Abiotic stress in plants: Late Embryogenesis Abundant proteins

Imen Amara

Abiotic stress in plants: Late Embryogenesis Abundant proteins

IMEN AMARA

Barcelona 2012

UNIVERSITAT DE BARCELONA

FACULTAT DE FARMACIA

DEPARTAMENTO DE BIOQUIMICA Y BIOLOGIA MOLECULAR

Abiotic stress in plants: Late Embryogenesis Abundant proteins

Memoria presentada por Imen Amara para obtener el grado de Doctor por Universidad de Barcelona.

Trabajo realizado en el Departamento de Genética Molecular del Centre de Recerca en Agrogenòmica (CRAG), Laboratorio de Genética Molecular de Plantas

Dra. **Adela Goday**

(Directora de Tesis)

Dra. **Josefa Badia Palcin**

(Tutora de Tesis)

Imen Amara

Caminar con sandalias hasta que la sabiduría te ofrece zapatos

Avicenne (980-1037)

A mes chers parents

Tous les mots du monde ne sauraient exprimer l'immense amour que je vous porte, ni la profonde gratitude que je vous témoigne pour tous les efforts et les sacrifices que vous n'avez jamais cessé de consentir pour mon instruction et mon bien-être. J'espère avoir répondu aux espoirs que vous avez fondés en moi. Je vous rends hommage par ce modeste travail en guise de ma reconnaissance éternelle et de mon infini amour.

A mes sœurs et mon frère

C'est à travers vos encouragements que j'ai opté pour cette noble profession, et c'est à travers vos critiques que je me suis réalisée. Puisse l'amour et la fraternité nous unissent à jamais. Je vous souhaite une belle vie.

A mon mari

Il me soit permis aujourd'hui de t'assurer mon profond amour et ma grande reconnaissance. Je ne peux exprimer à travers ses lignes tous mes sentiments d'amour et de tendresse envers toi ton aide, ta générosité, ton soutien ont été pour moi une source de courage et de confiance à surmonter toutes les difficultés rencontrées au cours de cette thèse.

A mon petit

Merci pour la joie que tu me procures.

A tous ceux ou celles qui me sont chers et que j'ai omis involontairement de citer.

Agradecimientos

Llegado este momento me gustaría agradecer a los que de forma directa o indirecta han contribuido en mi trabajo de tesis doctoral.

Este trabajo no hubiera sido posible sin la variada y generosa contribución de muchas personas e instituciones que han facilitado las cosas para que este trabajo llegue a un feliz término. Por ello, es para mí un verdadero placer utilizar este espacio para ser justo y consecuente con ellas, expresándoles mis agradecimientos.

- En primer lugar quiero agradecer a mi directora de Tesis: **Dra. Adela Goday** por su dirección, por haber confiado en mí, por tener paciencia y por haberme enseñado tantas cosas (no sólo a nivel profesional) y por todo el tiempo que me han dado, por sus sugerencias e ideas de las que tanto provecho he sacado.

-También me complace agradecer de manera especial a la **Prof. Montserrat Pagés**, por haberme recibido, por la oportunidad de poder realizar este trabajo en su laboratorio y por apoyarme hasta el final. Debo agradecer también su amabilidad durante mi estancia en su grupo, durante la cual he tenido todo el soporte profesional y logístico para alcanzar los objetivos perseguidos. Muchas gracias por permitirme vivir una experiencia tan importante para mi formación. He aprendido mucho en tu laboratorio.

- J'exprime toute ma gratitude et mes plus profonds remerciements à Monsieur **Prof. Khaled Masmoudi**, le chef du Laboratoire Protection et amélioration des Plantes au CBS de m'avoir accueilli dans son laboratoire, pour ses aides interminables me m'aider a trouvée une bourse de coopération et venir en Espagne.

- A mi tutora de tesis la **Dra. Josefa Badia Palacin** por aceptar la tutoría de esta tesis.

-A los miembros del tribunal por acceder a evaluar la presente tesis doctoral.

-Debo un especial reconocimiento a la Agencia Española de Cooperación Internacional para el Desarrollo (AECID) y L'Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) por las becas que me han dado.

-A la **Dra. Eliandre de Oliveira** y a **Maria Antonia Odena**, de la plataforma proteomica del Parc Científic de Barcelona, por los análisis de espectrometría de masas, y por las largas discusiones que han sido la clave del buen trabajo que hemos realizado juntos, el cual no se puede concebir sin su siempre oportuna participación. Le agradezco también el haberme facilitado siempre los medios suficientes para llevar a cabo todos los resultados.

- Quiero expresar también mi más sincero agradecimiento a la **Dra. Dolors Ludevid** por las largas discusiones, siempre enmarcadas en su orientación y rigurosidad; su capacidad para guiar mis ideas ha sido un aporte invaluable.

No puedo olvidar a mis compañeros y amigos con los cuales he compartido despacho e incontables horas de trabajo. Gracias por los buenos y malos momentos, por aguantarme y por escucharme.

- Quiero expresar mi agradecimiento especial a **Marta Riera** quien fue una compañera siempre generosa y dispuesta como pocos, que compartió conocimientos y experiencias de tipo profesional y personal que fueron de gran valor.

-A **VictoriaLumbreras**, gracias por tus consejos y tu ayuda.

- Gracias, **Belmiro**, tu amistad y ayuda que no tiene precio. Gracias por aguantarme tanto tiempo y por hacer de los momentos de convivencia dentro del laboratorio una experiencia muy rica tanto en lo profesional como personal. Eres mi amigo para toda la vida.

-Gracias **Isa**, mi profesora de español, gracias por tu constante apoyo y por animarme siempre que haga falta. Gracias por tu ayuda siempre incondicional , por escucharme y brindarme ayuda cuando ha sido necesario y por haber hecho de mi estancia en el CRAG una maravillosa aventura y una continúa caja de sorpresas! Qué rica es tu comida!.

- A los nuevos compañeros del grupo: **Tommaso**, **Elena** y **Agnese**, por llegar con tan buena onda. Gracias sobre todo por trasmitirme el entusiasmo y la energía necesaria para afrontar esta última parte de la tesis. Habéis sido unos compañeros fantásticos.

-A **Sami** por enseñarme, casi todo lo que se de proteínas lo he aprendido de ti.

-A la **Dra. Montserrat Capellades**, por su inestimable ayuda y sus buenos consejos y por estar siempre dispuesta a echar una mano. Y a **Eva** por tu ayuda.

-A los compañeros y compañeras del laboratorio que ya se fueron, por haber hecho tan agradable trabajar a su lado**.** Muchas gracias **Cristina**, **Michael** por vuestra sonrisa y vuestro buen humor durante todos estos años.

- A **Hedia**, por la gran amistad que me has ofrecido, por tu constante apoyo y por animarme siempre que haga falta.

- Con mucho cariño, a **Anahit** como me alegro de haberte conocido, muchas gracias por la gran amistad que me has ofrecido, por tu tierna compañía y tu inagotable apoyo.

- A los invernaderos: **Pilar**, **Eva**, **Jordi, Mina** y Alejandro. Gracias **Montse** por toda tu ayuda y tus consejos en la realización de los experimentos del confocal.

- Para todos los miembros del Departamento de Genética Molecular del CRAG: Maite, Maria, Jorge, Silvia, Nico, Merce, Pep, Norma, Judit, Mary Paz, mis más sinceros agradecimientos.

- Je tiens à avouer sincèrement ma profonde gratitude à mes amis (es): **Malika**, **Nadia**, **Ikram et Rania** pour l'aide et leurs encouragements.

-Je tiens à avouer sincèrement ma profonde gratitude à **Moez Hanin**, **Faïçal Brini,** pour leurs efforts incessants et leurs aides.

Merci à tous et à toutes.

INDEX

I. INTRODUCTION 19

1. ABIOTIC STRESS IN PLANTS: CELLULAR AND MOLECULAR MECHANISMS 19
2. DESICCATION TOLERANCE 21
3. LATE EMBRYOGENESIS ABUNDANT (LEA) PROTEINS 22
4. CLASSIFICATION OF LEA PROTEINS: SEQUENCE MOTIFS 23
4.1 GROUP 1 26
4.2 GROUP 2 26
4.3 GROUP 3 28
4.4 GROUP 4 28
4.5 GROUP 5 29
4.6 GROUP 6 29
4.7 GROUP 7 29
5. SUBCELLULAR LOCALIZATION AND EXPRESSION PROFILES. 30
6. BIOCHEMICAL PROPERTIES AND STRUCTURE OF LEA PROTEINS 31
7. LEA PROTEIN FUNCTIONS 36
7.1 Protein protection 36
7.2 Membrane protection 37
7.3 Ion binding and antioxidant function 38
7.4 Other functions 38
8. BIOTECHNOLOGICAL APPLICATIONS OF LEA GENES 39
9. GROUP 2 RAB17 AND GROUP 5 RAB28 LEA PROTEINS FROM MAIZE: A BRIEF SUMMARY. 41

II. AIMS OF THE PRESENT WORK 43

III. CHAPTER 1 47

1. INTRODUCTION 49
2. MATERIALS AND METHODS 51
2.1 Plant material and protein extraction: heat and acid treatments 51
2.2 Gel electrophoresis 52
2.3 Immunoblot analysis 52
2.4 In-Gel Digestion 52
2.5 Acquisition of MS/MS Spectra by LC-ESI-MSMS. 52
2.6 In-Solution Digestion 53
2.7 Acquisition of MS and MSMS Spectra by LC-MALDI-MSMS. 53
2.8 Data Analysis 53

3. RESULTS **54**
3.1 Comparison of the electrophoretic protein patterns of acid-insoluble and acid-soluble extracts from Arabidopsis seed thermostable proteins 54
3.2 Analysis of SDS gel bands by LC-ESI-MSMS 56
3.3 Analysis of soluble and insoluble acid fractions by LC-MALDI-MSMS. 59
4. DISCUSSION **65**
4.1 General comments on the analytical MS methods. 65
4.2 Physicochemical properties of the proteins. 66
5. CONCLUSION **69**

IV. CHAPTER 2 73

1. INTRODUCTION **75**
2. MATERIALS AND METHODS **77**
2.1 Heat and acid soluble proteins from maize embryos 77
2.2 Expression of recombinant LEA proteins in E.coli 77
2.3 In vitro protein aggregation assay 77
2.4 In vivo assay of E. coli stress tolerance 78
2.5 Maize stress treatments 78
2.6 RNA extraction and semi-quantitative RT-PCR analysis 78
2.7 Transient expression in N. benthamiana leaves and laser-scanning confocal microscopy. Transgenic Arabidopsis plants. 79
2.8 Cell stains 80
2.9 Lipid body isolation and Oleosin Coimmunoprecipitation 80
2.10 Phosphorylation assays 81
2.11 Mass Spectrometry (LC-MS/MS) 81
2.12 PTM analysis by mass spectrometry 81
3. RESULTS **83**
3.1 The unfolded subproteome in maize embryos: enriching for LEA proteins 83
3.2 Two dimensional analyses of Emb564, Rab17 and Mlg3 LEA proteins: posttranslational modifications. 87
3.3 Recombinant Emb564, Rab17 and Mlg3 LEA proteins: in vitro antiaggregation effects and in vivo protection. 92
3.4 LEA-GFP fusions in tobacco epidermal cells: group 3 Mlg3 LEA protein reduces dehydration-induced cell shrinkage effects 95
3.5 Group 2 Rab17 LEA protein preserves cell viability under heat shock and colocalizes to oil bodies. 100
4. DISCUSSION **105**
4.1 Maize embryo LEA proteome 105
4.2 Search for posttranslational modifications 105
4.3 Experiments in E.coli: differences in anti-aggregation and protective properties among LEA proteins 106
4.4 LEA-GFP fusion proteins in plant cells 107

