29a which shows total protein extracts from cotyledons from 16-32DAF. Again, its level appears to increase with age on a fresh weight basis but forms only a small part of the total protein mass.

A comparison of total protein samples from cotyledons and embryonic axes from completely desiccated pea seeds (shown in Fig 28c) confirms that the dry embryonic axis contains protein bands which are absent or present at lower levels in dry cotyledons; they include a band at ca.18K, ca.27K and possibly at ca.22K, ca.67K and ca.70-80K. None of these bands are particularly prominent in the dry axis total protein fraction.

5.1.9 The Effect of Seed Age and Premature Desiccation on Levels of the Molecular Chaperone Chaperonin 60

A Western blot of total cotyledonary protein probed with chaperonin 60 antiserum indicates that the presence of chaperonin 60 continues throughout seed development from 16 to 32DAF at a level comparable to that in the leaf positive control (Fig.29). From the results it appears that the

Figure 29. Western blot analysis of total protein extracted from developing cotyledons and leaf tissues reacted with antiserum for chaperonin 60.

a) Polyacrylamide gel electrophoresis of total protein from developing pea cotyledons and leaf tissues, extracted at 40mg/ml and run under reducing conditions

Track 1: SDS-7 marker

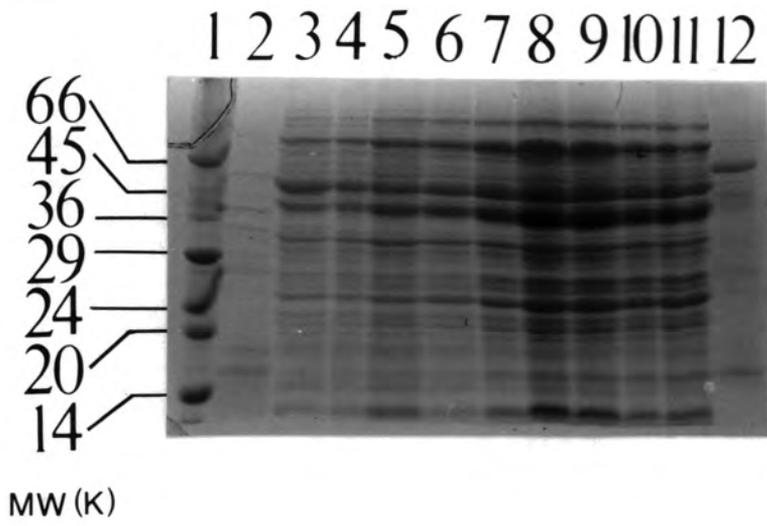
Tracks 3-10: total protein from 16, 18, 20, 22, 24, 26, 28, 30 and 32DAF cotyledons

Track 11: total protein from fresh pea leaf;

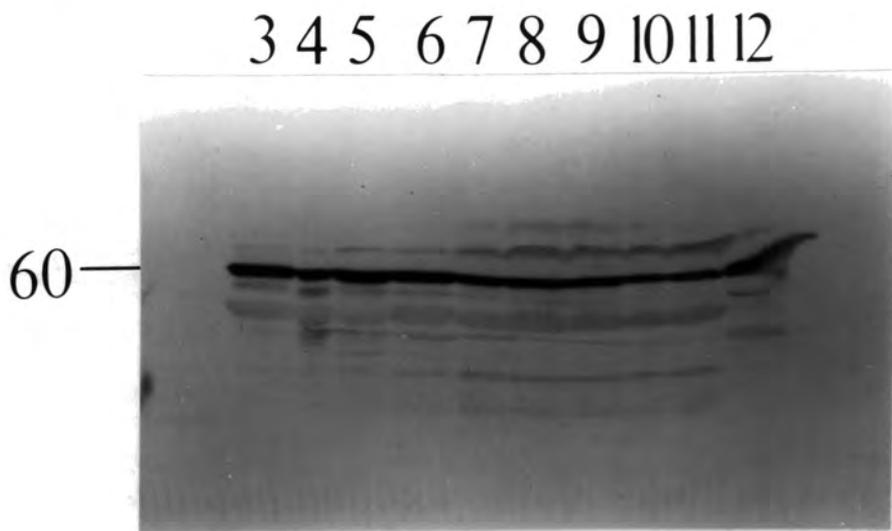
b) Western blot analysis of gel a) reacted with antiserum for chaperonin 60 at a dilution of 1/5000. Antiserum was kindly donated by Dr. L. Barnett and Prof. R.J. Ellis.

(Marker and protein band sizes shown are approximate)

a)



b)



amount of this protein decreases slightly after 24DAF during desiccation, but the signal at 60K remains strong.

Fig.30 shows a Western blot, probed with chaperonin 60 antiserum, of total cotyledonary protein from seeds which have been exposed to a desiccation treatment (air drying of isolated pods) for 24, 48 and 72 hours, compared to total protein extracted from cotyledons of peas of the same biological age (15, 16 and 17DAF) and batch isolated straight from the plant. It appears that the levels of chaperonin 60 are slightly reduced after 48 and 72 hours desiccation compared to the untreated samples, despite a higher concentration of protein in the prematurely desiccated samples (Fig.30a).

5.1.10 The Effect of Premature Desiccation on Cotyledonary Total Protein

Fig.30a shows total protein extracted at 40mg/ml from cotyledons taken from 14DAF pods which were excised from the plant and air dried for 24, 48 and 72 hours, together with protein samples taken from untreated cotyledons from the same plants at 15, 16 and 17DAF to give comparable biological ages. The protein levels increase slightly with age from 15 to 17DAF; however, protein samples from 24 hour air dried cotyledons appear to give higher protein levels, some of the bands continuing to increase in intensity with continuing desiccation, in particular bands at 50K, 33K and 19K (vicilin), ca.40K and ca.20K (major and minor legumin

Figure 30. Western blot analysis of total protein from pea cotyledons from immature and prematurely desiccated pods reacted with antiserum for chaperonin 60.

a) Polyacrylamide gel electrophoresis of total protein from immature cotyledons from untreated and air dried pods, extracted at 40mg/ml and run under reducing conditions.

Track 1: 14DAF cotyledons, air dried for 24 hours

Track 2: 15DAF cotyledons, untreated

Track 3: 14DAF cotyledons, air dried for 48 hours

Track 4: 16DAF cotyledons, untreated

Track 5: 14DAF cotyledons, air dried for 72 hours

Track 6: 17DAF cotyledons, untreated

Track 7: SDS-7 size marker;

b) Western blot analysis of total protein from immature cotyledons from untreated and air dried pods, reacted with antiserum for chaperonin 60 at a dilution of 1/10000

Track 1: SDS-7 size marker

Track 2: 14DAF cotyledons, air dried for 24 hours

Track 3: 15DAF cotyledons, untreated

Track 4: 14DAF cotyledons, air dried for 48 hours

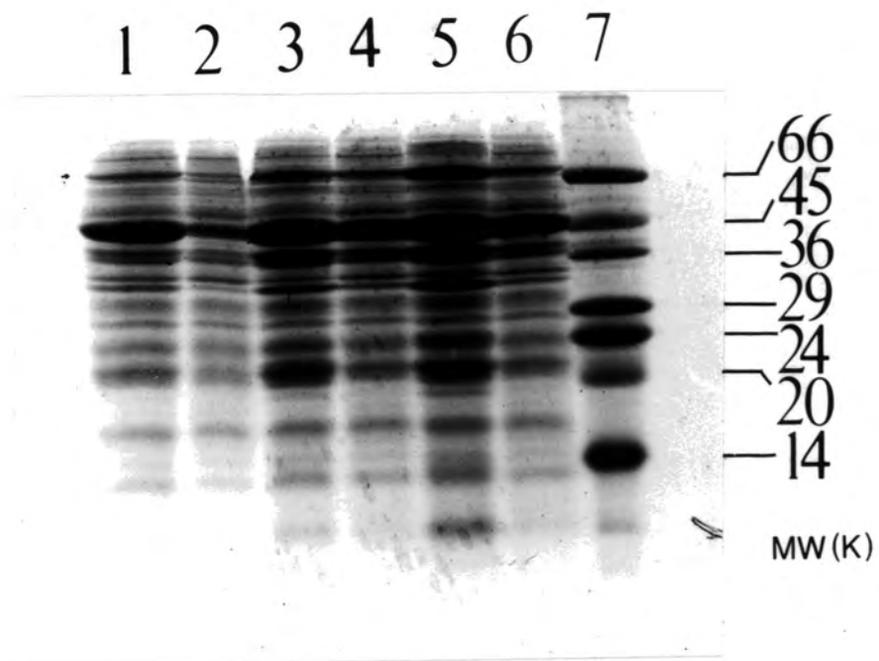
Track 5: 16DAF cotyledons, untreated

Track 6: 14DAF cotyledons, air dried for 72 hours

Track 7: 17DAF cotyledons, untreated.

(Marker and protein band sizes shown are approximate)

a)



b)



subunits respectively) and 26K (major albumin). All of these bands are seen to be present in cotyledons from 12DAF, to increase in abundance during natural maturation and desiccation (Fig.29a) and are stored in the dry seed (Fig. 28c). No prominent protein bands appear in the desiccated cotyledons which were absent in the untreated immature cotyledons at 15-17DAF.

Fig.31 shows the proteins being actively synthesised in immature and prematurely desiccated cotyledons pulse labelled with a ^{14}C -amino acid mix, loaded on a fresh weight basis. The major bands seem to represent the storage proteins vicilin (50K, 33K, 19K), legumin (60K) and convicilin (70K) with a fainter band above convicilin possibly representing the minor Leg S subunit pairs. In all cases the radiolabelled protein bands appeared stronger in the desiccated samples than in the corresponding untreated cotyledons, the greatest difference being observed between 15DAF and 24 hour dried samples. The levels of vicilin synthesis show the greatest increase upon 24 hours of desiccation, then appear to fall slightly upon further drying. In contrast, the levels of legumin and convicilin appear to increase slightly with continued dehydration. No abundant bands appear to be synthesised upon up to 72 hours premature desiccation that are not already present in the untreated cotyledons.

Figure 31. Fluorograph of total protein extracted from pea cotyledons from immature and prematurely desiccated pods, pulse labelled for four hours with a ^{14}C -amino acid mixture. The protein was extracted at 20mg/ml and electrophoresed under non-reducing conditions (20ul/track).

Track 1: 15DAF cotyledons, untreated

Track 2: 14DAF cotyledons, air dried for 24 hours

Track 3: 16DAF cotyledons, untreated

Track 4: 14DAF cotyledons, air dried for 48 hours

Track 5: 17DAF cotyledons, untreated

Track 6: 14DAF cotyledons, air dried for 72 hours

(Major protein bands are marked with approximate sizes)

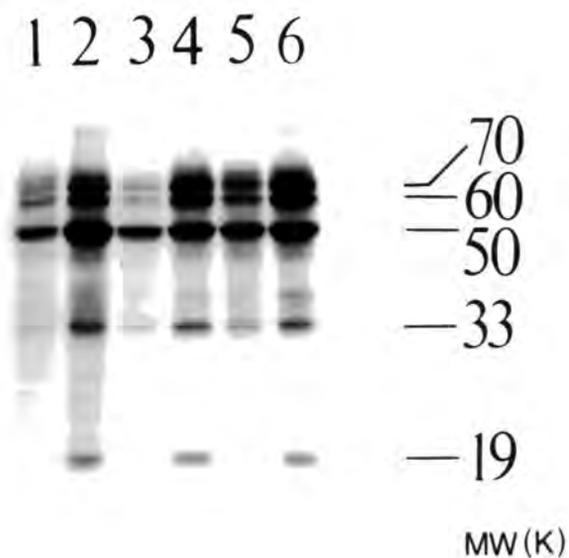


Table 3. Radioactivity (cpm) of total protein extracts from immature and prematurely desiccated cotyledons pulse labelled with a ^{14}C -Amino Acid Mixture. 1ul samples from 20mg/ml extracts.

Sample	cpm/ul
15DAF, untreated	175.2
air dried, 24 hours	176.6
16DAF, untreated	236.8
air dried, 48 hours	270.8
17DAF, untreated	205.0
air dried, 72 hours	233.4

5.2 DISCUSSION

5.2.1 The Analysis of Northern and Dot Blot Results

Several points must be kept in mind when assessing the patterns of message levels from Northern and dot blots. Firstly, it must be remembered that the actual mass of mRNA decreases during late embryogenesis as the majority of messages become less prevalent (see 1.2.5.1); thus, if the level of a certain transcript appears to rise with seed age it does not necessarily imply a large increase in transcription rate or specific stabilisation of the message. It could merely be that as fewer other abundant messages are present the transcript forms a larger proportion of the total

message fraction than formerly, achieving a greater prominence. Similarly, the rise in abundance of sets of messages such as those for seed storage proteins or long-lived mRNA for germination during mid and late maturation may cause an apparent decrease in the proportion of other messages, such as ubiquitin. Secondly, some autoradiographs showed slight differences between the changes in transcript level suggested by Northern blot and those derived from dot blot analysis. In such cases it would seem more advisable to take the results from the Northern blot, as it is possible to visualise the RNA before blotting and confirm that the same weight of RNA has been loaded in each well. A further point to consider is the variation in message level and timing of expression between different batches of the same cultivar. Thompson *et al* (1989) demonstrated that the absolute levels of *Leg A* mRNA at the peak level at 18DAF differed by 30-40% between two batches of pea cv. Feltham First; further, the results of Thompson (1989) on levels of the *Leg J* message showed that although peak levels occurred at 16 and 22DAF in one batch, in another they appeared at 18 and 24DAF. Thus although cultivars seemed to retain the same pattern and rough timing of expression the actual peaks of expression and concentration of transcripts may vary. This may explain the failure of the hybridisation to 12DAF total RNA to follow the general trend of expression (in *Leg A* and *J*), as the 12DAF RNA was extracted from a different batch of peas to the 14-30DAF RNA. To gain a more accurate representation, several

batches of peas should be used for each examination and treatment, and this was unfortunately outside the scope of this project. Consequently, when comparing the effects of premature desiccation and exogenous ABA on message levels (both treatments having been applied to two further batches of Feltham First) it is the pattern of increase or decrease in message level that should be examined rather than the absolute transcript levels. This caveat also applies to comparing levels of hybridisation to different probes: due to varying exposure times and differing specific activities of the probes used the absolute quantities of each species cannot be estimated or compared with any accuracy; however, it was noted during experimentation that longer exposure times and larger quantities of RNA were required to detect messages with PCU1 and PM46 probes, while the legumin probes gave the strongest hybridisation to total RNA in the shortest time.