5. CONCLUSIONS **111**

V. CHAPTER 3 115

1. INTRODUCTION **117**
2. MATERIALS AND METHODS **119**
2.1 Isolation and plasmid construction of Rab28 gene 119
2.2 Generation of Rab28 transgenic maize 119
2.3 Molecular analysis of transgenic plants. 120
2.4 Protein extraction, western blot and LEA antibodies 121
2.5 Stress treatments and measurement of physiological parameters 121
2.6 Nuclei isolation and fractionation 122
2.7 Immunocytochemistry 122
3. RESULTS **123**
3.1 Molecular analysis of transgenic maize plants 123
3.2 Phenotypic traits of Rab28 Transgenic Plants 125
3.3 Overexpression of Rab28 enhances resistance to osmotic stress in transgenic maize plants 126
3.4 Physiological parameters in Rab28 transgenic plants under water stress 129
3.5 Nucleolar localization of Rab28 protein in dry embryo cells and in transgenic maize roots 131
4. DISCUSSION **135**
5. CONCLUSION **138**

VI. GENERAL DISCUSSION 141

VII. CONCLUSIONS 153

VIII. SPANISH SUMMARY 157

IX. BIBLIOGRAPHY 171

Abbreviations

ABA:	Absicic Acid
bp:	Base pairs
CFU:	Colony Forming Units
CK2:	Casein kinase II
E.coli:	Escherichia coli
FTIR:	Fourier-transform infrared spectroscopy
GFP:	Green Fluorescent Protein
IPTG:	Isopropyl-beta-thio galactopyranoside
KDa:	kdalton
LEA proteins:	Late Embryogenesis Abundant proteins
LC-ESI-MSMS:	MS analysis using an electrospray ionization source coupled on-line to liquid chromatography
LC-MALDI-MSMS:	HPLC analysis via matrix assisted laser desorption/ionization
MDA:	Malondialdehyde
MS:	Mass spectrometry
MS medium:	Murashige and Skoog medium
OD:	Optical Density
PCR:	Polymerase Chain Reaction
PEG:	Polyethylene glycol
PTM:	Posttranslational modification
RT-PCR:	Reverse transcription polymerase chain reaction
RWC:	Relative Water Content
SDS-PAGE:	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
TCA:	Trichloroacetic acid
TLA:	Total leaf area
WT:	Wild-type
1D gel:	One-dimensional gel electrophoresis
2D gel:	Two-dimensional gel electrophoresis

INTRODUCTION

I. INTRODUCTION

1. Abiotic stress in plants: Cellular and molecular mechanisms

Plants are exposed to multiple environmental stresses along their life cycle. Abiotic stresses such as drought, high salinity and freezing temperatures affect most areas of the world and they impact in plants by directly reducing its survival in the natural environment and its productivity in agriculture. About half of the annual world crop production is lost due to abiotic stresses, especially drought (Boyer 1982; Vinocur and Altman 2005). Most plants encounter transient decreases in relative water content at some stages of their life, and many produce highly-desiccation tolerant structures, such as seeds, spores or pollen. Although in a few vascular plants, i.e. resurrection plants, desiccation tolerance also occurs in the vegetative tissues (Ingram and Bartels 1996) in general, plant vegetative tissues, leaves and roots, are very sensitive to water deficit and they can only withstand reduced and transient water losses (Blum 2011; Bray 1993).

Drought stress induces a range of physiological, cellular and molecular responses in plants towards stress tolerance. Physiological drought occurs not only by water deficit but it is also produced during cold and salt stresses. Freezing reduces the liquid water in extracellular fluids and can lead to intracellular dehydration; increased extracellular ion concentration also reduces intracellular water availability. Thus, it is not surprising that in plants drought, high salinity and freezing temperatures share common stress tolerance responses (Mian et al. 2011). At whole plant level, water stress in plants occurs when transpirational demands exceed water absorption by roots, affecting enzymatic reactions and physiological processes. The first response to water stress is a decrease in leaf turgor, which reduces cell elongation. This is followed by a decrease in protein synthesis and cell division. Stomatal closure behind loss in turgor is the primary cause for reductions in photosynthesis under mild water stress under severe drought; reductions in biochemical reactions generally follow reductions in photosynthetic capacity (Blum 2011; Chaves 1991).

Cellular turgor is not the only important transducer of plant water stress. The growth regulating hormone abscisic acid (ABA) is produced in response to desiccation, causing many of the known expressions and consequences of plant water deficit such as arrested growth, stomatal closure and reproductive failure (Blum 2011). Drought and high salinity result in strong increases of

plant ABA levels which induce the expression of stress-related genes and adaptive physiological responses (Cramer et al. 2011; Raghavendra et al. 2010). In seeds, ABA is a major factor regulating seed dormancy. ABA-deficient seeds as the viviparous mutants in maize show reduced dormancy whereas overexpressing of ABA biosynthetic enzymes leads to enhanced dormancy (Finkelstein et al. 2008). The transition between seed dormancy and germination is controlled by a dynamic balance of synthesis and catabolism of two antagonistic hormones, ABA and Gibberllins (GA) (Seo et al. 2009). Embryonic ABA inhibits the germination by suppressing GA-responsive genes essential for seed germination and seedling growth. This transition between seed dormancy and germination is a critical stage in the plant life cycle and therefore an important ecological and commercial trait (Finkelstein et al. 2008). Since many abiotic stresses ultimately result in desiccation of the cell and osmotic imbalance, there is an overlap in the expression pattern of stress genes after cold, drought, high salt or ABA application. Various stress signals and ABA share common elements in their signaling pathways, with common elements that cross talk with each other in order to maintain cellular homeostasis, (Cramer et al. 2011; Wasilewska et al. 2008). Drought, salt and cold stress produce extensive changes in the regulation of gene expression, gene activation/gene suppression, signal transduction pathways biochemical modulation and proteomic machinery which lead to the survival or death of the affected plants (Ahuja I 2010; Xiong et al. 2002; Yamaguchi-Shinozaki and Shinozaki 2006). In Arabidopsis gene networks of stress-inducible genes have been identified and have been classified into two groups: regulatory and functional genes (Cramer et al. 2011; Shinozaki and Yamaguchi-Shinozaki 2007) (Figure 1). The first group comprises regulatory proteins, i.e. protein factors involved in further regulation of signal transduction and stress responsive gene expression. These include various transcription factors, protein kinases, protein phosphatases, enzymes involved in phospholipid metabolism, and other signaling molecules such as calmodulin-binding protein. Many transcription factor genes are stress inducible, suggesting that various transcriptional regulatory mechanisms may function in regulating drought, cold, or high salinity stress signal transduction pathways. The second gene group includes functional genes that encode protective proteins that function in abiotic stress tolerance. These include molecules such as Late Embryogenesis Abundant (LEA) proteins, chaperones, osmotin, antifreeze proteins, mRNA-binding proteins, and key enzymes for osmolyte biosynthesis, water channel proteins, sugar and proline transporters, detoxification enzymes and various proteases.

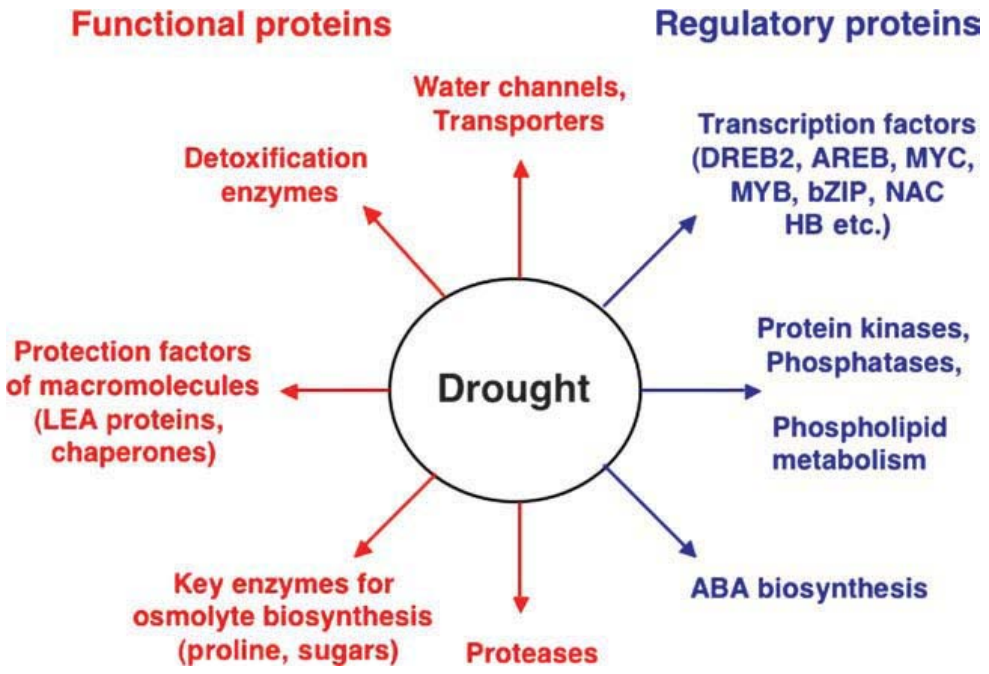


Figure 1. Functions of drought stress-inducible genes in stress tolerance and response. Gene products are classified into two groups**.** (Shinozaki and Yamaguchi-Shinozaki 2007)

The first group contains protein factors involved in regulation of signal transduction and gene expression that function in stress response (regulatory proteins) and the second group includes proteins that function in stress tolerance (functional proteins).

2. Desiccation tolerance

Desiccation tolerance is the ability to survive the removal of all, or almost all the cellular water without irreversible damage (Leprince and Buitink 2010). Desiccation tolerance constitutes the exception of the established principle that life consists of chemical interactions that take place in liquid water. Two types of tolerance are distinguished. Drought tolerance is the capacity to tolerate moderate dehydration, with no bulk cytoplasmic water, which represents about 23% water on a fresh weight basis. Desiccation tolerance refers to tolerance to further dehydration, when the dehydration shell is gradually lost (Hoekstra et al. 2001). Drought tolerance at cellular level is based on structural stabilization by preferential hydration, whereas desiccation tolerance involves the replacement of water by molecules that also form hydrogen bonds. Desiccation-tolerant cells are capable of rehydrating successfully (Hoekstra et al. 2001). Certain plants, animals and micro-organisms are able to dry completely and yet remain viable. Anydrobiotic organisms include micro-organisms such as yeast, invertebrates such as tardigrades, rotifers and

some nematodes and in plants, resurrection plants, most plant seeds and pollen (Hand et al. 2011; Hoekstra et al. 2001).

During successful drying, a multifaceted and complex response takes place, including structural and component alterations of the cell wall, organelles or organs, induction of repair system, removal of free radicals and accumulation of macromolecules (Buitink and Leprince 2004; Shih 2008). This contrasts with desiccation-sensitive cell types and organisms which suffer damage in all main cell components on drying; proteins lose their correct conformation and can form toxic aggregates; membranes undergo phase transitions causing leakiness and fusion; RNA is likely to experience structural rearrangements and DNA, RNA and chromatin stability will be compromised (Kanias and Acker 2006).

Upon water deficit a common mechanism is the synthesis of osmotically active compounds compatible with metabolism. These are low-molecular weight solutes that accumulate in the intracellular compartment for osmotic adjustment. Among the many existing osmoprotectants, the non-reducing disaccharide trehalose is most widespread in nature, in anhydrobiotic bacteria, fungi, invertebrates and resurrection plants. However, in most plants systems, sucrose and certain oligosaccharides such as raffinose, stachyose and cyclitols, instead of trehalose, accumulate in large quantities during seed maturation. Non-reducing disacharides function as water replacement molecules and vitrification agents and contribute to the formation of bioglasses (Cacela and Hincha 2006). Intracellular bioglasses confer long term stability in the dry state. Non-reducing disaccharides are, however, not an absolute requirement for anydrobiosis (Tunnacliffe and Lapinski 2003; Wolkers et al. 1998; Wolkers et al. 1999). In seeds the major seed proteins involved in bioglass formation are late embryogenic abundant (LEA) proteins (Blackman et al. 1995; Buitink and Leprince 2004; Oliver et al. 2000).

3. Late Embryogenesis Abundant (LEA) proteins

As the name suggests, Late Embryogenesis Abundant proteins were originally discovered in the late stages of embryo development in cotton seeds (Dure et al. 1981; Galau et al. 1986). In plants, most of LEA proteins and their mRNAs accumulate to high concentrations in embryo tissues during the last stages of seed development when desiccation occurs (Baker J 1988; Bies-Etheve et al. 2008; Hundertmark and Hincha 2008; Ingram and Bartels 1996). Embryogenesis in flowering plants represents a series of stages to develop a miniature plant within the seed. In this process the formed embryo will undergo a cellular expansion stage with dry mass increase in order to provide energy for the process of germination. In most plants the final stage of seed

development, maturation, is initiated by a reduction in seed water content, which will eventually drop to about 10%. During this stage and preceded by an increase in ABA content, gene expression and protein profiles change greatly and are associated with the acquisition of desiccation tolerance and development of capacity for seed germination (Goldberg et al. 1989; Skriver and Mundy 1990). LEA proteins are accumulated in this final stage, in contrast to storage proteins which appear earlier. Moreover, their mRNAs are maintained at high levels in the dehydrated mature embryos, while transcripts of storage protein genes are completely degraded during the last embryogenesis stage (Goldberg et al. 1989). Since orthodox seeds acquire the ability to withstand severe dehydration at this stage, LEA proteins have been associated with desiccation tolerance (Cuming 1999; Dure et al. 1981; Galau et al. 1986).

LEA proteins accumulate in vegetative tissues exposed to dehydration, osmotic, and/or low-temperature stress (Bies-Etheve et al. 2008; Bray 1993; Dure et al. 1989; Hundertmark and Hincha 2008). They are also found in anydrobiotic resurrection plants upon drying (Bartels and Sunkar 2005; Ingram and Bartels 1996). Members of the LEA family seem to be ubiquitous in the plant kingdom. Since their first description, hundreds of LEA proteins from vascular to nonvascular plants have been isolated. Their presence has been confirmed not only in angiosperms and gymnosperms (Bray 1993) but also in seedless vascular plants (e.g. Selaginella) (Alpert 2005; Oliver et al. 2000), bryophytes (e.g. Tortula, Physcomitrella) (Oliver and Solomon 2004), pteridophytes (e.g. ferns) (Reynolds 1993) and algae (Honjoh et al. 1995). Some LEA-like genes are only induced by ABA or other environmental clues (Battaglia et al. 2008; Shih 2008). In addition they have now also been identified in some microorganisms (Garay-Arroyo et al. 2000; Stacy et al. 1999), fungi (Abba et al. 2006; Borovskii et al. 2000), protozoa, rotifers, nematodes (Solomon et al. 2000; Browne et al. 2004), insects and crustacean (Hand et al. 2011; Hoekstra et al. 2001; Tunnacliffe and Wise 2007). The correlation of LEA proteins in seed maturation stages, during water stress in vegetative plant organs, and in anydrobiotic animals suggests that LEA proteins represent a widespread adaptation to water deficit; however, their precise functions remain unclear.

4. Classification of LEA proteins: sequence motifs

LEA proteins were initially classified in six subgroups on the basis of specific domains (Dure et al. 1989). With increasing information available on new described members, differences on expression profiles, description in organisms other that plants and especially with the new bioinformatic tools, the classification has been subjected to different rearrangements (Battaglia et

al. 2008; Bies-Etheve et al. 2008; Hundertmark and Hincha 2008; Wise and Tunnacliffe 2004b). Here, we will adopt a modification by Covarrubias's group (Battaglia et al. 2008) of the classification initially introduced by Dure (Dure et al. 1989) in which LEA proteins from cotton were categorized by virtue of similarities in their deduced amino acid sequences. This classification is based on the presence of specific motifs conserved across species which are unique to each family. Based on these characteristics and considering all available sequence information from different plant species LEA proteins are grouped into seven distinctive groups or families; nevertheless, groups 1, 2 and 3 are considered the major LEA groups containing most members of the protein family. Over 100 entries for each group 1, 2 and 3 are found in public databases, Gen Bank (http://www.ncbi.nlm.nih.gov) and TIGR database (http://plantta.tigr.org) (Shih 2008). Table I shows a comparison of the nomenclature used in the present work with the corresponding Pfam classification and the nomenclature used by several authors. The first inventory of LEA proteins was performed in the Arabidopsis thaliana genome and fifty *lea* genes were identified in Arabidopsis genome (Bies-Etheve et al. 2008). In rice the lea genome comprises 33 genes, 36 in grapevine and 33 in poplar (Hundertmark and Hincha 2008) (Figure 2).

Table I. Correspondence between different nomenclatures given to LEA protein groups (Battaglia et al. 2008).

This Work	Dure	Bies-Ethève	PFAM	PFAM No.	Name
1	D-19	1	LEA_5	PF00477	Em1, Em6
2	D-11	2	Dehydrin	PF00257	Dehydrin, RAB
3A	D-7	3	LEA_4	PF02987	ECP63, PAP240, PM27
3B	D-29	3*	LEA_4	PF02987	D-29
4A	–	4	LEA_1	PF03760	LE25_LYCES
4B	D-113	4	LEA_1	PF03760	PAP260, PAP051
5A	D-34	5	SMP	PF04927	PAP140
5B	D-73	6	LEA_3	PF03242	AtD121, Sag21, lea5
5C	D-95	7	LEA_2	PF03168	LEA14
6	–	8	LEA_6	PF10714	LEA18
7	–	–	ABA_WDS	PF02496	ASR

–, Used to denote groups that were not identified by these authors.

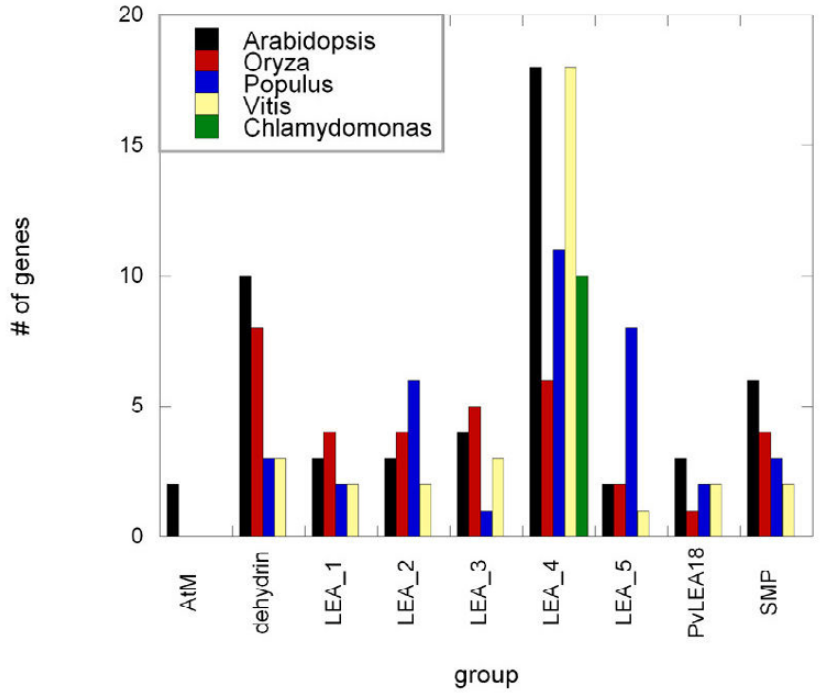


Figure 2: Comparison of the sizes of the different *LEA* gene groups *in* Arabidopsis, rice (Oryza), poplar (Populus), grapevine (Vitis) and Chlamydomonas (Hundertmark and Hincha 2008).

4.1 Group 1

This LEA group (Pfam PF00477) originally represented by the D-19 and D-132 proteins from cotton seeds contain an internal 20-mer sequence (TRKEQ [L/M] G[T/E] EGY[Q/K] EMGRKGG [L/E]) (Figure 3). This motif may be present in several copies arranged in tandem, from one to four in plant species, and up to eight in other organisms (Galau et al., 1993; Baker et al., 1995, Battaglia et al. 2008). The wheat Em proteins belong to this group (Cuming 1984). Group 1 LEA proteins have also been found in *Bacillus subtilis* (Stacy and Aalen, 1998), in other soil bacterial species and in the crustacean *Artemia franciscana*.

4.2 Group 2

This group of LEA proteins (Pfam PF00257), also known as‘‘dehydrins’’, was originally identified as the D-11family in cotton embryos. Group 2 LEA protein is the most characterized group of LEA proteins. A distinctive feature of group 2 LEA proteins is a conserved, Lys-rich 15-residue motif, EKKGIMDKIKEKLPG, named the K-segment (Campbell and Close 1997) which can be found in one to 11 copies within a single polypeptide (Figure 3). An additional motif also found in this group is the Y-segment, whose conserved consensus sequence is [V/T]D[E/Q]YGNP, usually found in one to 35 tandem copies in the N terminus of the protein; this motif has similar amino acid sequence to the nucleotide binding domain found in chaperones of plants and bacteria (Figure 3). Many proteins of this group also contain a tract of Ser residues, called the S-segment, acting as a site for protein phosphorylation (Vilardell et al. 1990). Less conserved motifs, the ø-segments, are usually rich in polar amino acids and lay interspersed between K-segments (Close 1996; Dure 1993).

The presence and arrangement of these different motifs in a single polypeptide allow the classification of group 2 LEA proteins into five subgroups (Scott and Close 1997). Proteins that only contain the K-segment are in the K-subgroup, and those that include the S-segment followed by K-segment are in the SK-subgroup. In addition, there are the YSK-, YK-, and KS-subgroups (Figure 3). To date group 2 proteins have been identified in a range of plants from different taxa, including angiosperms, gymnosperms, pteridophytes, bryophytes, fungi, algae and cyanobacteria (Shih 2008).