A further feature of the results which appears misleading is the apparent hybridisation of several probes to the supposedly negative control, *E.coli* rRNA. This was most noticeable with PCU1 and *PsMT_A*, but a small level of non-specific hybridisation also occurred with legumin probes, a result also observed by Thompson (1989). Although this non-specific binding does not necessarily invalidate the pattern of expression below the line of the negative control, it does indicate a low message content in those ages which show less hybridisation of the probe to the message than to the

negative control, an observation often supported by the corresponding Northern blots.

5.2.2 The Effect of Cotyledon Age on Specific Message Levels in Developing Pea Seeds

The transcript level of the *PsMT_A* putative metallothionein follows a pattern thought to be typical of non-seed specific messages, falling from mid-cotyledon expansion to a low level towards maturation then continuing to decline as the seed desiccates and the seed mRNA mass is reduced. The noticeable drop from 14-16DAF, before the more gradual decrease as general message levels fall, may be due to specific down-regulation, either transcriptional or post-transcriptional; however, it is also possible that the fall is only proportional as the levels of vicilin and then legumin increase (Boulter *et al*, 1990). Although a *PsMT_A* protein has not been isolated from pea seeds, the expression of its message during early and mid-embryogenesis implies that a certain amount of the protein may be synthesised and accumulated during seed filling and that the metallothionein is stored in this form rather than as long-lived mRNA, as the transcript accumulation pattern does not resemble that of stored messages. In relation to its proposed function as a detoxicant, there is little if any need for such a system in pea seeds, as during development the metallothioneins in roots and stems will sequester any heavy metals before they reach the seed, while after interruption of the vascular

connections there is no influx of metals until imbibition. Upon germination, the levels of $PsMT_A$ protein may be sufficient until the seedling is competent to produce its own metallothioneins; the results of Evans *et al* (1990) confirmed that the $PsMT_A$ message was present in 14DAF cotyledons, absent in the dry seed and present in the roots of seedlings 14 days after germination. It has also been suggested that metallothioneins may be important in the seed to bind and store metal ions; if valid, this could explain why the $PsMT_A$ transcript level falls as the seed matures and severs connection with the maternal plant, during which the supply of metal ions decreases then stops. This role was proposed for the Ec class II metallothionein in wheat, whose message is abundant in immature embryos, conserved at a lower level in the dry seed and undetectable after germination; however, only 5% of the wheat grain's Zn content was found to be associated with Ec (Kawashima *et al*, 1992). This does not necessarily rule out a role of metal ion storage for other metallothioneins, as the Ec message, unlike most metallothioneins, is not inducible by metal ions and may be more important in another aspect of metal relations. A third possible and as yet unsubstantiated reason for the apparent storage of metallothionein protein in the seed rather than the messages may be as a limited store of cysteine, due to the abundance of this residue in metallothioneins.

Although not seed-specific, the ubiquitin-coding message levels show a different pattern to those of $PsMT_A$.

First, like *PsMT_A*, the amount of RNA hybridising to the polyubiquitin probe falls from 12DAF to maturation, between 16 and 18DAF; the level of hybridisation then appears to increase to an isolated peak at 20DAF (at the onset of desiccation), falls again and then rises again after 24DAF to give strong expression during late desiccation. This pattern could be the result of several factors. Firstly, as well as being required in normal cellular metabolism during development, free ubiquitin and ubiquitin extension proteins would presumably be in great demand during germination for protein degradation, ribosome biogenesis and the other proposed functions of ubiquitin (see section 1.4.3); consequently, either the messages coding for ubiquitin or the ubiquitin proteins would be expected to be present at fairly constitutive levels during development and stored in the dry seed. The pattern of ubiquitin-coding transcript accumulation could thus be a steady level of message expression for translation and for storage, with apparent falls in message level to 18DAF and after 20DAF being caused by a proportionate rise in other seed transcripts, first of seed storage proteins then LEA genes, and the rise from 28-30DAF being a consequence of continued ubiquitin message accumulation while other gene transcript levels fall to low levels. However, the peaks of expression at 20 and 28-30DAF imply that it is possible that superimposed on this low level of expression is a response to stress: at 20DAF after the damage caused by the severance of vascular connections and

the onset of water stress, and later at 28-30DAF during extreme dehydration stress. This is supported by the fact that, instead of the ubiquitin-coding messages merely being stored for use upon germination, the ubiquitin protein has been shown to be synthesised during late embryogenesis, becoming prevalent at 24DAF and increasing in abundance to 30DAF (Haidar and Gatehouse, 1992, submitted). However, a third possibility for the changes in ubiquitin message level is the appearance during desiccation of two new transcripts which hybridise to the polyubiquitin probe (Fig.19b), one at ca.1400b appearing at 20DAF and a second at ca.1300b at 24DAF, both of which, along with the original ca.1800b message, increase in abundance to 30DAF. Ubiquitin is known to be encoded by a multigene family and the presence of several polyubiquitin transcripts of different lengths and exhibiting different expression between plant tissues and upon stress has been demonstrated in barley, *Arabidopsis* and rice (see 1.4.3). Watts and Moore (1989) have also shown the presence of two distinct polyubiquitin messages in pea leaf (PCU1 and PCU2) and the results from cotyledonary total RNA, probed with one of these pea leaf cDNAs, support results seen in Chapter 4, indicating the presence of a number of transcripts in the seed, two of which appear to be specific either to late embryogenesis or to increasing desiccation stress. It is conceivable that the 1.8kb band of hybridisation (seen in Fig.19b) represents a constitutively expressed ubiquitin gene, which is either responsive to

damage and desiccation stress or runs in parallel with a second, stress-responsive 1.8kb transcript, while further stress-responsive and/or late embryogenesis-specific transcripts are represented by the bands at ca.1.4 and 1.3kb. It is assumed that the three bands of hybridisation observed on the Northern blot of developing cotyledons represent polyubiquitin transcripts; although ubiquitin extension protein cDNAs have been isolated from 24-28DAF seed tissue, all other ubiquitin extension proteins examined have been shown to produce short transcripts of between approximately 600-800b, while polyubiquitin messages are commonly longer, from 1-2.5kb (Neves *et al*, 1991; Baker and Board, 1991; Garbarino, Rockhold and Belknap, 1992). No smaller transcripts which may correspond to the pea ubiquitin extension protein messages were noticeable on the Northern blots of total cotyledonary RNA probed with a pea polyubiquitin cDNA. As the majority of research has shown that ubiquitin extension protein messages are most prevalent in actively dividing cells or tissue (Ozkaynak *et al*, 1987; reviewed by Garbarino, Rockhold and Belknap, 1992) it is understandable that in the maturing and dehydrating seed the ubiquitin extension protein message level may be fairly low, producing a small amount of transcripts for ribosome synthesis and possibly for storage for use upon germination. To investigate this further it would be necessary to probe total RNA with the coding sequence just for the tail extension proteins of PCU3 (52 residue) and PCU4 (79 residue)

to ascertain the approximate lengths of these transcripts and to examine their expression during mid-late embryogenesis.

In contrast, although the *PsMT_A* gene is also part of a small multigene family, work by L. Gatehouse and others (unpublished) indicated that only one type of transcript is expressed in each tissue - a 600b transcript in the seed, 670b in stems and 640b in roots. As the transcripts are so similar in length the Northern blots of total cotyledon RNA neither confirm nor deny this, the *PsMT_A* probe hybridising to an indistinct band which may contain one or several transcripts. However, if different transcripts are expressed in pea seeds their expression would appear to be coordinate, as no change in band size during development was observed.

The *Leg A* message appears to exhibit a typical pattern of seed storage protein accumulation. The transcript level increases from 14DAF (it is assumed that the fall from 12 to 14DAF is caused by the difference in batches), the steepest rise occurring between 16 and 20DAF, and reaches a peak at 20DAF; the level then remains high before falling after 24DAF. This pattern agrees with those obtained for *Leg A* expression in cv. Feltham First by Thompson (1989), Thompson *et al* (1989) and Boulter *et al* (1990), although in most of the cases peaks occurred earlier at 18DAF and fell after 22DAF. A different pattern was obtained by Domoney and Casey (1987) with cv. Birte; although they described a similar rise between 15 and 19DAF, to give a steady message level until 22DAF, they also observed a second considerably

higher peak at 24DAF by dot blot analysis, although their Northern analysis did not suggest such a large increase in message level. This difference in transcript accumulation may be attributable to the use of different cultivars and/or environmental growth conditions.

Unexpectedly, the results shown here suggest a second isolated peak at 28DAF, after a substantial drop at 26DAF, which is also noted with *Leg J* transcript levels. This apparent peak disagrees with all recorded results of storage gene message levels, which show a steady and usually rapid fall after 24DAF (reviewed by Goldberg, Barker and Perez-Grau, 1989); it seems reasonable to conclude that this aberrant expression is an artifact caused either by an abnormally low level of legumin transcript expression in the 26DAF total RNA sample or an unexpectedly high level in the 28DAF sample, results from other groups supporting the latter option. Neither of these ages show uncharacteristic levels of the other probes examined, the effect appearing to be specific to legumin messages. Aside from this problem, the *Leg A* message is shown to fall to very low levels (just above hybridisation to the negative control) by late desiccation at 30DAF, confirming the results of Thompson *et al* (1989) and Domoney and Casey (1987).

The pattern of *Leg J* subfamily message accumulation also agrees very well with the results already obtained by Thompson (1989), Thompson *et al* (1989) and Boulter *et al* (1990), increasing first to 16DAF and then reaching a second

peak of expression at 24DAF, four days after the onset of desiccation. This demonstrates, as noted by Thompson *et al* (1991), that at least part of the *Leg J* subfamily expression increases during seed desiccation while, according to the majority of results, other storage protein message levels are falling. However, again, the pattern of expression of the *Leg J* subfamily minor legumin described by Domoney and Casey (1987) in cv. Birte is different, the first peak occurring later at 19DAF and being greater than that seen at 24DAF. This disparity in peak height between the two cultivars can be explained if the theory of the differential accumulation of members of the *Leg J* subfamily suggested by Thompson *et al* (1989) holds true; the first peak at 16-19DAF (depending on cv.) being composed mainly of the as yet uncharacterised *Leg L* and, in cv. Birte, the *Leg K* message, and the second peak representing the increased accumulation of the *Leg J* transcript. In cv. Feltham First the *Leg J* peak is higher than that of *Leg L*, suggesting a degree of differential regulation (*Leg K*, being translationally ineffective in this cv. does not accumulate), but in cv. Birte the combined mass of the *Leg L* and *K* messages is greater than that of *Leg J*, making the second peak of expression at 24DAF (consisting mainly of *Leg J* transcripts) smaller in comparison. Thus these results support the proposal that the apparent biphasic accumulation of *Leg J* subfamily messages is due to the differential expression of genes within the subfamily, with respect to both timing and the level of accumulation, with

Leg K and *L* acting as 'normal' seed protein genes and *Leg J* being specifically up-regulated during dehydration before the message levels fall at late desiccation to the minimum seen with other storage protein messages. This differential expression is supported by the changing levels of hybridisation of the *Leg J* probe to different bands in the total RNA samples (Fig.19c). The similar results of Thompson *et al* (1991) demonstrated that the lower band (1680b), showing a maximum at 24DAF, corresponds to the *Leg J* transcript and proposed that the upper band, with a maximum at 16DAF, represented the *Leg L* message. The third message at ca.1350b was not observed by other workers: it is possible that this faint message, whose level increases from 20 to 24DAF then disappears, is a general or specific degradation product of the *Leg J* message, as no other evidence has been obtained for the presence of another minor legumin gene with homology to the *Leg J* sequence in this cultivar.

Both the *Leg J* and the *Leg A* probes show a decreased level of hybridisation to the 18DAF sample, despite the overall increase in message level between 16 and 20DAF in both types of legumin subunit messages. This comparatively low level of hybridisation to the 18DAF sample was confirmed by other Northern blots, including those showing ABA treatments (Fig.26). It is possible that the fall is caused by a proportionate peak in the expression of other abundant seed messages: Domoney and Casey (1987) demonstrated a peak of hybridisation to a *Leg S* probe at 19DAF in cv. Birte, the

Leg S message being significantly more abundant than the level of transcripts hybridising to a *Leg J* family probe; further, Thompson *et al* (1989) noted a peak of lectin expression in pea cv. Feltham First at 18DAF (although, as with legumin messages, the timing of the peak varies slightly between different batches). It is also possible that the slightly lower level of hybridisation of the *Leg A* probe to 18DAF cotyledonary RNA is a result of differential expression of subfamily members, with the bulk of expression before and after 18DAF being caused by different members of the same subfamily, both running in parallel at ca.1800b. The phenomenon of differential expression of genes within a subfamily during pea seed development, already demonstrated by the ubiquitin multigene family and the *Leg J* subfamily, is also suggested by a Northern blot of the *Leg A* probe hybridised to total cotyledon RNA, where several bands of hybridisation below the major *Leg A* band were seen at 18, 22, 24 and possibly 28DAF. However, the RNA on this blot was fairly old and showed a certain amount of degradation and, further, there were no lower bands visible in the undegraded 20DAF sample; these points, together with the fact that the lengths of the lower transcripts did not correspond to the known lengths of *Leg A* subfamily members, suggested that the lower bands were merely products of degradation. Despite this, the pattern of *Leg A* subfamily message accumulation shown by Domoney and Casey (1987) in cv. Birte does suggest an almost biphasic distribution of message accumulation

similar to that seen in *Leg J*, with a first peak at 16DAF and a second at 24DAF (20-24DAF in cv. Feltham First); it is conceivable that, if the observed fall in the level of hybridisation of the *Leg A* probe to 18DAF RNA is not caused by the increased expression of other abundant messages, one of the members of the *Leg A* subfamily may exhibit a later peak of expression in a pattern similar to that of the *Leg J* transcript. Further screening of seed RNA with specific probes for the different *Leg A* subfamily members is necessary to investigate this possibility of *Leg A* subfamily differential expression. It does appear from a comparison of the Northern blots (Fig.19c and d) that while the bulk of hybridisation to the *Leg A* probe does appear to occur at 20, 22 and 24DAF the difference in the level of hybridisation between these later ages and the earlier (12-18DAF) stages is of a smaller magnitude than that seen between the 12-18DAF and 20-30DAF stages on the blot probed with PJC5.2.