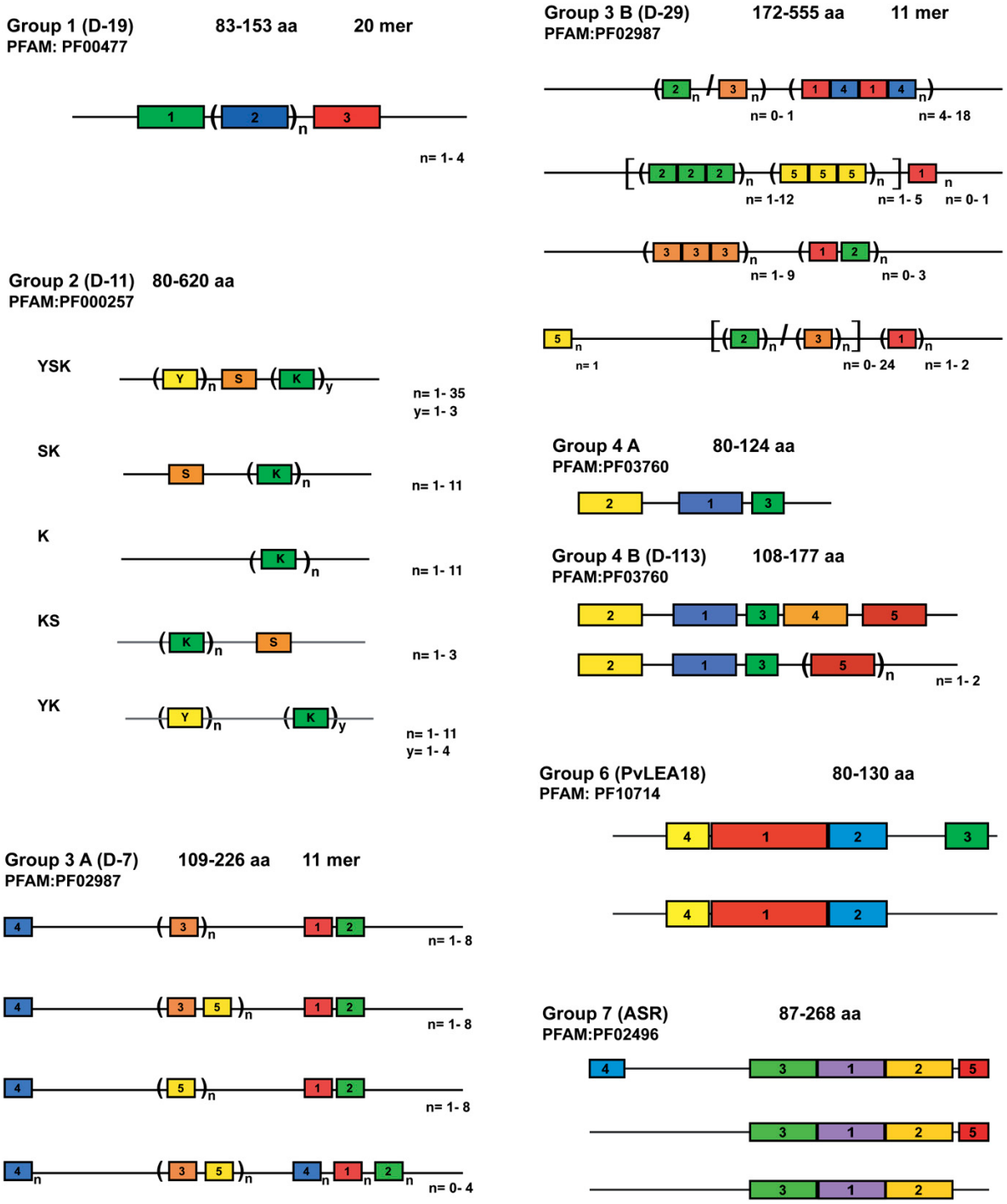


Figure 3: Array of the distinctive motifs in the LEA protein groups (Battaglia et al. 2008).

Each block contains a schematic representation of the arrangement of the motifs that distinguish each group of LEA proteins and their corresponding subgroups. Although similar colors and numbers indicate the different motifs for each group, they do not imply any sequence relation among the motifs in the different blocks. The range of protein sizes in each group is indicated at the top of each block, in number of amino acid (aa) residues.

4.3 Group 3

Group 3 LEA proteins (Pfam PF02987) are characterized by a repeating motif of 11 amino acids (Dure 1993). Differences found in the molecular mass in this group of proteins are usually a consequence of the number of repetitions of this 11-mer motif. In comparison with other groups of LEA proteins, the group 3 members are quite diverse.

The variability in the 11-mer motif leads to a subclassification of the group 3 LEA proteins into two subgroups: 3A, represented by the cotton D-7 LEA protein; and 3B, represented by the cotton D-29 LEA protein (Fig. 4). The first subgroup is highly conserved; two of the motifs characteristic of these proteins (motifs 3 and 5) correspond to almost the same 11-mer described originally for this subgroup, with some variation at positions 9 and 10 (TAQ [A/S] AK [D/E] KT[S/ Q] E). At the N-terminal portion of 3A proteins, we found motif 4 (SYKAGETKGRKT), and at the C-terminal portion, we found motifs 1 and 2 (GGVLQQTGEQV and AADAVKHTLGM). The other subgroup (3B) is more heterogeneous; four variations of the 11-mer were found (motifs 1–4), but the variability was restricted to the consensus sequence described above. Yet, motif 5 is highly conserved and is unique to this subgroup. Additionally, other conserved regions (motifs 1, 2, and 4 in subgroup D-7 and motif 5 in subgroup D-29), which may or not be repeated and whose sequences are completely different from the 11-mer, are also present.

Interestingly, proteins similar to plant group 3 LEA proteins accumulate in several non plant organisms in response to dehydration, in fungi, microbial and animal kingdoms. They have been found in prokaryotes *Deinococcus radiodurans* (Battista et al. 2001), *Haemophilus influenza* (Dure 2001) and in *Caenorhabditis elegans* (CeLEA-1), whose expression is correlated with the survival of this nematode under conditions of desiccation, osmotic, and heat stress (Gal et al. 2004). Other anhydrobiotic organisms such as the nematodes *Steinernema feltiae* (Solomon 2000) and *Aphelencus avenae* (Browne et al. 2004), as well as the bdelloid rotifer *Philodina roseola* (Tunnacliffe et al. 2005), the chironomid *Polypedilum vanderplanki* (Kikawada et al. 2006) and the crustacean *A. franciscana* (Hand et al. 2007; Wang et al. 2007) also accumulate group 3 LEA proteins in their desiccated states.

4.4 Group 4

Group 4 LEA proteins (Pfam PF03760) are of widespread occurrence in the plant kingdom, including nonvascular plants (bryophytes) and vascular plants (gymnosperms and angiosperms). A motif that has characterized the proteins in this group is motif 1, located at the N-terminal

region with the following consensus sequence: AQEKAEKMTA[R/H]DPXKEMAHERK [E/K][A/E][K/R]. However, four additional motifs can be distinguished in many group 4 LEA proteins. The presence or absence of motif 4 or 5 defines two subgroups within the family (Figure 3). The first subgroup (group 4A) consists of small proteins (80–124 residues long) with motifs 2 and/or 3 flanking motif 1. The other subgroup (group 4B) has longer representatives (108–180 residues) that, in addition to the three motifs in the N-terminal portion, may contain motifs 4 and/or 5 at the C-terminal region. D-113 protein from cotton, the first discovered of this group, belongs to group 4B.

4.5 Group 5

The first proteins described for this group (Pfam PF04927) were D-34, D-73, and D-95 from cotton (Baker J 1988; Cuming 1999). They represent an atypical LEA subgroup because they contain a higher proportion of hydrophobic residues (see below). They include maize Rab28 (Pla M 1991), carrot ECP31 (Kiyosue et al. 1992) and *Medicago trunculata* MtPM25(Boucher et al. 2010) among others (Figure 3).

4.6 Group 6

PvLEA18 protein from bean (*Phaseolus vulgaris*) was the first protein described from this group (Pfam PF03168) (Colmenero-Flores 1997). To date, 36 members of this family have been identified from different species of vascular plants. The proteins in this group are characterized by their small size (approximately 7–14 kD) and high conservation. Four motifs distinguish this group, two of which (motifs 1 and 2) are highly conserved. Noteworthy, the sequence LEDYK present in motif 1 and the Pro and Thr residues located in positions 6 and 7, respectively, in motif 2 show 100% conservation (Figure 3).

4.7 Group 7

The ASR proteins, considered to be members of the LEA family in some classifications, are small, heat-stable, and intrinsically unstructured proteins (Maskin et al. 2007; Silhavy et al. 1995). Several ASR genes have been identified from various species of dicotyledonous and monocotyledonous plants (Silhavy et al. 1995) as well as from gymnosperm species (Padmanabhan et al. 1997; Shen et al. 2005). However, no ASR-like genes are found in Arabidopsis. All known ASR proteins contain three highly conserved regions (motifs 1, 2, and 3; Figure 3). They share physiochemical properties with other LEA proteins and they accumulate in

seeds during late embryogenesis and in response to water-limiting conditions (Maskin et al. 2007).

5. Subcellular localization and expression profiles.

In plants, LEA proteins have been found localized in cytoplasm, nucleus, mitochondrion, chloroplast, endoplasmic reticulum, vacuole, peroxisome and plasma membrane (Tunnacliffe and Wise 2007). The different LEA groups show no preference for a specific subcellular localization. Most LEA proteins from the different groups are accumulated during embryo development in the dry seed.

In plants, group 1 LEA proteins are found mostly in seeds and they are not induced by stress conditions in vegetative tissues. They accumulate during seed development and they are considered as embryo-specific LEA proteins (Manfre et al. 2006; Manfre et al. 2009).

Group 2 LEA proteins accumulate during seed desiccation and in response to water deficit induced by drought, low temperature or salinity (Ismail et al. 1999; Nylander et al. 2001).

Most of them accumulate in all tissues upon water deficit although there are those that preferentially respond to particular stress conditions. Some dehydrins are strongly accumulated in response to low-temperature treatments but not to drought or salinity (Rorat et al. 2006); other group 2 members are not induced in response to low temperatures, while a small number of dehydrins show an unusual constitutive expression (Danyluk et al. 1998; Houde et al. 1992; Rorat et al. 2004). However, it is not possible to assign a specific accumulation pattern to particular group or subgroup. In addition, not all are ABA-induced and in some cases their response to stress is mediated by more than one pathway, one of which may be ABA dependent (Giordani et al. 1999; Nylander et al. 2001; Welling et al. 2004). Many group 2 LEA proteins accumulate in the cytoplasm, and some of them are also localized to the nucleus. For nucleus-directed SK2 proteins, the phosphorylated S-segment and the RRKK sequence are relevant for their nuclear localization (Plana et al. 1991; Riera et al. 2004). Some dehydrins are also found in other cell compartments, including the vicinity of the plasma membrane, mitochondria, vacuole, and endoplasmic reticulum (Borovskii et al. 2002; Borovskii et al. 2000; Danyluk et al. 1998; Heyen et al. 2002; Houde et al. 1992).

Expression and transcriptomic analysis of plant group 3 LEA proteins show their accumulation in mature seeds and in response to dehydration, salinity, or low temperatures (Cattivelli and Bartels 1990; Hsing et al. 1995; Romo et al. 2001). Some members also respond to hypoxia (Siddiqui et al. 1998) or to high-excitation pressure imposed by high light (NDong et al. 2002).

As for LEA proteins from other groups, the expression of group 3 LEA proteins appears to be regulated by ABA during specific developmental stages and/or upon stress conditions (Curry et al. 1991; Ried and Walker-Simmons 1993). Group 3 D-7 LEA protein from cotton accumulates to a concentration of about 200 mM in mature cotton embryos (Roberts et al. 1993). Studies of seeds have localized group 3 LEA proteins to the cytoplasm and protein storage vacuoles, as is the case for HVA1 from barley Hordeum vulgare (Marttila et al. 1996) whereas PsLEAm is distributed within the mitochondrial matrix of pea seeds (Grelet et al. 2005). Group 3 proteins are also detected in vegetative tissues; WAP27A and WAP27B are abundantly accumulated in endoplasmic reticulum of cortical parenchyma cells of the mulberry tree (*Morus bombycis*) during winter (Ukaji et al. 2001); and WCS19 accumulates specifically in wheat leaves and rye (*Secale cereale*) during cold acclimation, where it was localized within the chloroplast stroma (NDong et al. 2002).