5.2.3 The Effect of Plant Tissue - Specific Message Levels in the Developing Embryonic Axis and in Non-seed Tissues

In soybean, Meinke, Chen and Beachy (1981) found that the expression of seed storage protein genes was not only lower in the embryonic axis than in the cotyledons, but also followed a different accumulation pattern. The same effect was seen with the *Leg A* and *J* messages in total RNA from pea embryonic axes, isolated from the same batch of peas

as the cotyledonary total RNA: both messages are present at very low levels in the axes, exhibiting a small possible peak at 20-22DAF during the beginning of seed desiccation, then decreasing to minimal levels either just above (*Leg J*) or just below (*Leg A*) the apparent level of hybridisation to the negative control. It is assumed that the unexpectedly high level of hybridisation of the *Leg A* probe to 16DAF embryonic axis RNA on the dot blot in Fig.20c was due to unintentional overloading or the use of an improperly thawed sample, as this level of hybridisation was not seen on the Northern blot or on a dot blot screened with a probe for *Leg A*.

Thus storage protein genes show spatial regulation in the seed with the message failing to accumulate to high levels in the non-storage tissue. Further spatial regulation was seen when the total RNA from pea stems, leaves and roots was probed with *Leg A* and *J* subclones. In all non-seed tissue hybridisation to the *Leg A* probe is less than that to the negative control and the seed tissues, implying either no transcription of the message or a very low level of transcription and no accumulation. However, although hybridisation of the *Leg J* probe to stem and leaf RNA appears to be just below that to the negative control, the level of hybridisation to root RNA seems to be slightly higher than that to stem, leaf and *E.coli* rRNA. The results of Boulter *et al* (1990) also implied complete seed specificity of *Leg A*, but indicated a low level of transcription of *Leg J*, *S* and pea lectin in leaf nuclei; similar minimal levels of seed

storage proteins were also seen in leaf nuclei of soybean (Walling *et al*, 1986; see 1.2.5.4). It is therefore not clear whether the seed-specificity of legumin is absolute at the level of transcription, but the messages do not accumulate to detectable levels in most non-seed tissues and no translation products have been identified outside of the seed in either pea or transformed tobacco (Croy *et al*, 1988). These results lend some support to the maintenance of transcript diversity through development and growth suggested by Goldberg *et al* (1981), favouring up- and down-regulation of gene expression for temporal and spatial regulation for most genes, rather than initiation and inactivation of transcription.

Unlike the legumin messages, the levels of the *PsMT_A* and ubiquitin transcripts in the embryonic axis seem to follow the same pattern as that seen in the cotyledons. The level of the *PsMT_A* transcript in the axis seems to be slightly lower than in the cotyledons, but the transcript concentrations in the two tissues fall at the same time. In contrast, the embryonic axes appear to have higher levels of the ubiquitin transcript than the cotyledons, until 30DAF. This may be caused by a lower requirement for ubiquitin in protein degradation in the assimilative storage tissue during most of seed development, with an increased level of ubiquitin being necessary in the cotyledons of the dry seed for the degradation of cotyledon proteins upon imbibition. However, the apparent difference between tissues may also be

due to the lower level of seed storage protein message in the embryo, which means that even if ubiquitin has the same level of message accumulation in the axes and the cotyledons, the transcripts in the axis form a greater proportion of the mass. Although the level of the ubiquitin-coding messages in the axis was too low to detect hybridisation of the ubiquitin probe to 10ug samples of axis total RNA, and so separate transcripts could not be isolated, the similarity of the pattern of ubiquitin message accumulation in the embryo to that in the cotyledons supports the appearance of new sizes of transcripts at 20 and 24DAF in the axis, as in the cotyledons. It would be interesting to probe a Northern blot of embryonic RNA with sequences for the ubiquitin extension protein tails to try and ascertain whether there is a difference in regulation of ubiquitin extension protein messages with time between the axis, which will require rapid cell division and protein synthesis on germination, and the cotyledons, which contain a large amount of protein synthesis during development but change to a primary role of catabolism of stored proteins upon germination.

The levels of both *PsMT_A* and ubiquitin messages were considerably higher in stem and root than in either the seed or the leaf. Evans et al (1990) also found that *PsMT_A* messages are more abundant in roots than in leaves and cotyledons. The lower levels in leaf are slightly unexpected: with ubiquitin it is possible that leaf message levels (although greater than those seen during most of seed

development) are comparatively low (as suggested by Gausing and Barkardottir, 1986) until the leaf ages or undergoes stress and so had not accumulated in the leaves from which the RNA was extracted; while for detoxification and sequestration of metals *PsMT_A* would be required in greater amounts in the organs of uptake and transport than in the leaves. However, a simpler explanation could be that there is a greater diversity of abundant messages in the leaf (and, to some extent, the seed) than in the stem and root, and so these two messages attain greater prominence in the total RNA mass in roots and stems. A further point to remember is the presence of several different transcripts of ubiquitin-coding genes and *PsMT_A*; the phenomenon of members of the same multigene family being differentially regulated to reach different levels in different tissues has already been observed with ubiquitin transcripts in *Arabidopsis thaliana*, barley, potato and tomato (Burke, Callis and Vierstra, 1988; Gausing and Barkardottir, 1986; reviewed by Garbarino, Rockhold and Belknap, 1992). In most cases it was the ubiquitin extension protein messages which exhibited differential expression, between actively dividing and mature tissue (see section 1.4.3). However, Garbarino and co-workers (1992) demonstrated independent regulation of three potato polyubiquitin genes in response to stress, and it is possible that different members of the ubiquitin and *PsMT_A* gene families show different patterns of expression in different tissues.

5.2.4 The Effect of Premature Desiccation on Specific Message Levels in Immature Cotyledons

In the cases of *PsMT_A* and ubiquitin, premature desiccation applied to immature pea cotyledons within isolated pods appeared to accelerate aging, giving premature maturation comparable to that seen in soybean and castor bean (Rosenberg and Rinne, 1986, 1987; Kermode and Bewley, 1988). With *PsMT_A* this takes the form of a faster decline in message level; due to the problems with the control it is not clear whether the fall in transcript level is a direct result of the drying or a normal pattern of message degradation after removal of the pod from the plant. However, with a ubiquitin probe the message level in the treated cotyledons first falls, then increases on prolonged dehydration, to give a rise similar to that seen towards the end of natural maturation. This implies that removal from the plant and premature desiccation within the pod does not cause universal degradation of messages. Although it is possible that the rise in ubiquitin levels is caused by the stress of prolonged isolation of the immature seed from the maternal plant, work from other groups (Rosenberg and Rinne, 1986; Bewley and co-workers, see section 1.2.5.3) supports the theory that the rise in ubiquitin message is a result of premature maturation, with the ubiquitin rise in both naturally and prematurely aged cotyledons being due either to developmental programming (which is affected by premature maturation) or to

the state of desiccation stress experienced by the 30DAF and 8 day silica dried cotyledons, or a combination of the two. A polyubiquitin gene from rice has been shown to be responsive not only to heat shock (leaf blades, leaf sheaths and roots) but also to water stress (leaf blades and sheaths), although in all cases heat shock was most effective (Borkird *et al*, 1991). It was noted from these experiments that the rise in expression of these genes caused by water stress continued, and even increased, after the stressed plants were rehydrated in hydroponic cultures, the high level of expression continuing for a further three days. The authors suggest a relationship between heat shock genes and water adaptation. However, the effect could also be a direct result of the ubiquitin's response to protein damage, the water stress causing a certain amount of harm to proteins and intercellular structures and the rehydration giving further damage, so that the continued expression of ubiquitin is required.

It was noted that the severity of drying treatment appeared to have no significant effect, both drying regimes giving similar message levels, and it is the length of incubation that seems to be the most important factor in determining message levels. Further, it is noticeable that in pea seeds there is no apparent build-up of ubiquitin messages before the seed begins to desiccate. Apart from the peak of ubiquitin expression at the onset of desiccation, which may be involved with the breakdown of vascular

connections between the plant, pod and seed, a fairly severe degree of desiccation stress is required before the message level rises, both in the developing seed and in prematurely desiccated cotyledons. Thus it seems unlikely that ubiquitin plays a major role in the protection of the seed tissue and contents against the onset of desiccation, being instead more important in removing damaged proteins to allow new synthesis and metabolism. However, as it has been suggested that ubiquitin conjugates to a range of cytoplasmic and cell surface proteins, and that several ubiquitin proteins can associate with the same target protein, it is possible that the presence of ubiquitin in association with certain proteins may afford a degree of protection during severe desiccation stress. There is no evidence as yet for a protective function for ubiquitin, its rapid and prolonged response to stress being commonly associated with degradation of the damaged proteins etc., but its conjugative abilities suggest that a protective role may be worth investigating.

In contrast, *PsMT_A* showed no induction upon desiccation, either natural or applied, in the cotyledons; this agrees with other studies which, although showing that metallothioneins are in general responsive to a variety of stresses, have not recorded any positive response to dehydration.

The premature desiccation of immature cotyledons produced an increase in the level of hybridisation of both the *Leg A* and *Leg J* probes to total RNA. The samples

subjected to 24 hours air drying exhibit a slightly lower level of hybridisation to both the *Leg A* and *Leg J* probes than that seen to RNA from untreated cotyledons (although the other Northern blots screened did not show such a large difference between the two samples as that indicated by the blot in Fig.25, where the 14DAF sample was seen on gel to be slightly overloaded). This difference could be due to a disruption in the expression of legumin messages and/or degradation of existing messages upon the shock of pod removal and water deprivation; however, as the untreated and treated cotyledons were from different batches it may merely be that the first untreated batch had a naturally higher level of *Leg J* and *A* messages at this developmental age. The dot blot results do not appear to indicate a noticeably lower level of hybridisation of the legumin probes to the 24 hour air dried sample than to the 14 DAF sample; however, the high degree of saturation in these dot blots makes an accurate comparison very difficult. In contrast, the level of hybridisation to the 24 hour water control is unusually high; from the results of 24 hour water controls from another batch of pea plants in Fig.26 it can be seen that the removal of the pod from the plant and subsequent incubation in water does lead to an increase in the level of RNA hybridising to the legumin message probes, with continued incubation leading to a fall in the level of hybridisation to the *Leg J* probe (although not with the *Leg A* probe). However, the level of hybridisation to the 24 hour water control in the desiccation

treatment batch is still much higher than would be expected and is thought to be due to the extensive degradation of this sample, which led to problems in the accurate estimation of yield and so resulted in overloading.

With the *Leg J* probe, 48 and 72 hours air drying gives a significant rise in message level, as compared to the water and untreated controls. Silica drying has a similar effect, but continued drying with silica crystals led to a significant fall in message level, implying not only a lack of induction of the *Leg J* subfamily messages but also degradation of the large pool of messages accumulated after 72 hours silica drying. The effect of premature desiccation on *Leg A* subfamily message levels is less pronounced, with controls containing higher levels of *Leg A* subfamily transcripts so that the observed rise in hybridisation upon air drying loses some of its significance, although there is still a noticeable difference between the treated and untreated samples at 18S. However, in contrast to the *Leg J* results, silica drying does not appear to have a significantly inductive effect on the level of hybridisation to the *Leg A* probe, although 8 days of silica drying did give a large fall in message level similar to, but apparently greater than, that seen with the *Leg J* results.

Unfortunately, the possibility of different levels of legumin messages between different batches does not allow a direct accurate comparison between prematurely dried samples and samples of untreated pea seeds throughout

development. From the dot blot results the apparent steepness of the rise upon 48 and 72 hours drying would seem to indicate specific induction of the *Leg J* and *A* subfamily messages. However, comparing the Northern blots in Figs.19 and 25 it is possible that the rise in legumin message level is caused by premature maturation. With *Leg J* this would account for the peak of expression after 72 hours air and silica drying, and the fall in expression after 8 days silica drying; this fall appears similar to the developmentally programmed drop in transcript accumulation seen during desiccation in the naturally matured seed. However, the desiccation-induced increase in hybridisation to the *Leg A* probe is much smaller than that seen with the *Leg J* probe, despite a naturally greater abundance of the *Leg A* subfamily messages (Thompson *et al*, 1989), and, further, 72 hours of silica drying does not give a noticeable increase in the level of *Leg A* subfamily messages from 48 hours silica drying or 48 hours water control. This suggests that there may be an added induction of *Leg J* subfamily messages upon desiccation. This proposed induction does not seem to operate under severe desiccation stress, given the substantial fall in hybridisation to the PJC5.2 probe after 8 days silica drying. It is possible that a longer period of isolation from the plant and increased desiccation leads to a cessation of the induction response, followed by the degradation of messages that is seen in normal maturation. Alternatively, the normal developmental programme of down-

regulation at a certain age or state of dehydration may override any induction response, leading to a fall in transcript levels.

As with normal maturation there is a difference in the expression of the *Leg J* and the purported *Leg L* transcripts; analysis of a Northern blot taken from a gel which was electrophoresed at a very slow rate indicated that hybridisation to the 1680b band representing *Leg J* increases upon premature dehydration before that to the 1880b band, which also disappears first upon silica drying. As the level of the *Leg J* message increases naturally during seed desiccation it is quite understandable that its transcript level rises upon the dehydration induced by up to 72 hours air and silica drying, although the extent of the rise appears to be greater than expected; while the *Leg L* and *A* subfamily messages decrease in levels during late desiccation, the bulk of the messages having been produced by 24DAF. Thus the apparent specific induction by desiccation of the *Leg J* message may be a part of normal gene expression, with the changing levels of legumin message being controlled by the state of desiccation of the seed, and *Leg J* being more responsive to a greater degree of dehydration than *Leg A* and *L*. This theory is supported by two main points. Firstly, although the rapid and fairly large response of the legumin messages to desiccation could suggest that the polypeptides or the whole protein are involved in protection of the seed against the effects of moderate dehydration, the fact that

both the legumin polypeptides and the final protein exist enclosed within protein bodies makes a role as an osmotic protectant unlikely. Secondly, the response of legumin to another external environmental factor, sulphur deficiency, was shown to be at the level of post-transcription (Beach *et al*, 1985). The fact that the response to dehydration is, like the timing and seed-specificity of legumin genes, at the level of message accumulation (reviewed by Gatehouse *et al*, 1988) indicates that the response to desiccation is not typical of legumin response to environmental stress. This is supported by the fact that continued desiccation by 8 days silica drying leads to a large fall in legumin message levels, possibly as sufficient protein has been synthesised. Thus the proposed ability of the legumin genes to respond to the level of hydration in the seed would not only allow regulation of seed protein synthesis during normal maturation, but also allow increased synthesis of storage proteins upon the immature seed experiencing water stress, thus giving a selective advantage by allowing the seed to accumulate storage proteins under conditions of stress. This is discussed further in sections 5.2.11 and 6.2.