LEA proteins from group 4 in some classifications are partially included in group 3. Cotton D-113, was found homogeneously distributed in all embryo tissues at a concentration of nearly 300 mM (Roberts et al. 1993). Later, similar proteins were found to accumulate in vegetative tissues in response to water deficit. In tomato (*Solanum lycopersicum*) plants, group 4 LEA transcripts (LE25) accumulated in leaves in response to water deficit and ABA (Cohen et al. 1991). In Arabidopsis vegetative tissues, the transcripts of the group 4 LEA proteins also accumulated in response to water-deficit treatments (Delseny et al. 2001).

Transcripts from group 5 LEA proteins accumulate during the late stage of seed development and in response to stress conditions, such as drought, UV light, salinity, cold, and wounding (Kiyosue et al. 1994; Maitra and Cushman 1994; Stacy et al. 1999; Zegzouti et al. 1999). In maize, Rab28 has been found accumulated in the nucleolus of scutellar cells of mature dry embryos (Niogret et al. 1996). In *Medicago trunculata*, MtPM25 is highly induced by drought (Boudet et al. 2006b).

6. Biochemical properties and structure of LEA proteins

Most of the biochemical properties of LEA proteins arise from their hydrophilic nature and byased amino acid composition. Although significant similarities have not been found between the members of the different groups, a unifying and outstanding feature of most of them is their high hydrophilicity and high content of Gly and a lack or low proportion of Cys and Trp residues, and a preponderance of certain amino acid residues such as Ala, Glu, Lys/Arg, and Thr (Baker 1988; Dure 1993; Garay-Arroyo et al. 2000) (Figure 4).

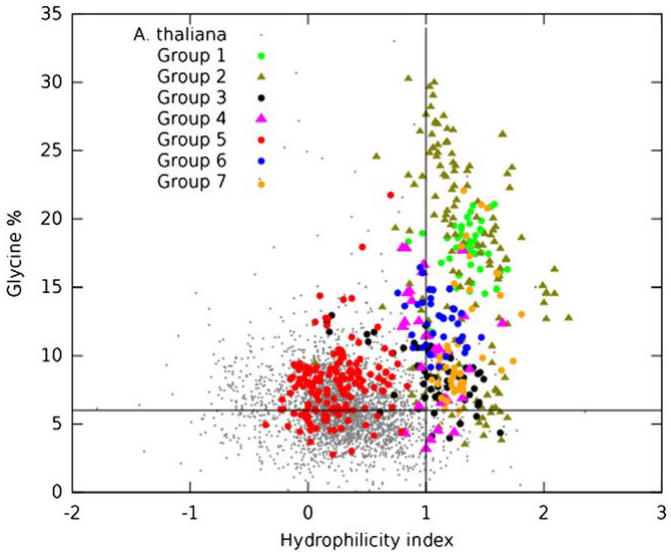


Figure 4: Graphic representation of LEA proteins according to the properties that define the hydrophilins (Battaglia et al. 2008).

All points on the top right quadrant correspond to proteins that can be considered as hydrophilins according to their definition (hydrophilicity index ≥ 1, Gly ≥ 6%). 3,000 randomly chosen Arabidopsis proteins are shown as reference (grey). To obtain the hydrophilicity index, the average hydrophobicity of all amino acids in each protein was multiplied by -1 (Kyte and Doolittle, 1982).

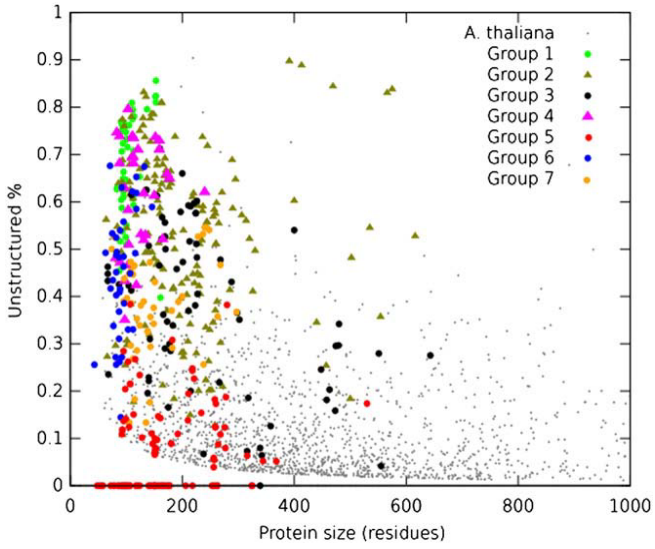


Figure 5: Protein size and disorder prediction of LEA proteins belonging to different groups (Battaglia et al. 2008).

3,000 randomly chosen Arabidopsis proteins are shown as reference (grey). Many of them (1,350 of 3,000) cannot be seen because they overlap the x axis, as they were predicted to be 0% unstructured. The unstructured fraction was calculated according to Dosztányi et al. 2005, using a sliding window of 21 amino acids.

The high hydrophilicity is likely to be responsible for their lack of conventional secondary structure in the hydrated state. Most of them exist principally as randomly coiled proteins in solution. While structure modeling and structure prediction programs suggest that at least some LEA proteins from particular families contain defined conformations (Close 1996; Dure 1993; Dure et al. 1989) all hydrophilic LEA proteins studied experimentally have revealed a high degree of unordered structure in solution. This has led them to be considered as intrinsically unstructured proteins (Figure 5) (Goyal et al. 2003; Ismail et al. 1999; Kovacs et al. 2008; Shih 2004; Soulages et al. 2003; Soulages et al. 2002; Tompa 2005; Wolkers et al. 2001). Similarly, the ability of LEA proteins to remain soluble at elevated temperatures can be attributed to their hydrophilic, unstructured nature. Heat-induce aggregation of proteins results from partial denaturation and association through exposed hydrophobic regions, something that cannot occur in a protein which is hydrophilic and natively unfolded (Tunnacliffe and Wise 2007).

LEA proteins are variable in size, ranging from 5 to 77 kD among most groups. They can be acid, neutral or basic. Group 1 proteins are acidic to neutral; group 2 comprises proteins with different pIs and group 3 are neutral to basic (Shih 2008). Typically, during SDS-PAGE, they migrate at a higher molecular mass than the one predicted from their deduced amino acid sequences due to their unstructured nature.

Since the discovery than group 1 protein Em, from wheat, was flexible in conformation with little secondary structure (ie. α-helice or β-sheet) and 70% of the protein behaving as random coil (McCubbin et al. 1985) most LEA proteins have been found to have an unfolded structure in the hydrated state (Tunnacliffe and Wise 2007). LEA proteins from groups 1, 2 and 3 are predicted to be at least 50% unfolded. Lack of conventional secondary structure means that members of the major LEA protein groups are included in the large class of protein variously called “natively unfolded”, “intrinsically disordered “or “intrinsically unstructured” (Dunker et al. 2001; Tompa 2002; Uversky et al. 2000) and it is thus not surprising that attempts to crystallize purified LEA proteins for X-ray crystallography have reportedly been unsuccessful (Goyal et al. 2003; McCubbin et al. 1985). Disorder prediction programs (http://www.pondr.com) suggest that 27-41% of all eukaryotic proteins contain unstructured region ≥ 50 residues long and that 6-17% of polypeptides are wholly disordered (Dunker et al. 2001); therefore, many proteins are natively unfolded or contain natively unfolded domains.

Such absence of structure has implications for the LEA function: it is unlikely that they have catalytic function. Nevertheless, some LEA proteins show some secondary structure and they

have some structural elements in equilibrium with unstructured states. Thus, a polyproline-type II (PPII) extended, left-handed helices have been described in groups 1 and 2 from soybean (Soulages et al. 2002), group 2 from Arabidopsis (Mouillon et al. 2006) and in group 3 from an anydrobiotic nematode (Goyal et al. 2003).

Many natively unfolded proteins are known to undergo increased folding under some conditions, usually when they bind a partner molecule or cation (Uversky et al. 2000). Environmental conditions can also affect folding and several LEA proteins become more structured when dried. LEA proteins exhibit the remarkable ability to become more ordered and to develop secondary structure as dehydration proceeds. For animal LEA proteins, Tunnacliffe's group first demonstrated this phenomenon by using Fourier-transform infrared (FTIR) spectroscopy (Goyal et al. 2003). FTIR spectroscopy allows for the assessment of protein secondary structure in the dry state by using the profile of the amide-I band, which provides information on the relative contributions of α-helix, β-sheet, and turn structures. A group 3 LEA protein from *Typha latifolia* become largely α-helical when dried rapidly; slow drying resulted in intermolecular β-sheet formation, as well as α-helix (Wolkers et al. 2001). Similarly, the group 3 LEA protein AavLEA1from nematode *Aphelenchus avenae* (Goyal et al. 2003) and LEAM from pea mitochondria (Tolleter et al. 2007) also gain structure on drying. Boudet et al. (Boudet et al. 2006b) used FITR to study both group 1 (MtEm6) and group 5 (MtPm25) proteins from *M. trunculata* and found them both to have increased folding in the dried state. Koag et al. (Koag et al. 2003) have also reported gain of structure, α-helical, when cowpea DHN1 is incubated with small unilamellar vesicles (SUVs).

A recent modeling study by Li and He (Li and He 2009) utilized a 66-amino-acid fragment of AavLEA1 and documented, through molecular dynamics simulation, many of these properties. Water was removed from 83.5 wt% to 2.4 wt% (Figure 6). As water is removed, the protein assumes progressively a more folded conformation. At 83.5 wt% LEA protein is completely solvated. At 50% water between 83.5% and 50.4%, the protein is unstructured. In this range and below this point, water molecules no longer are sufficient to fully solvate the protein. At less than 20% water the protein becomes more dehydrated and begins to adopt a significant amount of secondary structure. α-helical structure is apparent, and hairpin-like structures are formed (Figure 6). At 2.4% water the structure is very similar to that in the complete absence of water. The structural changes observed at very low water percentages (Goyal et al. 2003; Li and He 2009; Tolleter et al. 2007) suggest a functional role for the protein in the dry state rather than in the hydrated state. The propensity of some LEA proteins to gain structure under some

conditions, may be a general property of these proteins, and may have important functional implications in their physiological roles.