5.2.5 The Effect of Exogenous ABA on Specific Message Levels in Immature Cotyledons

Accumulation of both the *PsMT_A* and the ubiquitin-coding transcripts was not altered by treatment of immature

cotyledons in their pods with $10^{-4}M$, $10^{-5}M$ or $10^{-6}M$ ABA. For ubiquitin this concurs with previous results from Borkird *et al* (1991) who found that the expression of a polyubiquitin message isolated from rice was unaffected by treatment with ABA. Further, the developmental- or dehydration-induced rise in the level of hybridisation to the polyubiquitin probe during late natural or premature desiccation occurs when the sensitivity of the seed tissues to ABA is decreased (see section 1.2.4.2), supporting the theory that ABA does not play an important role in the expression of ubiquitin in the seed.

The messages of one metallothionein gene, wheat *Ec*, show increased accumulation in germinating seeds on the application of exogenous ABA, and the gene contains an identified ABA-responsive element found in other wheat, rice and cotton genes; however, the absence of a metal responsive element in the *Ec* promoter indicates that the expression of this metallothionein is atypical (Kawashima *et al*, 1992). More recently a cDNA isolated from barley aleurone by Klemsdal *et al* (1991) which also contains ABA-responsive elements, the accumulation of its messages in immature embryos being suppressed by added ABA, has been identified as having similarities to known metallothionein genes (Robinson *et al*, 1993). The presence of abundant levels of *Ec* messages throughout seed development and upon germination, suggest that plants may contain different types of metallothioneins with specialised functions; while *PsMT_A* appears to exist

primarily in the root and stem tissues, there may be further pea metallothioneins which are confined to the seed and whose levels change throughout development in a manner similar to that of *Ec*, which has a proposed role in zinc homeostasis (Kawashima *et al*, 1992).

These results imply that although metallothionein and ubiquitin messages and proteins are known to be responsive to a variety of stresses in plants, the response of ubiquitins and some metallothioneins does not appear to be universally mediated by ABA (although there are certain exceptions among the metallothioneins which may indicate specialised expression or functions). This casts doubt on the suggestion by Zeevaart and Creelman (1988) that ABA acts as a general transducer of stress. Further, these and previous results indicate that ABA is not always involved in the response of genes to water stress; pea ubiquitin shows an increased level of message accumulation upon severe dehydration in the seed (due either to developmental regulation or to stress response) and the transcript level of rice polyubiquitin rises upon drying in leaves, but both the rice polyubiquitin message and the pea ubiquitin transcripts showed no response to ABA.

Incubation of immature seeds within pods with ABA of all three concentrations and both incubation times gave significantly increased levels of both legumin messages when compared to the water controls from the same batch of peas and to the untreated controls of 14 and 18DAF. As with the

response to desiccation, *Leg A* message levels do not seem to increase upon ABA treatment as much as those of the *Leg J* subfamily, relative to the controls. There also appears to be a slightly different pattern of response to the length of incubation. 72 hours incubation with the two lower concentrations of ABA seem to give a slight increase in the level of hybridisation of the RNA to the *Leg A* probe. Part of this increase may be accounted for by the rise in message level seen between the 14 and 18DAF untreated controls, as the natural level of *Leg A* message increases at this developmental stage. In contrast, there appears to be a small decrease in the level of the *Leg J* subfamily message between the 24 and 72 hour treatments, despite a slight increase in transcript level in untreated cotyledons at this developmental stage, although this may just correspond to the fall in message level seen between the 24 and 72 hour water controls. It was not possible to distinguish any significant differences between the effects of the three concentrations of ABA; the details of the effects of ABA concentration on message levels in the seed might be better examined with a larger range of ABA concentrations, incubation periods and age of the treated seeds.

The observation of a pea seed storage protein response to ABA contradicts the conclusions of Barratt *et al* (1989). These workers took 3 day old pods and cultured them for 14 days on medium either with no additives, or supplemented either with 10^{-5} M ABA or fluoridone (to inhibit

ABA biosynthesis); upon translation and immunoprecipitation of the total embryo RNA they found no apparent differences in the protein patterns between the treatments, and the group concluded that ABA had no significant effect on pea seed proteins. These results are partly invalidated by a later experiment (Barratt, Domoney and Wang, 1989) where two ABA-responsive proteins (which were not induced by the pod culture experiment) were isolated from pea embryos. The proteins were induced in embryos of 30-80mg fresh weight cultured for 3 days on low osmotica and their induction was significantly enhanced by the inclusion of 10^{-5} M ABA in the culture media. However, the culture of young embryos with ABA still failed to produce a change in the immunoprecipitable translation products which hybridised to vicilin or legumin antibodies. It is possible that the age of the seed is an important factor in its response to ABA - the seed may have to be within a certain age limit where it has the programmed ability to accumulate legumin messages and polypeptides (usually after 9DAF) but has not lost its sensitivity to ABA. Barratt, Domoney and Wang (1989) took embryos that were still in the late stages of cell division and had not begun significant legumin message accumulation; it is conceivable that such young embryos were not able to initiate or maintain abundant legumin synthesis, while the 14DAF seeds used in these studies were already demonstrating an increasing accumulation of legumin transcripts before treatment. This proposed developmental control of a response

to ABA suggests that ABA only enhances the accumulation of storage protein synthesis, rather than initiating gene expression; such a theory is in accordance with the review by Skriver and Mundy (1990) which suggests that the regulation of LEA and *rab* genes by ABA, which can induce abundant message and protein synthesis in immature, germinating and even non-seed tissues, occurs by a different mechanism than that of storage protein genes. If this is the case, the pod culture of Barratt *et al* (1989) might have been expected to show an increase in legumin message level, as the pods were incubated to a developmental age of 17DAF and should have been actively synthesising legumin messages and proteins. One possible theory for this lack of a response to ABA from the storage protein messages may be that the continued (14 days) exposure of 3 day pods to 10^{-5} M ABA could either accustom the seed tissues to high internal ABA concentrations from an early age, or induce rapid ABA metabolism to reduce the internal levels and so negate the effect. Alternatively, as the translation products were only assessed at the end of the experiment, it is possible that the storage protein message levels did increase in response to ABA but further exposure led to a drop in levels and degradation of the accumulated messages, as seen with the *Leg J* subfamily message levels after 72 hours treatment with ABA, so that no lasting effect in the mRNA population was observed.

Despite the results of Barratt and co-workers, levels of seed storage protein messages from other related

plants (e.g. soybean, *Vicia faba*) have been shown to be responsive to ABA, as reviewed by Skriver and Mundy (1990); however, this review concluded that as the increase in message levels is slow and only a few fold greater than the natural level the storage proteins are unlikely to be involved in the seed's protection from or response to desiccation, in contrast to the more rapid and effective induction of LEA message and proteins. The results shown here with legumin dispute this assumption, showing a fairly rapid response to ABA in immature seeds which is maintained, like the response to desiccation stress, for at least three days. This does not necessarily prove any involvement of ABA in the response to premature dehydration - as ABA has been identified as an inhibitor of precocious germination it may merely have a general enhancing effect on the major developmental seed proteins. The experiments of Wang *et al* (1987) and Ross and McWha (1990) indicate that endogenous levels of ABA increase during pea seed filling as seed protein accumulation increases; this effect could be multiplied when the seed is exposed to much higher concentrations of ABA than are found in normal seed development, and may decrease slightly with time as the seed loses sensitivity to ABA with age. It is extremely difficult to assess the effects of ABA in normal development using an exogenous source of ABA, because of the added and in many cases inseparable effects of isolation from the plant or seed environment, often involving damage, and the delivery of ABA

at concentrations which may be too high and at an incorrect location. It is generally agreed that the action of hormones within the plant is extremely complex, and different concentrations of hormones applied at different stages of development or to different tissues can have varying or even opposite effects (Morris and Bowles, 1987). Studies on how the absolute ABA levels change *in planta* may give a more accurate picture of how the endogenous seed ABA relates to gene expression. In many cases such experiments have brought into doubt the significance of results obtained *in vitro*, showing for example that the maximal expression of ABA-inducible LEAs occurs when the seed ABA levels and sensitivity are both falling, while the use of ABA-deficient and non-responsive mutants in plants such as tomato (*sit*) and *Arabidopsis* (*aba* and *abi*) has shown that in some cases ABA is unnecessary for normal seed development (see Zeevaart and Creelman, 1988). In the experiments conducted here the only difference between the water controls and the treated seeds was the presence of ABA, so any effects of damage or isolation from nutrient sources should be common to both water and ABA treatments. It was interesting to note the difference between the state of degradation of total RNA from water controls from this batch of peas and the batch used for premature desiccation experiments, compared to the RNA from desiccated and ABA-treated cotyledons. Le Deunff and Rachidian (1988) demonstrated that immature pea seeds had the ability to germinate when allowed to imbibe. It is possible

that isolation of the pod and incubation with water encourages the seeds to begin a germinative pattern of mRNA and protein synthesis while within the pod, degrading its unwanted messages. Incubation without water, as in the premature desiccation experiments, does not cause this extensive degradation as the seed cannot begin to germinate without water. If this is the case, addition of ABA appears to halt the general degradation process - total RNA samples from 24 and 72 hour ABA treatments showed no degradation - as well as increasing the accumulation of the developmentally specific legumin messages. This supports the consensus that ABA prevents precocious germination and encourages development in the immature seed. This factor may also contribute to the low levels of legumin messages in the water controls - if the seed is attempting to pursue a germinative pathway the cessation of transcription and specific degradation of legumin messages would be expected. However, although the fall in *Leg J* transcript concentration between 24 and 72 hours incubation in water supports this proposal, there is no observed fall in *Leg A* subfamily message level upon continued incubation in water, so no conclusions can be drawn.

5.2.6 Em Message Levels in Pea

The inability to detect hybridisation of the Em probe to total RNA blots of treated or untreated cotyledons or pea embryonic axes indicates that the pea seed RNA does

not contain a nucleotide sequence with homology to the Em message, despite the presence of similar polypeptides showing extensive amino acid homology in both monocots and dicots (see section 1.4.2). This does not preclude the presence of a pea embryogenic protein with functional and structural similarities to the Em polypeptide. However, in pea cotyledons such a protein cannot be as abundant as in wheat grains, as storage proteins form up to 80% of the total protein. The case may be different in the embryonic axis, where legumin message levels have been shown to be substantially lower. Manickam and Carlier (1980) isolated a small soluble albumin from mung bean axes with a MW of 12K and a high glutamic acid/glutamine content - although tentatively identified as a storage protein it was later found to have some similarities to the wheat Em protein (Grzelczak *et al*, 1982); this protein was found to be very abundant in dry axes, but present at a substantially lower level in the dry cotyledons, the axis protein level being 3.5-fold greater than the cotyledon level and forming 5% of the dry axis protein. It would seem more sensible to concentrate further searches for hydrophilic proteins which may act in desiccation protection on the axes.

5.2.7 Cowpea Trypsin Inhibitor Levels in Pea

Pea seeds do not appear to contain a trypsin inhibitor protein with sufficient nucleotide homology to the

CpTI sequence to allow hybridisation. Hilder *et al* (1989) found that the heterogeneity between different classes of protease inhibitor within a plant species was greater than the dissimilarity between protease inhibitors of the same class from different species. This indicates that pea seeds, unlike a range of other legumes (e.g. soybean, azuki bean, mung bean, lima bean and *Macrotyloma axillare*; Hilder *et al*, 1989) may not contain a protease inhibitor of the Bowman-Birk trypsin/trypsin type seen in cowpea. This does not preclude the presence of other trypsin and protease inhibitors, but the absence of a similar protein in pea seeds meant that the effect of seed age and tissue, desiccation and ABA on the regulation of such genes could not be examined.

5.2.8 Hybridisation of the PM46 cDNA Insert to Pea Total RNA

From the circumstances of the difficulty of detecting a clear signal to the PM46 probe, the fact that the signal hybridises both to the 18S area and, less strongly, to the 25S area, and the fact that although the cDNA was isolated from a differential screen no pattern of differential expression in developing cotyledons was observed suggest that the results seen are due to non-specific binding of the probe to ribosomal RNA, especially at 18S, under the conditions of high RNA concentration and long exposure times used. It is possible that concomittant with this non-specific binding is a small degree of specific binding to a message in the 18S area, or that there is specific binding to

a very rare message in another part of the blot which is so faint as to be hidden by the high level of background hybridisation. It would be necessary to screen poly(A)⁺RNA rather than total RNA to ascertain whether the PM46 cDNA is a true seed message or a cloning artifact, despite the presence of cloning sites at either end of the cloned sequence. It is clear that even if the PM46 message is present in developing peas it is of extremely low abundance and cannot be classed as an LEA message.