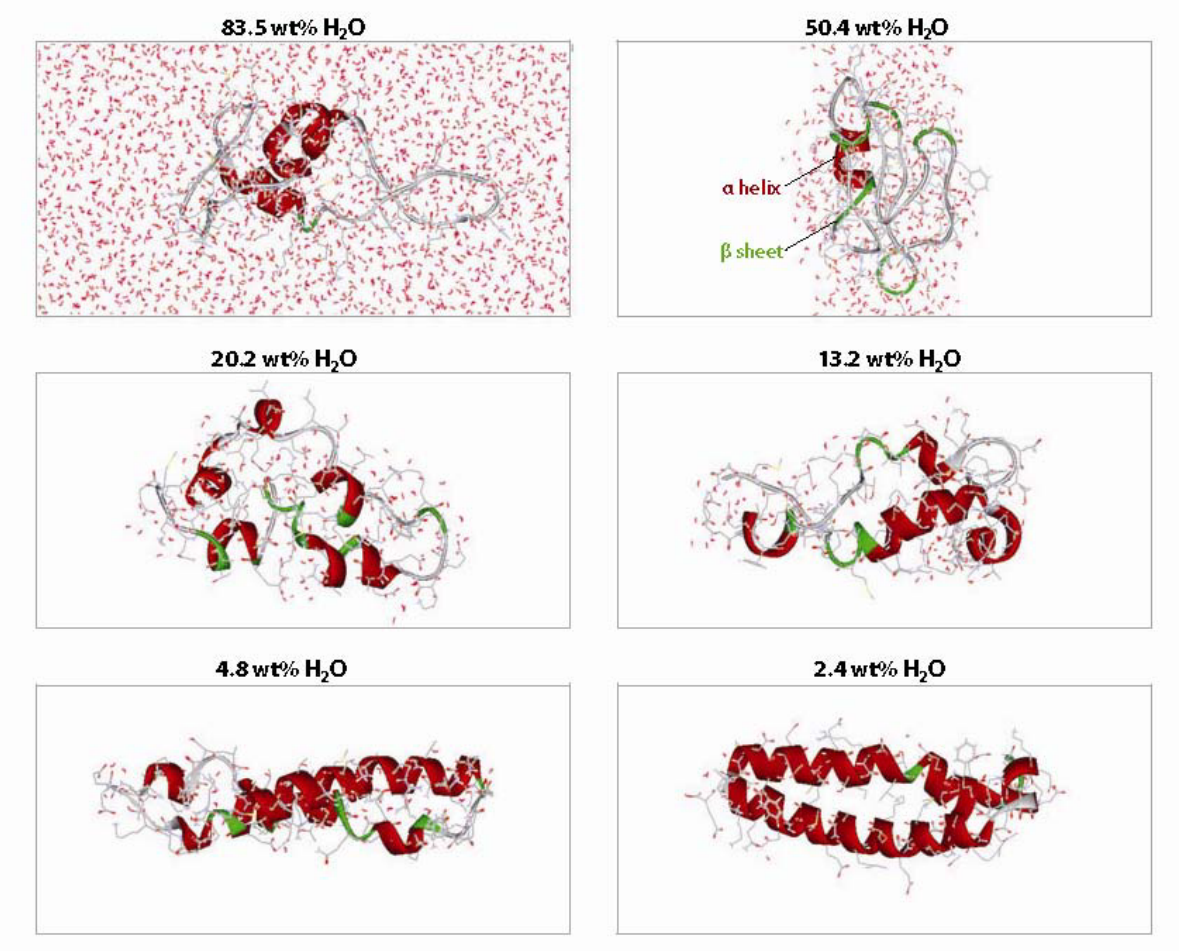


Figure 6. Representative conformations of the 66-amino-acid fragment of a LEA protein (AavLEA1) from the nematode *Aphelenchus avenae* are shown at different water contents (Hand et al. 2011 modified from Li and He 2009).

The smaller water molecules (grey and red) are depicted in the line style, and the larger LEA protein molecules are denoted using the solid ribbon style (α-helix, red; β-sheet, green; random coil, grey)

7. LEA protein functions

There is an extensive bibliography showing the correlation between the expression of LEA proteins or their genes with stress resistance (Hand et al. 2011; Shih 2008; Tunnacliffe et al. 2010). Many studies show the protection conferred by LEA proteins during salt and osmotic stress. Introduction of heterologous LEA proteins into plants and microorganisms results in an enhanced stress tolerance. Transgenic approaches have shown that over-expression of LEA proteins from different species in Arabidopsis, tobacco, rice, wheat, lettuce or cabbage produces improve abiotic stress resistant phenotypes (Leprince and Buitink 2010). However, the precise molecular function of LEA proteins is still unclear and so far LEA proteins have been suggested to act as stabilizers, hydration buffers, membrane protectants, antioxidants, organic glass formers and/or ion chelators (Tunnacliffe and Wise 2007).

7.1 Protein protection

LEA proteins have the capacity to protect target proteins from inactivation and aggregation during water stress. A role in protein stabilization is supported by the fact that some LEA proteins preserve enzyme activity *in vitro* after desiccation or freezing (Goyal et al. 2005; Grelet et al. 2005; Reyes et al. 2005) . One mechanism of protection is the prevention of water stress-induced aggregation of proteins (Chakrabortee et al. 2007; Goyal et al. 2005; Kovacs et al. 2008; Nakayama et al. 2007).

Many proteins, including the enzymes citrate synthase and lactate dehydrogenase, form insoluble aggregates when dried or frozen, but aggregation is reduced in the presence of LEA proteins from groups 1, 2 and 3. Group 2 proteins also prevent protein aggregation from heat stress (Kovacs et al. 2008). This protein anti-aggregation activity extends to the protection of complex mixtures of proteins, such as the water-soluble proteomes of human and nematode cells (Chakrabortee et al. 2007). Due to their hydrophilic, unstructured nature, LEA proteins themselves are not susceptible to aggregation on desiccation, freezing or boiling (Tunnacliffe et al. 2010).

The antiaggregation properties of a group 3 LEA protein have been demonstrated also in living cells; mammalian cells overexpressing aggregation-prone proteins reduce the formation of aggregates in the presence of nematode *A. avena* LEA protein (Chakrabortee et al. 2007). The group of Tunnacliffe has proposed that LEA proteins may exert a “molecular shield” activity (Goyal et al. 2005; Wise and Tunnacliffe 2004a). In the increasing crowded environment of the

dehydrating cytoplasm LEA proteins could decrease the interaction between partially denatured polypeptides and avoid their aggregation. The shield proteins might also have a space-filling role and help to prevent collapse of the cell as its water is lost.

Another functional hypothesis is the chaperone activity (Kovacs et al. 2008). The anti-aggregation activity of LEA proteins resembles a "holding" molecular chaperone, which function in the cell would be to stabilize passively protein species in a partially unfolded state, preventing aggregation while the stress lasts. They resemble holding chaperones, in their functioning without ATP, in contrast to classical folding chaperones which require ATP. But they are distinct in that they lack structure and do not form transient complexes with their client proteins through hydrophobic surfaces, since they are hydrophilic. The two functional mechanisms may not be so distinct and in the context of desiccation tolerance both activities may contribute to damage avoidance (Tunnacliffe 2010).

7.2 Membrane protection

During desiccation membrane protection is essential to preserve the cellular and organellar integrity. Some LEA proteins could contribute with sugars to H-bonding networking and protect membranes in the dry state (Hoekstra et al. 2001). LEA proteins being highly hydrophilic are not expected to interact with cellular membranes in hydrated conditions, but interaction cannot be excluded through α-helices in a dehydrating cell.

Among group 2 LEA proteins, wheat WCOR414 and Arabidopsis LTi29 acidic dehydrins were immunodetected in the vicinity of the plasma membrane during cold acclimation (Danyluk et al. 1998; Puhakainen et al. 2004) and maize dehydrins were observed in association to membranes of protein and lipid bodies (Egerton-Warburton et al. 1997). Maize native and recombinant dehydrin DHN (Rab17) binds *in vitro* to anionic lipid vesicles and binding produces an increase in amphipathic α-helix (a structural element that interacts to membranes and proteins); both binding and gain in conformation were attributed to the K-segment (Koag et al. 2003; Koag et al. 2009). Other dehydrins, ERD10 and ERD14, bind *in vitro* to acidic phospholipid vesicles (Kovacs et al. 2008). Recently, the contribution of dehydrin K-segments from group 2 Lti30 LEA protein has been analyzed *in vitro* in the binding to membranes and the relevance of the flanking His side chains as regulators of the interaction between the K-segments and membranes in a ph-dependent manner has been shown (Eriksson et al. 2011).

Group 3 LEA protein, LEAM, is located in pea mitochondria and is able to interact with a membrane and afford protection in the dry state (Tolleter et al. 2007). When drying, this

unstructured polypeptide is able to fold into amphipathic α-helices. The interaction between LEAM and phospholipids and the protective effect of LEAM was demonstrated by differential scanning colorimetry using a liposome desiccation assay (Tolleter et al. 2007).

To summarize, although their hydrophylicity and lack of structure suggest that LEA proteins should act in soluble compartments of cells, specific LEA proteins upon folding during desiccation or freezing may contribute to membrane protection (Tunnacliffe 2010).

7.3 Ion binding and antioxidant function

One consequence of dehydration is the increase in concentration of intracellular components, including ions. Increased ionic concentration can affect macromolecular structure and function. It has been proposed that LEA proteins, because of their many charged amino acid residues might act to sequester ions (Danyluk et al. 1998; Dure 1993). A dehydrin-like protein from celery is located in the vacuole and binds Ca++ when phosphorylated (Heyen et al. 2002). Acidic group 2 proteins, ERD10, COR47 and ERD14 also exhibit phosphorylation-dependent Ca++ binding. Phosphorylation sites are located in the serine motif (Alsheikh et al. 2005). Group 2 LEA proteins also bind a number of other metal ions; the interaction was proposed to occur through His residues which are over-represented in most group 2 LEA proteins (Svensson et al. 2000). Binding of metals ions by group 2 LEA proteins may be linked to the antioxidant properties reported for citrus CuCOR19 protein and *in vitro* scavenging activity for hydroxyl radicals (Hara et al. 2005). LEA proteins might reduce oxidative stress in dehydrating cells by scavenging ROS and/or by sequestering metal ions that generate ROS (Tunnacliffe and Wise 2007)

7.4 Other functions

LEA proteins might act as hydration buffers, slowing down the rate of water loss during dehydration; during partial drought, osmotic or freezing stress, hydration buffers allow sufficient water activity for proteins to retain function (Cuming 1999; Garay-Arroyo et al. 2000). Using a knockout mutant of Arabidopsis, whose seed exhibited premature dehydration, a role for group1 LEA protein Atm6 as hydration buffer was proposed (Manfre et al. 2006).

In a desiccating cell, when the water content falls below 10% on a dry weight basis, the cytoplasm vitrifies and enters in the “glassy state” (Buitink and Leprince 2004). In plants the formation of intracellular glasses is indispensable for survival in the dry state (seeds and pollens). LEA proteins accumulate to high levels in seeds (2-4% of the water soluble proteome) (Roberts et al. 1993) and they increase the density of the sugar glasses by strengthening the H-

bonding network (Buitink and Leprince 2004). Thus a potential role of LEA proteins is their contribution to the formation of biological glasses.

While among LEA proteins, some proteins which display similar sequences show different structure and *in vitro* properties, a prevalent issue is that a single LEA protein might have more than one function. The chloroplast LEA-like protein COR15am protects both membranes and proteins (Lin and Thomashow 1992; Nakayama et al. 2007); the mitochondrial group 3 protein LEAM, also exhibits both functions (Tolleter et al. 2007); group 2 from *Citrus* shows ionic binding, antioxidant and nucleic acid binding properties (Hara et al. 2005; Hara et al. 2009).This versatility may be a general feature of LEA proteins. Performance of more than one function,"moonlightening", is not uncommon among proteins and it is more likely to evolve in unfolded, rather than in folded proteins (Tompa 2005).

8. Biotechnological applications of LEA genes

The effects of over-expressing LEA genes from all LEA groups from different species of origin into different plant hosts is shown in Table II. In general, the phenotypes of the transgenic plants show enhanced stress tolerance, often related to drought or salt stress. Most studies report enhanced growth rates and reduced wilting of the aerial parts under stress under laboratory conditions and in some field trials, demonstrating a real potential of LEA proteins in engineering crops more tolerant to stress (Leprince and Buitink 2010). Apart from agronomical purposes, LEA proteins could be useful for other biotechnological applications in relation to their capacity to prevent aggregation of proteins. The use of a group 3 LEA protein as a fusion partner facilitates recombinant expression of recalcitrant proteins in *E.coli* in a soluble form (Singh et al. 2009). Moreover, the antiaggregation properties of another group 3 LEA protein have been applied to reduce the formation of *in vivo* aggregates; thus, coexpression of aggregation-prone proteins containing long polyglutamine (polyQ) or polyalamine (polyA) sequences, with group 3 AavLEA1 LEA protein in mammalian cells reduces substantially the expansion of protein aggregates associated to neurodegenerative diseases (Chakrabortee et al. 2007).

LEA gene				Species of origin	Host	Phenotypes
Pfam	Dure	Seed specific[a]	Gene name			
PF00477 LEA_5	D-19	Yes	*PMA1959*	Wheat	Rice	Reduced ion leakage in leaves after salt stress, increased plant height and shoot biomass
PF00257 dehydrin	D-11	No	*PMA80*	Wheat	Rice	Reduced ion leakage in leaves after salt stress, increased plant height and shoot biomass
			Rab17	Maize	*Arabidopsis*	Higher sugar and proline content after drying, more tolerant to salinity and recovery after mannitol stress
			Rab16A, Rab21[b]	Rice	Tobacco	Sustained growth rates under salinity, increased osmolyte production, reduced H_2O_2 levels and lipid peroxidation under drought stress
			pcC6-19	*Craterostigma plantagineum*	Tobacco	No obvious drought-tolerant phenotype
			RAB18/CORZ[c] *LTI29/LTI30*	Arabidopsis	Arabidopsis	Improved freezing tolerance, no effect on drought tolerance
			CuCOR19	Citrus ssp.	Tobacco	Enhancement of cold tolerance
PF02987 LEA_4	D-7	No	*HVA1*	Barley, chinese cabbage	Rice, wheat, chinese cabbage, creeping bentgrass	Increased biomass during water stress, decreased leaf wilting and membrane leakage
			pcC 3-06	*Craterostigma plantagineum*	Tobacco	No obvious drought-tolerant phenotype
			OsLEA3-1	Rice	Rice	Decreased spikelet fertility, no effect on absolute grain yield during drought and salt stress but increased biomass
PF03760 LEA_1	D-113	No	*BnLEA4-1*	Oilseed rape	Arabidopsis, lettuce, radish, bean	Enhanced growth ability under salt and drought stress
			BhLEA1,2	*Boea hygrometrica*	Tobacco	Higher relative water content, increased photosystem II activity (Fv/Fm), decreased electrolyte leakage and increased peroxidise and superoxide dismutase activity during drought stress
PF04927 SMP	*D-34*	Yes	*Atrab28*	Arabidopsis	Arabidopsis	Improved germination under salt and desiccation stress
PF03168- LEA_2	*D-95*	No	*CaLEA6*	Hot pepper	Chinese cabbage	Higher level of maximal photosynthetic rate of O_2 evolution (Pmax) during drought, salt, and heat.
PF03242 LEA_3	*D-73*	No	*DQ663481*[d]	*Tamarix androssowii*	Tobacco	Lower malonyldialdehyde content and conductivity; improved growth rate upon drought stress

Only those genes corresponding to PFAM domain associated to a LEA protein were taken into account. For comparison purposes, Dure's classification is also given. (a) Seed specific expression according to Iling et al.(Iling et al. 2005); (b) two names exist for this gene; (c) genes tested combination; (d) accession number because there is no name given for this gene.

Table II: Survey of functional genomics studies aiming at testing the effect of over-expressing LEA protein genes towards crop improvement (Leprince and Buitink 2010)**.**

9. Group 2 Rab17 and group 5 Rab28 LEA proteins from maize: a brief summary.

Maize rab17 gene (NCBI accession gi/126632258) encodes for a dehydrin protein which is highly accumulated in mature maize seeds and can be induced in immature embryos upon abscisic acid (ABA) incubation; in vegetative tissues Rab17 is readily induced by water stress and ABA (Goday et al. 1988; Pla et al. 1989).

The protein is a glycine rich protein with high content of polar amino acids and absence of cystein and tryptophan. Rab17 protein contains the lysine rich repeat K-segment (KIKEKLPG) twice in the carboxy terminus, and a cluster of eight serines, the S-segment, present in many but not all group 2 LEA proteins. The protein also holds a Y-domain (DEYGNP) similar to the nucleotide-binding site of plant and bacterial chaperones. The remaining protein consists of ø domains, predicted to form random coils. The molecular weight is 17 kD but in SDS gels it runs at 23 kD.

Rab17 is the most abundant phosphorylated protein in mature maize embryos. Rab17 phosphorylation occurs in the S-segment and it is regulated by CK2 (Plana et al. 1991; Riera et al. 2001). Due to its phosphorylation Rab17 runs in two-dimensional gels as a multisport pattern in the acidic end of the gel, although its theoretical pI is 8.75. In embryo cells, the protein is distributed in the cytoplasm and nucleus (Goday et al. 1994). By transgenic approaches the contribution of Rab17 protein to water deficit tolerance was shown in Arabidopsis plants over-expressing the dehydrin under a constitutive promoter; the generated plants were more tolerant to osmotic stress (Figueras et al. 2004). The use of a mutated version of the protein, mRab17, carrying an alteration in the EDD consensus site of CK2 has shown that absence of S-segment phosphorylation favors the retention of the nuclear in the nucleolus and reduces the stress-induced germination arrest observed in transgenic Arabidopsis plants expressing the phosphorylated Rab17 (Riera et al. 2004). In the present study we will use Rab17 as representative of maize group 2 LEA protein.

Group 5 *Rab28 lea* gene (NCBI accession gi/22459) was initially isolated from a maize embryo cDNA library (Pla et al. 1991). It is expressed along the embryogenesis and it is induced by dehydration and ABA in maize vegetative tissues (Pla et al. 1991). It encodes for a polypeptide of 278 amino acids, with a predicted molecular of 27.712 kDa and apparent molecular weight of 30kda, and theoretical pI of 6.01 (Niogret et al. 1996). In dry mature embryos Rab28 protein is basically located in the nucleolus (Niogret et al. 1996). As other members of the LEA group 5 the protein holds the QPRRP nuclear localization signal (Borrell et al. 2002). The Arabidopsis

ortholeg of Rab28, AtRab28, has been found associated preferentially with embryo provascular tissues (Arenas-Mena et al. 1999) and by immunoelectron microscopy the protein was also detected in the nucleus (Borrell et al. 2002). Arabidopsis transgenic seedlings over-expressing Atrab28 showed increased cation tolerance (Borrell et al. 2002).

II. Aims of the present work

In order to improve our understanding on LEA proteins and their molecular functions in drought tolerance, the present work analyzes in the first place, the composition of LEA subproteomes from Arabidopsis seeds and maize embryos; second, three maize embryo LEA proteins from groups 1, 2, and 3 are analyzed in order to detect functional differences among them and finally, transgenic maize plants over-expressing group 5 *rab28* lea gene are characterized. The following results are presented:

Chapter 1. Proteomic approach to analyze the composition of LEA subproteomes from Arabidopsis seeds by mass spectrometry. The main objective was the development and isolation method to obtain enriched LEA populations from Arabidopsis seeds. LEA subproteomes were obtained using an extraction procedure that combines heat stability and acid solubility of LEA proteins. To identify the protein content, we followed two approaches: first, a classical 1D (SDS-PAGE) gel-based procedure associated with MS analysis using an electrospray ionization source coupled on-line to liquid chromatography (LC-ESI-MSMS) and second, a gel-free protocol associated with an off-line HPLC and analysis via matrix assisted laser desorption/ionization (LC-MALDI-MSMS).

Chapter 2. Proteomic analysis of LEA proteins accumulated in maize mature seeds. Identification of LEA protein content by mass spectrometry and selection of three LEA proteins, Emb564, Rab17 and Mlg3, as representatives of groups 1, 2 and 3 for further study. Comparative functional analysis covering different aspects of maize Emb564, Rab17 and Mlg3 proteins, post-translational modifications, subcellular localization and their properties in *in-vitro* and *in- vivo* assays.

Chapter 3. Characterization of transgenic maize plants expressing maize group 5 *rab28* LEA gene under the ubiquitin promoter. Evaluation of *Rab28* transcripts and protein levels, phenotype and stress tolerance traits of transgenic plants under drought stress. Investigation of the subcellular localization of transgenic and wild-type Rab28 protein using immunocytochemical approaches.

CHAPTER 1

III. Chapter 1

Chapter 1

LC-MSMS IDENTIFICATION OF ARABIDOPSIS THALIANA HEAT STABLE SEED PROTEINS: ENRICHING FOR LEA-TYPE PROTEINS BY ACID TREATMENT

Oliveira, E., Amara, I., Bellido, D., Odena, M.A., Dominguez, E.,

Pages, M. and Goday, A.

J. Mass Spectrom (2007) 42: 1485-1495

Abstract

Protein identification in systems containing very highly abundant proteins is not always efficient and usually requires previous enrichment or fractionation steps in order to uncover minor proteins. In plant seeds the identification of late-embryogenesis-abundant (LEA) proteins is often masked by the presence of the large family of storage proteins. LEA proteins are predicted to play a role in plant stress tolerance. They are highly hydrophilic proteins, generally heat-stable, and correlate with dehydration in seeds or vegetative tissues. In the present work we analyze the protein composition of heat stable *Arabidopsis thaliana* seed extracts after treatment with trichloroacetic acid (TCA). The composition of the proteins that precipitate and those that remain in solution in 3% TCA was analyzed by two different approaches: 1D SDS-PAGE coupled to LC-ESI-MSMS analysis and a gel-free protocol associated with LC-MALDI-MSMS.

Our results indicate that treating total heat-soluble extracts with 3% TCA is an effective procedure to remove storage proteins by selective precipitation and this fractionation step provides a soluble fraction highly enriched in LEA type proteins. The analysis and determination of protein identities in this acid soluble fraction by MS technology is a suitable system for large-scale identification of LEA proteins present in seeds.

Keywords: Heat stable proteins, LEA proteins, *Arabidopsis thaliana* seeds, acid-soluble proteins, LC-MSMS.

1. INTRODUCTION

Plants have evolved different physiological, biochemical and molecular mechanisms to tolerate environmental stress conditions. Among others, a large set of proteins so called late embryogenesis abundant (LEA) proteins accumulate naturally in some desiccation tolerant plant structures such as the seed and they are also produced in plant vegetative tissues during exposure to abiotic challenges (drought, salinity and cold) (Bray 1997; Ingram and Bartels 1996; Roberts et al. 1993). The specific functions of LEA proteins are still unclear but their presence correlates with the acquisition of stress tolerance (Bartels and Salamini 2001; Cuming 1999). The protective role of LEA proteins has been proposed to take place through different mechanisms such as the binding or replacement of water, ion sequestration, maintenance of protein and membrane structure acting as developmental regulators or a combination of them (Tunnacliffe and Wise 2007).

The improvement in stress tolerance by the protective effect of LEA-type proteins has been reported by different approaches, in *in vitro* assays or in transgenic plants (Figueras et al. 2004; Imai et al. 1996; Kasuga et al. 1999; Kazuoka and Oeda 1994; Riera et al. 2004; Swire-Clark and Marcotte 1999; Xu et al. 1996). However, stress tolerance is a complex process in which other protective components, such as sugar, detoxifying molecules or heat-shock proteins, participate. In seeds, the simultaneous presence of LEA proteins and non-reducing disaccharides seems to be especially important for dehydration tolerance (Bartels and Sunkar 2005). In addition, not only the accumulation but also the post-translational modifications, such as phosphorylation, of stress proteins may be also relevant for their protective function (Alsheikh et al. 2003; Heyen et al. 2002; Riera et al. 2004; Röhrig et al. 2006).