5.2.9 The Effect of Seed Age and Tissue on the Total and Albumin Protein Fractions in Developing Pea Seeds

The apparent increase in protein content of the samples from 12-32DAF is due in part to extensive protein synthesis for seed filling, but during later development it is mainly caused by the dehydration of the seed - as samples were loaded on a fresh weight basis, the loss of water gives more concentrated protein samples with age. In the cotyledons there was no obvious appearance or disappearance of any abundant protein bands with age, both in the albumin and the total protein fractions (Figs.28a and 29a). Although the unpublished results of I.M. Evans and R. Swinhoe (see section 1.5) showed new synthesis of certain albumin bands during desiccation (at 30K, ca.21K, 17.5K and ca.9K, labelled with ¹⁴C-amino acid mix) these bands were not seen to be abundant or newly synthesised in the total cotyledon albumin

fraction shown here (although the gel did not show bands below 10K). The only indication of any abundant albumin bands specific to late embryogenesis in these results were in the embryonic axis. Here, Evans and Swinhoe noted new synthesis during desiccation of albumin bands at 14.3K and 9K in cotyledons pulse-labelled with ³⁵S-methionine. As the gel did not show bands below 10K the presence of the 9K band could not be assessed, but the 14.3K band was not shown to be abundant in the total albumin fractions from 14-32DAF. However, it is possible that the band at ca.18K which is shown to increase in the embryonic axis albumin fraction from 24DAF corresponds to the 17.5K band seen to be newly synthesised in cotyledons; possibly the synthesis of this protein begins or increases upon desiccation in both tissues but is more abundant in the embryonic axis, either by proportion or by tissue-specific expression. This has been noted in other species, for example the low molecular weight (12K) albumin which accumulated during late embryogenesis in mung bean and was abundant in the embryonic axis but present at much lower levels in the cotyledons (Manickam and Carlier, 1980). The fact that the 18K or 17.5K band was not detected by the experiment of Evans and Swinhoe could merely indicate that it does not contain methionine. The presence of a band of the same size in total protein from desiccated embryonic axes (Fig.28c) indicates that the protein is stored in the dry seed and appears to be rare or absent in the cotyledons. A high molecular weight band at 70-80K also appears to be

present during late development in the axis, stored in the desiccated axis and absent from the cotyledon; the synthesis of this protein was not detected in cotyledons by Evans and Swinhoe and so may be axis-specific.

Analysis of the total protein fraction from developing and desiccating pea seeds shows a similar absence of newly-synthesised abundant proteins in the cotyledons (Fig.29a). Thus although the results of Evans and Swinhoe indicate that the synthesis of new proteins does occur during desiccation, when the synthesis of the majority of non-seed specific, early and mid-embryogenesis specific proteins is decreasing or absent, such proteins do not appear to form an abundant portion of the total albumin or protein mass in the cotyledons. This could be partly a result of the bulk of protein mass being formed by legumin and vicilin protein; consequently any LEA proteins could form only a small proportion of the cotyledonary protein. The embryonic axis has a much lower content of globulin storage proteins and does appear to contain at least one LEA protein at ca.18K, which may correspond to a newly synthesised albumin at 17.5K, and possibly a high MW albumin at 70-80K, both of which proteins are sufficiently abundant to be seen in a total protein extract from dry seeds, although they are not prominent. These results indicate that pea contains few if any abundant proteins whose levels are high during late embryogenesis in the cotyledons; LEA proteins may be present in the embryonic axes, possibly at 17.5-18K and 70-80K,

although they do not appear to form a large percentage of the total albumin or total protein fraction. It is recommended that further investigations should be conducted on the axes rather than the cotyledons to identify and isolate protein bands which show an increase in abundance during mid and late embryogenesis.

5.2.10 The Effect of Cotyledon Age and Premature Desiccation on the Levels of Chaperonin 60 Protein

It is clear from the Western blot (Fig.29) that the chaperonin 60 protein is present throughout seed development from 16 to 32DAF. This is in agreement with the observation that molecular chaperones are abundantly and constitutively synthesised in the plant. The level of protein in the samples appears to fall slightly after 22 DAF; this could be caused simply by accumulation of other proteins such as the seed storage proteins and lectins, so that chaperonin 60 forms a lower proportion of the seed weight as the synthesis of the protein follows the general pattern of decrease during desiccation. Levels still remain easily detectable in the desiccating seed, an observation that complies with a molecular chaperone's function of protein protection during cellular activity, protein assembly and transport; all of these functions occur during seed development and would be necessary immediately upon imbibition. However, a further feature of the protection of active sites during transport

and assembly may be a degree of protection from damage during desiccation, the formation of damaged proteins being prevented by their association within the large molecular chaperone macromolecule. It has not been documented whether chaperonin 60 has any ability to maintain hydration of itself or associated protein molecules, or if its protein or message levels are increased upon desiccation stress in the whole plant. Molecular chaperones have been shown to be responsive to several environmental stresses, most notably heat shock, but Gething and Sambrook (1992) proposed that the induction factor in such cases was the presence of an increased number of unfolded or incorrectly assembled proteins rather than the stress per se.

The levels of chaperonin 60 protein do not appear to increase upon premature desiccation of immature cotyledons. Compared to the fresh untreated samples the desiccated cotyledons seem to have lower amounts of the chaperonin 60 protein, despite the protein concentration being higher in dried cotyledons due to water loss; this agrees with the observed slight decrease in protein level with natural seed dehydration and seems to support the theory that premature desiccation applied to immature cotyledons within the pod at this age leads to premature maturation, taking the message and protein populations of the plant to a later developmental stage while the seeds are relatively young. As the samples were loaded on a fresh weight basis, the apparent decrease is probably due in part to the

chaperonin 60 forming a smaller proportion of the total protein mass as the levels of other proteins increase and partly a result of the premature maturation of the seed, leading to a decrease in message levels and protein synthesis of the majority of non-LEA seed specific proteins.

5.2.11 The Effect of Premature Desiccation on Cotyledonary Total Protein

Again, dehydration of the seed leads to a higher concentration of total protein per unit of fresh weight and this may explain part of the apparent increase in protein content of prematurely desiccated cotyledons compared to untreated controls. However, the protein levels after a further 48 and 72 hours of desiccation appear to be slightly less than those seen after 24 hours of desiccation, even though protein levels are usually increasing at this developmental stage, as seen with the 15-17DAF samples. It is possible that the stress of isolation of the pod from the plant and premature desiccation causes first an increase in protein synthetic mechanisms, then a slight decline in protein synthesis upon further water stress, as has been noted in non-seed tissues upon dehydration (see section 1.1). This is supported by the pattern of proteins being actively synthesised in immature and prematurely desiccated cotyledons pulse labelled with ^{14}C -containing amino acids (Fig.31). The major proteins being synthesised at this developmental stage are all storage proteins, and these also form the most

abundant part of the total stored protein (Fig.30a); upon desiccation, the levels of all of the storage proteins appear to increase, with vicilin showing the fastest and largest increase after 24 hours drying, then declining slightly, and legumin showing a slower but continuous increase, in agreement with the increase in message level of *Leg A* and *J* observed at this stage. The fact that vicilin and, to a lesser extent, convicilin and the major storage albumin, appear to share the response to desiccation seen with legumin polypeptide messages indicates that the effect of premature desiccation is general to storage proteins. The faster response of vicilin to desiccation gives further support to the theory that the levels of storage proteins and their messages are related to the state of hydration of the seed - in normal development, the abundant synthesis of vicilin starts before that of legumin and the message levels also decline before the fall in legumin messages (reviewed by Boulter *et al*, 1990). Thus if the premature dehydration of the seed acts to take the seed's message and protein complement to a later state of development, accumulation of vicilin would be expected to increase and end more rapidly than legumin, as indicated by these results. The level of active synthesis of legumin continues to rise during the artificial desiccation period; this, together with the rise in legumin message up to 72 hours drying, indicates that the premature desiccation conditions used do not take the seed to a state equivalent to late desiccation, when storage protein

synthesis is falling. A longer or more severe period of desiccation could lead to a premature fall in protein synthesis, as suggested by the decline in legumin polypeptide message levels after 8 days silica drying.

No abundant protein bands were shown in these results to accumulate or be newly synthesised upon premature desiccation of immature cotyledons. This may simply be a result of the state of desiccation of the seed achieved by the air drying treatment; an unquantitative densitometric analysis of the effect on premature desiccation on the levels of a non-seed specific message (*PsMT_A*, Fig.24a) suggest that 72 hours air drying takes the immature cotyledons to a state equivalent to 22DAF. The only proteins whose pattern of expression agrees with late embryogenesis abundance, the 17.5-18K and 70-80K albumins in Figs. 28b and 28c, do not appear to become abundant until after 24DAF and so may not be induced by this level of dehydration; further, these proteins were not detected in the cotyledons and may be embryo-specific. With the high observed concentration of storage proteins during natural and induced dehydration in the cotyledons it really seems necessary to use specific antibody probes for known LEA proteins to examine their expression during seed development and upon desiccation. The discovery of pea dehydrin (Group 2 LEA) by Robertson and Chandler (1992) and the description by Barratt and Clark (1991) of two ABA-responsive pea LEA albumins show that pea does contain some typical LEA proteins and the use of antibodies to these

proteins and other types of LEA proteins, together with nucleic acid probes where available, should provide a more successful method of detection and characterisation of pea LEA proteins and genes and allow investigation of their expression in the pea seed and non-seed tissues in both unstressed and stressed conditions.

CHAPTER 6 - CONCLUSIONS

6.1 CHANGES IN MESSAGE AND PROTEIN LEVELS DURING PEA SEED DEVELOPMENT

Differential screening of a cDNA library from desiccating 24,26,28DAF pea cotyledons with cDNA from 18DAF cotyledons before full desiccation demonstrated that abundant message populations vary between mature cotyledons and cotyledons undergoing desiccation, as has been indicated by other work (reviewed by Goldberg, Barker and Perez-Grau, 1989). This was confirmed by an examination of the changing message levels of selected abundant and stress-related messages in the developing and desiccating seeds. However, attempts to isolate LEA messages which may be involved with protection of the seed tissues against dehydration were unsuccessful. The cDNA libraries obtained from desiccating cotyledons were of poor quality, perhaps due in part to the degradation of mRNA, producing damaged ends and partial cDNAs, occurring during desiccation, and inappropriate and prolonged storage and growth of the libraries appeared to lead to loss of cDNA inserts. Although differential screening did indicate a change in abundant message populations, the screening of the whole library with total 18DAF cDNA and the preparation of a subtracted probe indicated that the majority of messages present at 24,26,28DAF were already expressed at 18DAF, albeit at different levels. As other results have implied that

expression of LEA messages begins during mid-maturation it may have been more productive to screen the library with cDNA from an earlier plant stage, subtracted to remove abundant storage proteins.

The screening of the library with probes for putative stress proteins confirms the results indicated by the probing of total seed RNA. The *PsMT_A* transcript is extremely rare or absent in desiccating cotyledons and embryonic axes; the decline in message levels in the seed with age is thought to be typical of non-seed specific genes and does not follow the pattern of messages to be stored for germination. Its expression during early to mid-seed development may indicate that some protein is synthesised and stored in the seed for germination; alternatively the requirements for metallothioneins in the seed may be fulfilled by another as yet unidentified metallothionein with a pattern of expression similar to that of the Ec metallothionein from wheat. Both the wheat Em and the *CpTI* messages appeared to be absent from pea tissue.

Isolation of cDNA clones from a 24,26,28DAF cDNA library which hybridised to cDNA clone for pea polyubiquitin did demonstrate that more than one type of ubiquitin transcript is expressed in the seed. cDNAs coding for both previously identified types of ubiquitin extension proteins - one with a 52 amino acid tail and one with a 79 amino acid tail - were described and shown to have a high conservation of amino acid sequence with the extension protein tails from

other plants, animals, yeast and micro-organisms, including the site identified as being involved in nuclear localisation and a putative zinc-binding, nucleic acid binding domain. These proteins and their messages had not been previously isolated from pea, although they are thought to be universal, with a common role in ribosome biogenesis. Characterisation of a second partial polyubiquitin cDNA provided by F.Z. Watts demonstrated that pea contains at least two different polyubiquitin genes, giving a ubiquitin multigene family of at least four members. It is not known if these transcripts are expressed differentially between tissues, as in some other plants, but it does appear from Northern blots of total RNA from developing seeds that the levels of transcripts of different sizes change during seed development, some (thought by their length to represent polyubiquitin messages) appearing to increase in abundance upon desiccation of the seed. This increase may represent a response to severe water stress, which has been tentatively indicated in rice; or an accumulation of message and/or protein for use upon germination; or a simply a response to the presence of damaged proteins upon dehydration; or a combination of these factors. By the sizes of transcripts hybridising to the ubiquitin probe in developing and desiccating pea seeds it does not appear that the messages for the pea ubiquitin extension proteins are abundant during mid-late embryogenesis, although their presence has been demonstrated by the method of their isolation. This observation agrees

with other work which indicates that ubiquitin extension proteins and their messages are abundant only in actively dividing tissue; the low observed level of message in desiccating pea cotyledons is presumably present either for a basal level of ribosome synthesis or for storage as long lived mRNA for use upon germination. The use of the tail nucleic acid sequences as specific probes may give further insight into the expression of ubiquitin extension proteins in pea.

An examination of the accumulation of messages for major and minor polypeptides of the legumin protein confirms that the developmental expression of storage protein genes is tightly controlled and varies not only between storage proteins but also between messages for components of the same protein and even members of the same gene subfamily. There is no clear indication as to why the maximum accumulation of the *Leg J* message should be after that of other legumin polypeptides; Evans and Swinhoe (unpublished) found that the levels of the *Leg J* polypeptide increased three-fold between 20 and 24DAF in cotyledons pulse-labelled with ¹⁴C-amino acids. It would be interesting to investigate whether the heterogeneous nature of the legumin protein remains the same throughout seed development, or whether proteins synthesised later during desiccation contain a higher level of the *Leg J* polypeptide; and, if so, if this alters the properties of the protein in any way. The *Leg J* gene lacks a core enhancer element GCCACCTC which is present in the 5' flanking

sequences of pea *Leg A*, vicilin and homologous genes, as well as similar storage protein genes from other species (Thompson *et al*, 1991) and is thought to be involved in storage protein gene control.

Analysis of the partial cDNA clone isolated from a differential screen of the 24,26,28DAF library gave a sequence - M46 - which shows no significant homology to sequences in the Daresbury data bank or the SwissProt data bank. The sequence remains unidentified and attempts to demonstrate the presence and expression of a message corresponding to the partial cDNA sequence were inconclusive and did not definitely prove the presence of such a message. Further work is required to ascertain whether this is a true cDNA or, despite the presence of cloning adaptors, an artifact; a screen of poly(A)⁺RNA may demonstrate whether the faint binding of the cloned cDNA to total RNA was to a message or merely unspecific binding to ribosomal RNA.

Despite the changes in message populations there was no obvious appearance or disappearance of any abundant protein bands in total and albumin protein extracts from cotyledons. This could be partly due to the large proportion of stored proteins in aging seeds so that proteins expressed towards the end of embryogenesis would form only a small portion of the total protein mass. In the embryonic axis, where levels of storage proteins are substantially lower, at least one band was observed during late development that appeared to be at low levels or absent in the undesiccated

axis and in cotyledons, and this band was also found in the dry axis (ca.18K). This observation, together with results from Manickam and Carrier (1980) which indicate that some LEA messages and proteins may be expressed at higher levels in the axis than in the storage tissue suggest that it would have been more productive to construct a cDNA library from embryonic axes to isolate and examine the expression of genes with a possible function in desiccation protection.

6.2 THE EFFECT OF PREMATURE DESICCATION ON MESSAGE AND PROTEIN LEVELS IN DEVELOPING PEA COTYLEDONS

The application of premature desiccation treatments to immature cotyledons affected the expression of all of the pea messages examined. In the cases of the messages for *PsMT_A* and ubiquitin the drying treatment appeared to act to instigate premature maturation, an effect already recorded in other species (see sections 1.2.3.4 and 1.2.5.3). *PsMT_A* exhibited a slightly accelerated decline in transcript level, while the level of hybridisation to ubiquitin shows an increase upon prolonged desiccation, caused either by 'speeding up' of the developmental programme or by the severe desiccation stress seen both in 30DAF and 8 day silica dried seeds. The size of the band hybridising to the ubiquitin probe in 8 day silica dried seeds indicates that this increase upon desiccation is due primarily to polyubiquitin. A similar premature maturation response was seen with the

levels of chaperonin 60. In normal development the protein level seems fairly constant, with a slight decline on desiccation due either to higher levels of other proteins, or to the fall in message and protein synthesis seen at this developmental stage, or to a combination of the two; upon premature desiccation of immature cotyledons the same slight fall in protein level is seen, presumably for one or both of the same reasons.

A slightly different scale of response to desiccation was seen with the levels of *Leg A* and *J* messages. A short term drying treatment greatly increased the levels of hybridisation to *Leg J* and, to a lesser extent, *Leg A* probes to total cotyledonary RNA. Although it could be inferred from this that the legumin protein or polypeptides could have a function in the protection of seed tissues against the effects of dehydration, the fact that legumin exists in the seed inside protein bodies, together with the observation that a longer period of silica drying causes a substantial fall in legumin message levels, makes a role for legumin in desiccation protection unlikely. Instead it is suggested that the timing and/or level of storage protein expression, and possibly that of other genes expressed in the seed, is somehow controlled, via accumulation of transcripts, by the state of hydration of the seed tissues. This would explain the larger response of the *Leg J* message to desiccation, as its usual maximal expression occurs at a time of greater dehydration than that of *Leg A*.

It is possible that premature desiccation could lead not only to an increase in message level due to premature maturation, but also to an added induction to allow the seed tissues to synthesise all of the legumin necessary for successful seedling growth in a shorter time period; this would give the plant the ability to accelerate maturation under conditions of environmental stress and produce a seed with an almost full storage protein complement well before normal maturation. This is supported by the results from the pulse labelling of immature and prematurely desiccated cotyledons. These demonstrate that desiccation increases the levels of all pea storage proteins, and further that the order of storage protein synthesis appears to be maintained, with the peak of vicilin synthesis occurring first then decreasing as the levels of legumin increase more slowly. Several groups have suggested that the loss of water during seed development could be an important factor in the production of a mature viable seed with the ability to germinate upon imbibition (reviewed by Bewley, 1979; Saab and Obendorf, 1989). Complete desiccation may not be essential for all seeds - in many species the state of dehydration reached during maturation {e.g. soybean (Rosenberg and Rinne, 1986), cowpea (Ishibashi and Minamikawa, 1989), pea (Le Deunff and Rachidian, 1988)} or that engendered by a minimal drying treatment (e.g. castor bean and *Phaseolus vulgaris* (Bewley, Kermode and Misra, 1989)) may be sufficient to give a seed with all of the properties necessary to give efficient

germination and seedling growth. As examination of the *Leg A* promoter in transgenic tobacco has already identified several areas containing possible control elements for seed specificity and enhancement of transcription; such experiments could possibly be extended to investigate whether the desiccation response of *Leg A* is retained in transgenic tobacco and, if so, which part of the promoter controls this response. Such experiments could allow the isolation of a control element common to seed storage genes, and possibly other seed genes showing controlled changes in message accumulation during seed development. They could also further understanding of how the expression of genes in the seed is related to the stage of seed growth and the state of hydration and, conversely, if or how the loss of water during seed development is involved in the production of a viable seed. To investigate these theories further, premature desiccation experiments should be repeated with several batches of pea plants, using untreated pods from each batch as controls so that a direct comparison between the message and protein levels in treated and untreated tissues can be made. The desiccation treatments should also be applied at different developmental ages, especially to see if pea storage proteins or their messages can be induced by desiccation after maximal storage protein synthesis has occurred. The accumulation of separate storage proteins in the dry seed after normal development and after premature desiccation to the same state of dehydration could also be

compared to see if premature desiccation leads to an increase in yield.

The results from the analysis of total and actively synthesised proteins from immature and prematurely desiccated cotyledons did not identify any abundant protein showing an increase upon dehydration of immature cotyledons within seeds other than the major storage proteins, which were already undergoing abundant synthesis at this developmental stage. It is possible that the desiccation treatments were not severe enough or of a long enough time period to induce pea non-storage protein genes usually expressed during late embryogenesis, such as the newly synthesised albumins described by Evans and Swinhoe and the pea dehydrins isolated by Robertson and Chandler (1992), even though the latter could be induced by 24-48 hours desiccation in seedlings. It is possible that the peas may have been too immature to synthesise such proteins; alternatively, the high concentration of storage proteins overwhelmed the lower levels of LEA proteins, which might be visible in samples from which the storage proteins have been extracted, or in pea embryonic axes which have lower storage protein levels. The use of specific probes for known pea and plant LEAs may be a more effective method for the investigation of pea LEAs and their expression.

The dehydration treatments also gave no indication of the presence of proteins which may be involved in desiccation tolerance. Le Deunff and Rachidian (1988) found

that desiccation tolerance in pea seeds was acquired gradually throughout seed filling (from 10-20DAF). Several groups have isolated proteins which they believe to be involved in (but not necessarily cause) desiccation tolerance {e.g. barley (Bartels, Singh and Salamini, 1988); soybean (Blackman *et al*, 1991)}. The protein results shown here indicate that either such proteins have been synthesised and stored before 15DAF and are not induced further by premature desiccation, or that the levels of protein synthesis at this developmental stage and/or upon dehydration are too low to show up in the pulse labelled protein samples.

6.3 THE EFFECT OF EXOGENOUS ABA ON MESSAGE LEVELS IN IMMATURE PEA COTYLEDONS

Results from experiments with polyubiquitin and PsMT_A indicate that ABA is not a universal transducer of stress, as suggested by Zeevaart and Creelman (1988). However, treatment of immature cotyledons within their pods with exogenous ABA did demonstrate that ABA significantly increases the levels of legumin storage protein message accumulation. The results of Barratt, Domoney and Wang (1989) indicate that this effect only acts when the messages are being abundantly synthesised, suggesting that any regulation by ABA is in the form of either enhancement of transcription or specific stabilisation of messages rather than initiation or significant up-regulation of transcription, as with LEAs. The observed response to both

ABA and osmotic stress implies that this hormone may be involved in the regulation of storage proteins during natural and premature maturation; however, the response could also be merely a result of the encouragement by ABA of the accumulation of developmental-specific messages, with the use of concentrations much higher than those seen in normal seed development causing an overproduction of storage protein messages.

The results presented in this thesis have demonstrated that the message levels of selected genes vary during seed development, genes with different functions following very different patterns of accumulation. The levels of all messages and proteins examined that were present in the seed were altered by premature desiccation, which affected the timing of gene expression, producing in the seed the levels of messages and proteins which are normally found at a later stage of maturation. Seed storage protein genes were shown to be induced to high levels by premature desiccation, giving rapid storage protein synthesis upon stress. Storage protein genes in immature pea seeds were also shown to be responsive to some extent to ABA, the hormone appearing to enhance rather than initiate message accumulation.

REFERENCES

- Ackerson, R.C. (1984). Abscisic acid and precocious germination in soybeans. *Journal of Experimental Botany*, 35, 414-421.
- Baker, J., C. Steele, and L. Dure III (1988). Sequence and characterisation of 6 *Lea* proteins and their genes from cotton. *Plant Molecular Biology*, 11, 277-291.
- Baker, R.T. and P.G. Board (1991). The human ubiquitin-52 amino acid fusion protein gene shares several structural features with mammalian ribosomal genes. *Nucleic Acids Research*, 19, 1035-1040.
- Barratt, D.H.P., C. Domoney and T.L. Wang (1989). Purification and partial characterisation of two ABA-responsive proteins induced in cultured embryos of *Pisum sativum* L. *Planta*, 180, 16-23.
- Barratt, D.H.P., P.N. Whitford, S.K. Cook, G. Butcher and T.L. Wang (1989). An analysis of seed development in *Pisum sativum* VIII. Does ABA prevent precocious germination and control storage protein synthesis? *Journal of Experimental Botany*, 40, 1009-1014.
- Barratt, D.H.P. and J.A. Clark (1991). Proteins arising during the late stages of embryogenesis in *Pisum sativum* L. *Planta*, 184, 14-23.

- Bartels, D., M. Singh and F. Salamini (1988).** Onset of desiccation tolerance during development of the barley embryo. *Planta*, 175, 485-492.
- Bartels, D., K. Schneider, G. Terstappen, D. Piatkowski and F. Salamini (1990).** Molecular cloning of abscisic acid-modulated genes which are induced during desiccation of the resurrection plant *Craterostigma plantagineum*. *Planta*, 181, 27-34.
- Baumlein, H., W. Boerjan, I. Nagy, R. Panitz, D. Inze and U. Wobus (1991).** Upstream sequences regulating legumin gene expression in heterologous transgenic plants. *Molecular and General Genetics*, 225, 121-128.
- Beach, L.R., D. Spencer, P.J. Randall and T.J.V. Higgins (1985).** Transcriptional and post-transcriptional regulation of storage protein gene expression in sulphur deficient pea seeds. *Nucleic Acids Research*, 13, 999-1013.
- Benech Arnold, R.L., M. Fenner and P.J. Edwards (1991).** Changes in germinability, ABA content and ABA embryonic sensitivity in developing seeds of *Sorghum bicolor* (L.) Moench. induced by water stress during grain filling. *New Phytologist*, 118, 339-347.
- Bewley, J.D. (1979).** Physiological aspects of desiccation tolerance. *Annual Review of Plant Physiology*, 30, 195-238.

- Bewley, J.D. and M. Black (1985). Chapter 3 in: *Seeds - Physiology of Development and Germination*. Plenum Press, New York, pp.40-164.
- Bewley, J.D., A.R. Kermode and S. Misra (1989). Desiccation and minimal drying treatments of seeds of castor bean and *Phaseolus vulgaris* which terminate development and promote germination cause changes in protein and mRNA synthesis. *Annals of Botany*, 63, 3-17.
- Birnboim, H.C. and J. Doly (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research*, 7, 1513-1523.
- Blackman, S.A., S.H. Wettlaufer, R.L. Obendorf and A.C. Leopold (1991). Maturation proteins associated with desiccation tolerance in soybean. *Plant Physiology*, 96, 868-874.
- Bogue, M.A., R.A. Vonder Haar, M.L. Nuccio, L.R. Griffing and T.L. Thomas (1990). Developmentally regulated expression of a sunflower 11S seed protein gene in transgenic tobacco. *Molecular and General Genetics*, 222, 49-57.
- Borkird, C., B. Claes, A. Caplan, C. Simoens and M. van Montagu (1991). Differential expression of water-stress associated genes in tissues of rice plants. *Journal of Plant Physiology*, 138, 591-595.

- Boulter, D., R.R.D. Croy, I.M. Evans, J.A. Gatehouse, N. Harris, A. Shirsat and A. Thompson (1990). The molecular biology of pea seed development with particular reference to the storage protein genes. Chapter 11 in: *Genetic Engineering of Crop Plants*. Eds. G.W. Lycett and D. Grierson, Butterworth, London, pp. 127-141.
- Bown, D., M. Levasseur, R.R.D. Croy, D. Boulter and J.A. Gatehouse (1985). Sequence of a pseudogene in the legumin gene family of pea (*Pisum sativum* L.). *Nucleic Acids Research*, 13, 4527-4538.
- Bray, E.A. (1991). Regulation of gene expression by endogenous ABA during drought stress. *Journal of Experimental Botany Abstracts*, 42, No.238, Supplement, 9.
- Burke, T.J., J. Callis and R.D. Vierstra (1988). Characterisation of a polyubiquitin gene from *Arabidopsis thaliana*. *Molecular and General Genetics*, 213, 435-443.
- Butler, W.M. and A.C. Cuming (1991). Molecular responses to ABA and osmotic stress in embryos of viviparous maize mutants. *Journal of Experimental Botany Abstracts*, 42, 11.

Cabrera y Poch, H.L., C.Arribas and N. Izquierdo (1990).

Sequence of a *Drosophila* cDNA encoding a ubiquitin gene fusion to a 52-amino acid ribosomal protein tail.

Nucleic Acids Research, 18, 3994.

Callis, J., J.A. Raasch and R.D. Vierstra (1990). Ubiquitin extension proteins of *Arabidopsis thaliana*. *Journal of Biological Chemistry*, 265, 12486-12493.

Casey, R. (1979). Genetic variability in the structure of the alpha-subunits of legumin from *Pisum* - a two-dimensional gel electrophoretic study. *Heredity*, 43, 265-272.

Casey, R., C. Domoney and T.H.N. Ellis (1986). Legume storage proteins and their genes. *Oxford Surveys in Plant Molecular Cell Biology*, 3, 1-95.

Chapman, A.B., M.A. Costello, F. Lee and G.M. Ringold (1983). Amplification and hormone-regulated expression of a mouse mammary tumour virus-*Eco* *gyl* fusion plasmid in mouse 3T6 cells. *Molecular Cell Biology*, 3, 1421.

Close, T.J., A.A. Kortt and P.M. Chandler (1989). A cDNA-based comparison of dehydration-induced proteins (dehydrins) in barley and corn. *Plant Molecular Biology*, 13, 95-108.

Cohen, A. and E.A. Bray (1992). Nucleotide Sequence of an ABA-induced tomato gene that is expressed in wilted

vegetative organs and developing seeds. *Plant Molecular Biology*, 18, 411-413.

Comai, L. and J.J. Harada (1990). Transcriptional activities in dry seed nuclei indicate the timing of the transition from embryogeny to germination. *Proceedings of the National Academy of Science, USA*, 87, 2671-2674.

Cook, S.K., H. Adams, C.L. Hedley, M.J. Ambrose and T.L. Wang (1988). An analysis of seed development in *Pisum sativum*. VII. Embryo development and precocious germination *in vitro*. *Plant Cell, Tissue and Organ Culture*, 14, 89-101.

Croy, R.R.D. and J.A. Gatehouse (1985). Genetic Engineering of Seed Proteins: Current and Potential Applications. In: *Plant Genetic Engineering*, ed. J.H. Dodds, Cambridge University Press, Cambridge, UK, pp.143-268.

Croy, R.R.D., I.M. Evans, J.N. Yarwood, N. Harris, J.A. Gatehouse, A.H. Shirsat, A. Kang, J.R. Ellis, A. Thompson and D. Boulter (1988). Expression of pea legumin sequences in pea, *Nicotiana* and yeast. *Biochemie und Physiologie der Pflanzen*, 183, 183-197.

Cuming, A.C. (1984). Developmental regulation of gene expression in wheat embryos: molecular cloning of a DNA sequence encoding the Em polypeptide. *European Journal of Biochemistry*, 145, 351-357.

- Dasgupta, J. and J.D. Bewley (1982). Desiccation of axes of *Phaseolus vulgaris* during development causes a switch from a developmental pattern of protein synthesis to a germinative pattern. *Plant Physiology*, 70, 1224-1227.
- De Bruijn, S.M. and D. Vreugdenhil (1992). Abscisic acid and assimilate partitioning to developing seeds. I. Does abscisic acid influence the growth rate of pea seeds? *Journal of Plant Physiology*, 140, 201-206.
- Domoney, C., D.R. Davies and R. Casey (1980). The initiation of legumin synthesis in immature embryos of *Pisum sativum* L. grown *in vivo* and *in vitro*. *Planta*, 149, 454-460.
- Domoney, C., T.H.N. Ellis and D.R. Davies (1986). Organisation and mapping of legumin genes in *Pisum*. *Molecular and General Genetics*, 202, 280-285.
- Domoney, C. and R. Casey (1987). Changes in legumin messenger RNAs throughout seed development in *Pisum sativum* L. *Planta*, 170, 562-566.
- Dure III, L.S., M. Crouch, J. Harada, T-H.D. Ho, J. Mundy, R. Quatrano, T. Thomas and Z.R. Sung (1989). Common amino acid sequence domains among the LEA proteins of higher plants. *Plant Molecular Biology*, 12, 475-486.
- Dure III, L. (1993). A repeating 11-mer amino acid motif and plant desiccation. *The Plant Journal*, 3, 363-369.

- Ellis, J. (1987). Proteins as molecular chaperones. *Nature*, 328, 378-379.
- Ellis, J.R., A.H. Shirsat, A. Hopher, J.N. Yarwood, J.A. Gatehouse, R.R.D. Croy and D. Boulter (1988). Tissue-specific expression of a pea legumin gene in seeds of *Nicotiana plumbaginifolia*. *Plant Molecular Biology*, 10, 203-214.
- Ellis, R.J. (1990). Molecular chaperones: the plant connection. *Science*, 250, 954-959.
- Espelund, M., S. Saeboe-Larssen, D.W. Hughes, G.A. Galau, F. Larsen and K.S. Jakobsen (1992). Late embryogenesis-abundant genes encoding proteins with different numbers of hydrophilic repeats are regulated differentially by abscisic acid and osmotic stress. *The Plant Journal*, 2, 241-252.
- Evans, I.M., R.R.D. Croy, P. Hutchinson, D. Boulter, P.I. Payne and M.E. Gordon (1979). Cell-free synthesis of some storage protein subunits by polyribosomes and RNA from developing seeds of pea (*Pisum sativum* L.). *Planta*, 144, 455-462.
- Evans, I.M., L.N. Gatehouse, J.A. Gatehouse, N.J. Robinson and R.R.D. Croy (1990). A gene from pea (*Pisum sativum* L.) with homology to metallothionein genes. *FEBS Letters*, 262, 29-32.

- Evans, M., M. Black and J. Chapman (1975).** Induction of hormone sensitivity by dehydration is the one positive role for drying in cereal seed. *Nature*, 258, 144-145.
- Feinberg, A.P. and B. Vogelstein (1983).** A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry*, 132, 6-13. And addendum (1984), 137, 266-267.
- Finkelstein, R.R. and M.L. Crouch (1986).** Rapeseed embryo development in culture on high osmoticum is similar to that in seeds. *Plant Physiology*, 81, 907-912.
- Finley, D. and A. Varshavsky (1985).** The ubiquitin system: functions and mechanisms. *Trends in Biochemical Science*, 10, 343-347.
- Finley, D., B. Bartel and A. Varshavsky (1989).** The tails of ubiquitin precursors are ribosomal proteins whose fusion to ubiquitin facilitates ribosome biogenesis. *Nature*, 338, 394-401.
- Garbarino, J.E., D.R. Rockhold and W.R. Belknap (1992).** Expression of stress-response ubiquitin genes in potato tubers. *Plant Molecular Biology*, 20, 235-244.
- Gatehouse, J.A., I.M. Evans, R.R.D. Croy and D. Boulter (1986).** Differential expression of genes during legume seed development. *Philosophical Transactions of the Royal Society of London*, B314, 367-384.

Gatehouse, J.A., D. Bown, J. Gilroy, M. Levasseur, J.

Castleton and T.H.N. Ellis (1988). Two genes encoding 'minor' legumin polypeptides in pea (*Pisum sativum* L.): Characterisation and complete sequence of the *Leg J* gene. *Biochemical Journal*, 250, 15-24.

Gausing, K. and R. Barkardottir (1986). Structure and

expression of ubiquitin genes in higher plants.

European Journal of Biochemistry, 158, 57-62.

Genschik, P., Y. Parmentier, M.C. Criqui and J. Fleck (1990).

Sequence of a ubiquitin extension protein tail in

tobacco. *Nucleic Acids Research*, 18, 4007.

Gething, M-J. and J. Sambrook (1992). Molecular Chaperones.

Nature, 355, 33-45.

Goday, A., D. Sanchez-Martinez, J. Gomez, P. Puigdomenech and

M. Pages (1988). Gene expression in developing *Zea*

mays embryos: regulation by ABA of a highly

phosphorylated 23-to 25-kD group of proteins. *Plant*

Physiology, 88, 564-569.

Goldberg, R.B., G. Hoschek, S.H. Tam, G.S. Ditta and R.W.

Breidenbach (1981). Abundance, diversity and

regulation of mRNA sequence sets in soybean

embryogenesis. *Developmental Biology*, 83, 201-217.

Goldberg, R.B., S.J. Barker and L. Perez-Grau (1989).

Regulation of gene expression during plant embryogenesis. *Cell*, 56, 149-160.

- Gomez, J., D. Sanchez-Martinez, V. Stiefel, J. Rigau, P. Puigdomenech and M. Pages (1988). A gene induced by the plant hormone ABA on response to water stress encodes a glycine-rich protein. *Nature*, 334, 262-264.
- Goupil, P., P. Hatzopoulos, G. Franz, F.D. Hempel, R. You and Z.R. Sung (1992). Transcriptional regulation of a seed-specific carrot gene, DC8. *Plant Molecular Biology*, 18, 1049-1063.
- Grzelczak, Z.F., M.H. Sattolo, L.K. Hanley-Bowdoin, T.D. Kennedy and B.G. Lane (1982). Synthesis and turnover of proteins and mRNA in germinating wheat embryos. *Canadian Journal of Biochemistry*, 60, 389-397.
- Gubler, U. and B.J. Hoffman (1983). A simple and very efficient method for generating cDNA libraries. *Gene*, 25, 263-269.
- Haidar, A. and J.A. Gatehouse (1992). Expression of ubiquitin during late embryogenesis in pea (*Pisum sativum* L.). *Plant Molecular Biology* (submitted).
- Hall, T.C., Y. Ma, B.U. Buchbinder, J.W. Payne, S.M. Sun and F.M. Bliss (1978). Messenger RNA for G1 protein of french bean seeds: cell-free translation and product characterisation. *Proceedings of the National Academy*

of Science, USA, 75, 3196-3200.

- Hames, B.D. (1981). An introduction to polyacrylamide gel electrophoresis. In: *Gel Electrophoresis of Proteins: A Practical Approach*. Eds. B.D. Hames and D. Richmond, IRL Press, Oxford, pp.1-91.
- Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology*, 166, 557-580.
- Harada, J.J., C.S. Baden and L. Comai (1988). Spatially regulated genes expressed during seed germination and postgerminative development are activated during embryogeny. *Molecular and General Genetics*, 212, 466-473.
- Harada, J.J., R.A. Dietrich, L. Comai and C.S. Baden (1988). Regulation of gene expression during seed germination and postgerminative development. Chapter 2 in: *Plant Gene Research*, 5, eds. D.P.S. Verna and R.B. Goldberg, pp.85-107.
- Harada, J.J., A.J. DeLisle, C.S. Baden and M.L. Crouch (1989). Unusual sequence of an ABA-inducible mRNA which accumulates late in *Brassica napus* seed development. *Plant Molecular Biology*, 12, 395-401.
- Hershko, A. and A. Ciechanover (1982). Mechanisms of intracellular protein breakdown. *Annual Review of*

Biochemistry, 51, 335-364.

Hershko, A. (1991). The ubiquitin pathway for protein degradation. *Trends in Biochemical Science*, 16, 265-269.

Hilder, V.A., A.M.R. Gatehouse and D. Boulter (1990). Genetic engineering of crops for insect resistance using genes of plant origin. Chapter 5 in: *Genetic Engineering of Crop Plants*, eds. G.W. Lycett and D. Grierson, Butterworth, London, pp.51-66.

Hilder, V.A., R.F. Barker, R.A. Samour, A.M.R. Gatehouse, J.A. Gatehouse and D. Boulter (1989). Protein and cDNA sequences of Bowman-Birk protease inhibitors from the cowpea (*Vigna unguiculata* Walp.). *Plant Molecular Biology*, 13, 701-710.

Hilder, V.A., A.M.R. Gatehouse, S.E. Sheerman, R.F. Barker and D. Boulter (1987). A novel mechanism of insect resistance engineered into tobacco. *Nature*, 330, 160-163.

Hoffman, N.E., K. Kenton, D. Milkowski and E. Pichersky (1991). Isolation and characterisation of tomato cDNA and genomic clones encoding the ubiquitin gene *ubi3*. *Plant Molecular Biology*, 17, 1189-1201.

Hong, B., R. Barg and T-H.D. Ho (1992). Developmental and organ-specific expression of an ABA- and stress-induced

protein in barley. *Plant Molecular Biology*, 18, 663-674.

Huynh, T.V., R.A. Young and R.W. Davis (1985). Constructing and screening cDNA libraries in lambda-gt10 and lambda-gt11. In: *DNA Cloning Techniques: A Practical Approach*, volume 1. Ed. D.M. Glover, IRL Press, Oxford, pp.49-65.

Ishibashi, N. and T. Minamikawa (1989). Cotyledonary mRNA species and germinability of immature *Vigna unguiculata* seeds. *Plant Cell Physiology*, 30, 373-379.

Jacobson, A. (1987). Purification and fractionation of poly(A)⁺RNA. *Methods in Enzymology*, 152, 254-261.

Johnson, E.S., B. Bartel, W. Seufert and A. Varshavsky (1992). Ubiquitin as a degradation signal. *EMBO Journal*, 11, 497-505.

Jung, V., S.B. Pestka and S. Pestka (1990). Efficient cloning of PCR-generated DNA containing terminal restriction endonuclease recognition sites. *Nucleic Acids Research*, 18, 6156.

Kafatos, F.C., C.W. Jones and A. Efstratiadis (1979). Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridisation procedure. *Nucleic Acids Research*, 7, 1541.

- Kagi, J.H.R. and A. Schaffer (1988). Biochemistry of metallothionein. *Biochemistry*, 27, 8509-8515.
- Kawashima, I., T.D. Kennedy, M. Chino and B.G. Lane (1992). Wheat E_c metallothionein genes: like mammalian Zn²⁺ metallothionein genes, wheat Zn²⁺ metallothionein genes are conspicuously expressed during embryogenesis. *European Journal of Biochemistry*, 209, 971-976.
- Kermode, A.R. and J.D. Bewley (1985). The role of maturation drying in the transition from seed development to germination. I. Acquisition of desiccation tolerance and germinability during development of *Ricinus communis* L. seeds. *Journal of Experimental Botany*, 36, 1906-1915.
- Kermode, A.R. and J.D. Bewley (1988). The role of maturation drying in the transition from seed development to germination. V. Responses of the immature castor bean embryo to isolation from the whole seed: a comparison with premature desiccation. *Journal of Experimental Botany*, 39, 487-497.
- Kermode, A.R., S.K. Pramanik and J.D. Bewley (1989). The role of maturation drying in the transition from seed development to germination. VI. Desiccation-induced changes in mRNA populations within the endosperm of *Ricinus communis* L. seed. *Journal of Experimental Botany*, 40, 33-41.

- Kermode, A.R., E.B. Dumbroff and J.D. Bewley (1989). The role of maturation drying in the transition from seed development to germination. VII. Effects of partial and complete desiccation on ABA levels and sensitivity in *Ricinus communis* L. seeds. *Journal of Experimental Botany*, 40, 303-313.
- Klemsdal, S.S., W. Hughes, A. Lonneborg, R.B. Aalen and O-A. Olsen (1991). Primary structure of a novel barley gene differentially expressed in immature aleurone layers. *Molecular and General Genetics*, 228, 9-16.
- Kuhlemeier, C., P.J. Green and N-H. Chua (1987). Regulation of gene expression in higher plants. *Annual Review of Plant Physiology*, 38, 221-257.
- Kyhse-Anderson, J. (1984). Electrophoretic transfer of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. *Journal of Biochemical and Biophysical Methods*, 10, 203-209.
- Lalonde, L. and J.D. Bewley (1986). Desiccation during germination of pea axes and its effects on the mRNA population. *Journal of Experimental Botany*, 37, 754-764.
- Lane, B.G., R. Kajioka and T.D. Kennedy (1987). The wheat-germ Ec protein is a zinc containing metallothionein.

Biochemical Cell Biology, 65, 1001-1005.

Le Deunff, Y. and Z. Rachidian (1988). Interruption of water delivery at physiological maturity is essential for seed development, germination and seedling growth in pea (*Pisum sativum* L.). *Journal of Experimental Botany*, 39, 1221-1230.

Levitt, J. (1980). Water Stress. Chapter 2 in: *Responses of Plants to Environmental Stresses*. Volume 2, 2nd Ed., Academic Press, London, pp.25-280.

Litts, J.C., G.W. Colwell, R.L. Chakerian and R.S. Quatrano (1987). The nucleotide sequence of a cDNA clone encoding the wheat E_m protein. *Nucleic Acids Research*, 15, 3067-3618.

Litts, J.C., M.B. Erdman, N. Huang, E.E. Karrer, A. Noueir, R.S. Quatrano and R.L. Rodriguez (1992). Nucleotide sequence of the rice (*Oryza sativa*) E_m protein gene (Emp1). *Plant Molecular Biology*, 19, 335-337.

Logemann, J., J. Schell and L. Willmitzer (1987). Improved method for the isolation of RNA from plant tissues. *Analytical Biochemistry*, 163, 16-20.

Lycett, G.W., R.R.D. Croy, A.H. Shirsat and D. Boulter (1984). The complete nucleotide sequence of a legumin gene from pea (*Pisum sativum* L.). *Nucleic Acids Research*, 12, 4493-4506.

- Lycett, G.W., R.R.D. Croy, A.H. Shirsat, D.M. Richards and D. Boulter (1985). The 5' flanking regions of three pea legumin genes; comparison of the DNA sequences. *Nucleic Acids Research*, 13, 6733-6743.
- Manickam, A. and A.R. Carlier (1980). Isolation and function of a low molecular weight protein of mung bean embryonic axes. *Planta*, 149, 234-240.
- March, J.F., D.J.C. Pappin and R. Casey (1988). Isolation and characterisation of a minor legumin and its constituent polypeptides from *Pisum sativum*. *Biochemical Journal*, 250, 911-915.
- Marcotte, W.R., C.C. Bayley and R.S. Quatrano (1988). Regulation of a wheat promoter by ABA in rice protoplasts. *Nature*, 335, 454-457.
- Marcotte, W.R., S.H. Russell and R.S. Quatrano (1989). ABA-responsive sequences from the Em gene of wheat. *The Plant Cell*, 1, 969-976.
- Mason, H.S. and J.E. Mullet (1990). Expression of two Soybean vegetative storage protein genes during development and in response to water deficit, wounding and jasmonic acid. *The Plant Cell*, 2, 569-579.
- Matta, N.K., J.A. Gatehouse and D. Boulter (1981). Molecular and subunit heterogeneity of legumin of *Pisum sativum* L. (garden pea) - a multidimensional gel

electrophoretic study. *Journal of Experimental Botany*, 32, 1295-1307.

McBride, L.J., S.M. Koepf, R.A. Gibbs, W. Salser, P.E.

Mayrand, M.W. Hunkapillar and M.N. Kronick (1989).

Automated DNA sequencing methods involving polymerase chain reaction. *Clinical Chemistry*, 35, 2196.

McCubbin, W.D., C.M. Kay and B.G. Lane (1985). Hydrodynamic

and optical properties of the wheat germ Em protein.

Canadian Journal of Biochemistry and Cell Biology, 63, 803-811.

Meakin, P.J and J.A. Gatehouse (1991). Interaction of seed

nuclear proteins with transcriptionally-enhancing

regions of the pea (*Pisum sativum* L.) *leg A* gene

promoter. *Planta*, 183, 471-477.

Meinke, D.W., J. Chen and R.N. Beachy (1981). Expression of

storage protein genes during soybean (*Glycine max* cv.

Provar) seed development. *Planta*, 153, 130-139.

Mierendorf, R.C. and D. Pfeffer (1987). Direct sequencing of

plasmid DNA. *Methods in Enzymology*, 152, 556-562.

Miller, K. (1987). Gel Electrophoresis of RNA. *Focus*

(Bethesda Research Laboratories), 9, No.3, 14-15.

Misra, S. and J.D. Bewley (1985). The mRNA population in the

embryonic axes of *Phaseolus vulgaris* during development

and following germination. *Journal of Experimental Botany*, 36, No.171, 1644-1652.

Morris, C.F., R.J. Anderberg, P.J. Goldmark and M.K. Walker-Simmons (1991). Molecular cloning and expression of abscisic acid-responsive genes in embryos of dormant wheat seeds. *Plant Physiology*, 95, 814-821.

Morris, P.C. and D.J. Bowles (1987). Abscisic acid and embryogenesis. In: *Growth Regulators and Seeds*, eds. N.J. Pinfield and M. Black, *British Plant Growth Regulator Group Monographs*, 15, Long Ashton, UK, 65-76.

Morris, P.C., E.W. Weiler, S.E. Maddock, M.G.K. Jones, J.R. Lenton and D.J. Bowles (1988). Determination of endogenous ABA in immature cereal embryos during *in vitro* culture. *Planta*, 173, 110-116.

Morris, P.C., A. Kumar, D.J. Bowles and A.C. Cuming (1990). Osmotic stress and ABA induce expression of wheat Em genes. *European Journal of Biochemistry*, 190, No.3, 625-630.

Morris, P.C., P.C. Jewer and D.J. Bowles (1991). Changes in water relations and endogenous ABA content of wheat and barley grains and embryos during development. *Plant, Cell and Environment*, 14, 443-446.

Mundy, J. and N-H. Chua (1988). ABA and water stress induce the expression of a novel rice gene. *EMBO Journal*, 7,

2279-2286.

- Neves, A.M., P. Guerreiro, L. Miquerol and C. Rodrigues-Pousada (1991). Molecular cloning and expression of a *Tetrahymena pyriformis* ubiquitin fusion gene coding for a 53-amino acid extension protein. *Molecular and General Genetics*, 230, 186-192.
- Oeda, K., J. Salinas and N-H. Chua (1991). A tobacco bZip transcription activator (TAF-1) binds to a G-box-like motif conserved in plant genes. *EMBO Journal*, 10, No.7, 1793-1802.
- Oishi, M.Y. and J.D. Bewley (1992). Premature drying, fluridone-treatment and embryo isolation during development of maize kernels (*Zea mays* L.) induce germination, but the protein synthetic responses are different. Potential regulation of germination and protein synthesis by abscisic acid. *Journal of Experimental Botany*, 43, 759-767.
- Ozkaynak, E., D. Finley, M.J. Solomon and A. Varshavsky (1987). The yeast ubiquitin genes: a family of natural gene fusions. *EMBO Journal*, 6, 1429-1439.
- Payne, P.I. (1976). The long-lived mRNA of flowering plant seeds. *Biological Reviews*, 51, 329-363.
- Pollmann, L., J. von Kampen and M. Wetterlind (1991). Ubiquitin in a lower plant. Characterisation of ubiquitin-

encoding DNA and RNA from *Chlamydomonas reinhardtii*.
European Journal of Biochemistry, 202, 197-204.

- Raynal, M., P. Gaubier, F. Grellet and M. Delseny (1990).
Nucleotide sequence of a radish cDNA clone coding for a
late embryogenesis abundant (LEA) protein. *Nucleic
Acids Research*, 18, 6132.
- Rerie, W.G., M.I. Whitecross and T.J.V. Higgins (1990).
Nucleotide sequence of an A-type legumin gene from pea.
Nucleic Acids Research, 18, 655
- Richardson, M. (1977). The Protease Inhibitors of Plants and
Micro-organisms. *Phytochemistry*, 16, 159-169.
- Robertson, M. and P.M. Chandler (1992). Pea dehydrins:
Identification, characterisation and expression. *Plant
Molecular Biology*, 19, 1031-1044.
- Robinson, N.J., A.M. Tommey, C. Kuske and P.J. Jackson
(1993). Plant metallothioneins. *Biochemical Journal*
(submitted).
- Rosenberg, L.A. and R.W. Rinne (1986). Moisture loss as a
prerequisite for seedling growth in soybean seeds
(*Glycine max* L. Merr.). *Journal of Experimental
Botany*, 37, No.184, 1663-1674.
- Rosenberg, L.A. and R.W. Rinne (1987). Changes in seed
constituents during germination and seedling growth of
precociously matured soybean seeds (*Glycine max*).

Annals of Botany, 60, 705-712.

Rosenberg, L.A. and R.W. Rinne (1988). Protein synthesis during natural and precocious soybean seed (*Glycine max* [L.] Merr.) maturation. *Plant Physiology*, 87, 474-478.

Rosenberg, L.A. and R.W. Rinne (1989). Protein synthesis during rehydration, germination and seedling growth of naturally and precociously matured soybean seeds (*Glycine max*). *Annals of Botany*, 64, 77-86.

Ross, G.S. and J.A. McWha (1990). The distribution of ABA in *Pisum sativum* plants during seed development. *Journal of Plant Physiology*, 136, 137-142.

Rouse, D., C.A. Gehring and R.W. Parish (1992). Structure and sequence of a dehydrin-like gene in *Arabidopsis thaliana*. *Plant Molecular Biology*, 19, 531-532.

Saab, I.N. and R.L. Obendorf (1989). Soybean seed water relations during *in situ* and *in vitro* growth and maturation. *Plant Physiology*, 89, 610-616.

Sambrook, J., E.F. Fritsch and T. Maniatis (1989). *Molecular Cloning: A Laboratory Manual*. 2nd Ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, USA.

Sanger, F., S. Nicklen and A.R. Coulson (1977). DNA sequencing with chain termination inhibitors. *Proceedings of the National Academy of Science, USA*,

74, 5463-5467.

- Schobert, B. (1977). Is there an osmotic regulatory mechanism in algae and higher plants? *Journal of Theoretical Botany*, 68, 17-26 (In Levitt, 1980).
- Shirsat, A., N. Wilford, R. Croy and D. Boulter (1989). Sequences responsible for the tissue-specific promoter activity of a pea legumin gene in tobacco. *Molecular and General Genetics*, 215, 326-331.
- Shirsat, A.H., P.J. Meakin and J.A. Gatehouse (1990). Sequences 5' to the conserved 28bp Leg box element regulate the expression of pea seed storage protein gene Leg A. *Plant Molecular Biology*, 15, 685-693.
- Skriver, K. and J. Mundy (1990). Gene expression in response to ABA and osmotic stress. *The Plant Cell*, 2, 503-512.
- Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology*, 98, 503.
- Thompson, A.J. (1989). Regulation of Gene Expression in Developing Pea Seeds. Ph.D. thesis, University of Durham.
- Thompson, A.J., I.M. Evans, D. Boulter, R.R.D. Croy and J.A. Gatehouse (1989). Transcriptional and post-transcriptional regulation of seed storage protein gene expression in pea (*Pisum sativum* L.). *Planta*, 179,

279-287.

Thompson, A.J., D. Bown, S. Yaish and J.A. Gatehouse (1991).

Differential expression of seed storage protein genes in the pea leg J subfamily: sequence of gene leg K. *Biochemie und Physiologie der Pflanzen*, 187, 1-12.

Thompson, E.W. and B.G. Lane (1980). Relation of protein synthesis in imbibing wheat embryos to the cell-free translational capacities of bulk mRNA from dry and imbibing seeds. *Journal of Biological Chemistry*, 255, 5965-5970.

Tommeey, A.M., J. Shi, W.P. Lindsay, P.E. Urwin and N.J. Robinson (1991). Expression of the pea gene *PsMT_A* in *E.coli*: metal binding properties of the expressed protein. *FEBS*, 292, No.1,2, 48-52.

Walling, L., G.N. Drews and R.B. Goldberg (1986).

Transcriptional and post-transcriptional regulation of soybean seed protein mRNA levels. *Proceedings of the National Academy of Science, USA*, 83, 258-259.

Wang, T.L., S.K. Cook, R.J. Francis, M.J. Ambrose and C.L.

Hedley (1987). An analysis of seed development in *Pisum sativum*. VI. Abscisic acid accumulation. *Journal of Experimental Botany*, 38, 1921-1932.

Watts, F.Z. and A.L. Moore (1989). Nucleotide Sequence of a full length cDNA clone encoding a polyubiquitin gene

from *Pisum sativum*. *Nucleic Acids Research*, 17, No.23,
10100.

Williamson, J.D., R.S. Quatrano and A.C. Cuming (1985). Em
polypeptide and its mRNA levels are modulated by ABA
during embryogenesis in wheat. *European Journal of
Biochemistry*, 152, 501-507.

Williamson, J.D. and R.S. Quatrano (1988). ABA-regulation of
two classes of embryo-specific sequences in mature
wheat embryos. *Plant Physiology*, 86, 208-215.

Zeevaart, J.A.D. and R.A. Creelman (1988). Physiology of
abscisic acid. *Annual Review of Plant Physiology*, 39,
439-473.

ACKNOWLEDGEMENTS

I would like to thank Dr. J.A. Gatehouse for supervision of this project, analysis of results and the provision of technical facilities. I am also indebted to the following researchers for advice and assistance: Mr. J. Davies, Ms. G. Davison, Dr. D. Bown, Miss H. Edmonds, Mrs. L. Edwards, Dr. T. Fawcett, Dr. T. Fordham-Skelton, Mr. J. Gilroy, Mr. L. Gatehouse, Mr. D. Hutchinson (photography), Dr. K. Johnston, Dr. P. Meakin, Mr. R. Swinhoe, Dr. A. Tommey and especially Dr. I.M. Evans and Dr. N. Robinson for advice and encouragement and Ms. J. Bartley for excellent sequencing services. I am grateful to the following for generous donations of plasmid DNA: Dr. R.R.D. Croy (PRC3.1), Dr. A.C. Cuming (PMuge Kla), Ms. A.S. Dale (PASD1), the John Innes Institute (PJC5.2), Dr. A. Tommey (*PsMT_A*) and Dr. F.Z. Watts (PCU1 and PCU2); and to Dr. L. Barnett and Prof. R.J. Ellis for the kind donation of chaperonin 60 antiserum. I also acknowledge the support of a grant from the Agriculture and Food Research Council.