The presence of LEA proteins in the seed and their abundance makes this plant structure especially useful as a source to obtain large numbers of LEA proteins for analytical studies on a proteomic scale. LEA proteins are generally non-globular, low-complexity proteins. They are natively unstructured but may become structured under stress conditions. They have been classified in different groups on the basis of amino acid sequence similarities and expression pattern (Tunnacliffe and Wise 2007). LEA proteins are usually rich in hydrophilic amino acids with a high proportion of charged and polar amino acids and they fail to coagulate upon boiling (Cuming 1999; Wise 2003). We have previously used this thermostability property in the initial purification step prior to the isolation of phosphorylated LEA proteins in Arabidopsis seeds (Irar

et al. 2006). However, one third of the Arabidopsis seed dry weight consists of storage proteins, comprising the 12S globulins and 2S albumins, their precursors and mature forms. These proteins are also salt-soluble and thus represent a serious source of contamination when trying to isolate LEA-type proteins from plant seeds (Irar et al. 2006; Mills et al. 2002).

Mass spectrometry (MS) has been used in protein identification since the beginning of the 90^{ths} (Mann et al. 1993; Yates et al. 1993). Proteins are identified by comparison of the MS information against protein or nucleotide databases. Two kinds of information can be obtained by MS: the protein peptide map (PMF, peptide mass fingerprinting) or the structural information obtained via fragmentation of the peptides (MSMS). The PMF is the simplest approach to identify one single protein, but it fails to analyze samples containing mixture of proteins. In such cases, the most employed technique is a combination of peptide separation via high pressure liquid chromatography (HPLC) coupled on/off line to the mass spectrometer, which should be able to perform MSMS analysis of a large number of peptides. The confidence of the identification of a protein will be higher as the number of fragmented peptides matched in the database search increases. The score associated with the identification is calculated based on the quality and number of peptides fragmented, among other parameters such as the size of the database.

Achievement of good peptide fragmentation is usually related to the peptide ionization. The intensity of the peptide signal is associated with its ability to ionize. And it depends fundamentally on the amino acid sequence, but the concentration or abundance of the molecule plays also an important role. The ionization is not a very well understood phenomenon, but the suppression suffered by some ions in presence of others has been observed. In fact, one of the major problems in the identification of proteins in complex samples via MS is the capability of the mass spectrometer to resolve and fragment a big number of peptides. In order to reduce the sample complexity, different separation steps are employed. In plant proteomics, diverse enrichment and separation steps are combined prior to the mass spectrometry analysis(Rossignol et al. 2006). Because of plant proteome complexity and the lack of new technology in the high-throughput plant proteomics field, the most used protocol to separate proteins is still the 2-DE (two-dimensional electrophoresis) followed by mass spectrometry analysis of digested selected spots. However, other combinations of separation and mass analysis like MudPIT, (multidimensional protein identification technology) (Link et al. 1999) and BN-SDS-PAGE (bidimensional blue-native) (Heinemeyer et al. 2004) have been employed.

Unstructured proteins have been isolated in different systems, for proteomic identification using either their solubility during acid treatment (Cortese et al. 2005) or their thermostability upon heating (Csizmók et al. 2006; Galea et al. 2006). These studies aimed to an efficient enrichment of unstructured proteins by depletion of proteins containing folded globular domains. Since LEA are thermostable and natively unfolded proteins, we have investigated whether the acid solubility treatment could be used as an enriching step in this kind of proteins. In the present study, the total salt soluble protein extracts from Arabidopsis seeds were heat-treated and the resulting thermostable protein fraction was subsequently subjected to acid treatment. With these two consecutive fractionation steps it is possible to isolate a protein fraction devoid of major storage protein contaminants and highly enriched in LEA-type proteins. With this procedure, the complexity of the resulting acid soluble (S) and acid insoluble (I) fractions was lower than the initial thermostable seed protein extract. To identify the protein content we followed two approaches suitable for protein identification in less complex protein mixtures or subproteomes. First, a classical 1D (SDS-PAGE) gel-based procedure associated with MS analysis using an electrospray ionization source coupled on-line to liquid chromatography (LC-ESI-MSMS). And second, a gel-free protocol associated with an off-line HPLC and analysis via matrix assisted laser desorption/ionization (LC-MALDI-MSMS)(Bodnar et al. 2003).

2. Materials and methods

2.1 Plant material and protein extraction: heat and acid treatments

Seeds of *Arabidopsis thaliana* RLD ecotype were used in all experiments. 150 mg of dry seeds were embedded in sterile water for 2h at 4ºC and ground with the aid of an eppendorf-shaped glass pestle and granular quartz. Proteins were solubilized in salt buffer (20 mM TES-KOH, pH 8.0, 0.5 M NaCl) at 4ºC containing standard protease inhibitors (1 mM PMSF, 50 uM Leupeptin and 10 mM E-64). Protein extracts were clarified twice by centrifugation at 10.000 g for 10 min at 4ºC. The obtained supernatant containing the salt soluble extracts were heated at 100ºC for 10 min, cooled on ice for 30 min and centrifuged at 10.000 g for 10 min at 4 ºC. The total heat-soluble proteins were then subjected to TCA treatment by addition of 40% TCA to give a 3% TCA final concentration. After 15 min on ice, the insoluble 3% TCA fraction was removed by centrifugation at 10.000 g for 15 min. The acid soluble proteins in the corresponding supernatant were concentrated by precipitation with 15% TCA. The pellets were washed twice with cold methanol containing 0.1 M ammonium acetate and air-dried.

2.2 Gel electrophoresis

Extracts from each fraction were solubilized in SDS sample buffer and proteins were analyzed by one-dimensional electrophoresis (SDS-PAGE) on 15% polyacrylamide gels. The gels were stained with Coomassie Brillant Blue R (Sigma) or by silver staining (Shevchenko et al. 1996). The Coomassie stained protein bands from the soluble and insoluble fractions were carefully excised prior to MS analysis.

2.3 Immunoblot analysis

Protein separated on one-dimensional gels were transferred to nitrocellulose membranes (Towbin et al. 1979). A 1:2000 dilution of a rabbit polyclonal antiserum, specific for the maize LEA protein Rab17 was used to identify the maize protein in transgenic Arabidopsis seeds. The antiserum also recognizes the homologous LEA protein in *Arabidopsis thaliana*, AtRab18. Western detection was performed by the ECL method with a secondary anti-rabbit antibody conjugated to peroxidase (ECL Detection System, Amersham Pharmacia Biotech).

2.4 In-Gel Digestion

Proteins were in-gel digested with trypsin (Sequencing grade modified, Promega) in the automatic Investigator ProGest robot of Genomic Solutions. Briefly, excised gel bands were washed sequentially with ammonium bicarbonate (NH_4HCO_3) buffer and acetonitrile (ACN). Proteins were reduced and alkylated respectively, by treatment with 10 mM dithiothreitol (DTT) solution during 30 minutes at 56 ºC, and treatment with a 55 mM solution of iodine acetamide (IAA). After sequential washings with buffer and acetronitrile, proteins were digested overnight, at 37 ºC with 0.27 nmol of trypsin. Tryptic peptides were extracted from the gel matrix with 10% formic acid (FA) and acetonitrile; the extracts were pooled and dried in a vacuum centrifuge.

2.5 Acquisition of MS/MS Spectra by LC-ESI-MSMS.

Proteins excised from the gel were analyzed by ESI-MS-MS (Q-TOF Global, Micromass-Waters). Tryptic digested peptides were analyzed by online liquid chromatography tandem mass spectrometry (CapLC-nano-ESI-Q-TOF) (CapLC, Micromass-Waters). Samples were resuspended in 15 μL of 1% formic acid solution and 4 μL were inject to chromatographic separation in reverse-phase capillary C_{18} column (75 μm of internal diameter and 15 cm length, PepMap column, LC Packings). Peptides were eluted at a gradient of 5 min at 5 %B; 5 to 50 %B in 30 min; 50 to 95 %B in 5 min (A: 2% ACN/ 0.1% FA and B: 90% ACN/ 0.1% FA with a 200

nL flow rate. Ionization was performed in coated nano-ES needles (PicoTip™, New Objective). A capillary voltage of 1800-2200 V was applied together with a cone voltage of 80 V. The collision in the CID (collision-induced dissociation) was 20-35 eV and argon was employed as collision gas.

2.6 In-Solution Digestion

Proteins were in-solution digested with trypsin (Sequencing grade modified, Promega). Samples were resuspended in 100 µL of NH_4HCO_3 buffer 50 mM. Proteins were reduced and alkylated respectively, by adding 50 µL of 10 mM DTT solution and reacting at 56 ºC during 45 minutes, followed by addition of 50 µL of 55 mM solution of IAA for 30 minutes at rt (room temperature). The samples were digested with 1µg of trypsin overnight at 37ºC and dried in a speed-vac.

2.7 Acquisition of MS and MSMS Spectra by LC-MALDI-MSMS.

Peptides were resuspended in 100 µl of 1% formic acid and separated in a CapLC-Waters instrument, where 10 µl of sample were injected. The gradient used was: 5 min at 5 %B; 5 to 70 %B in 95 min; 70 to 95 %B in 10 min (A: 2% ACN/ 0.1% FA and B: 90% ACN/ 0.1% FA). The flow rate was 200 nL/min. The matrix CHCA (α-cyano-4-Hydroxy-Cinnamic Acid, Sigma) at 2 mg/mL in 80% ACN was dispensed from a syringe at 970 nL/min. Fractions were recovered at each 25 seconds on a MALDI plate using a Probot robot. Spectra were acquired in a MALDI-TOF/TOF 4700 Proteomics Analyzer (Applied Biosystems). The LC-MALDI acquisition method was used to acquire MS and MSMS spectra. For each spot a maximum of 15 most intense peaks with a signal/noise higher than 35 were fragmented.

2.8 Data Analysis

Data generated in ABI 4700 were analyzed by GPS workstation software (v 2.1) with MASCOT resident against NCBI Database. The searches were also performed manually with online MASCOT. PKL files were generated with data from ESI-QTOF analyses. Files were used to search online with MASCOT against NCBI. The following parameters were used to run searches: Database: NCBI, Taxonomy: Arabidopsis Thaliana, Enzyme: Trypsin 1 missed cleavage, Fixed modification: Carbamidomethylation of Cys (variable modification), Oxidation of Met, Peptide mass tolerance: 200 ppm and Fragment tolerance: 0.25 Da.

3. Results

3.1 Comparison of the electrophoretic protein patterns of acid-insoluble and acid-soluble extracts from Arabidopsis seed thermostable proteins

Proteins from Arabidopsis seeds were extracted by heat treatment and by 3% trichloroacetic acid (TCA) procedure. The SDS-PAGE electrophoretic profiles of the different protein fractions obtained are shown in Figure 1. The predominant protein bands which run in the molecular range of 20 to 35 kD correspond to major *Arabidopsis* storage proteins. With the heat treatment, part of the storage proteins become insoluble, although a substantial amount is still found in the heat-stable supernatant and cannot be removed by longer centrifugation times. However, the bulk of storage proteins seem to become insoluble by treatment with 3% TCA (Figure 1A, lanes 3, 4). The combination of both treatments, heat followed by TCA precipitation, produced a silver-stained protein pattern with clear differences: an acid-insoluble fraction (I) containing the classical storage protein pattern and an acid-soluble fraction (S) with specific different bands (Figure 1B).

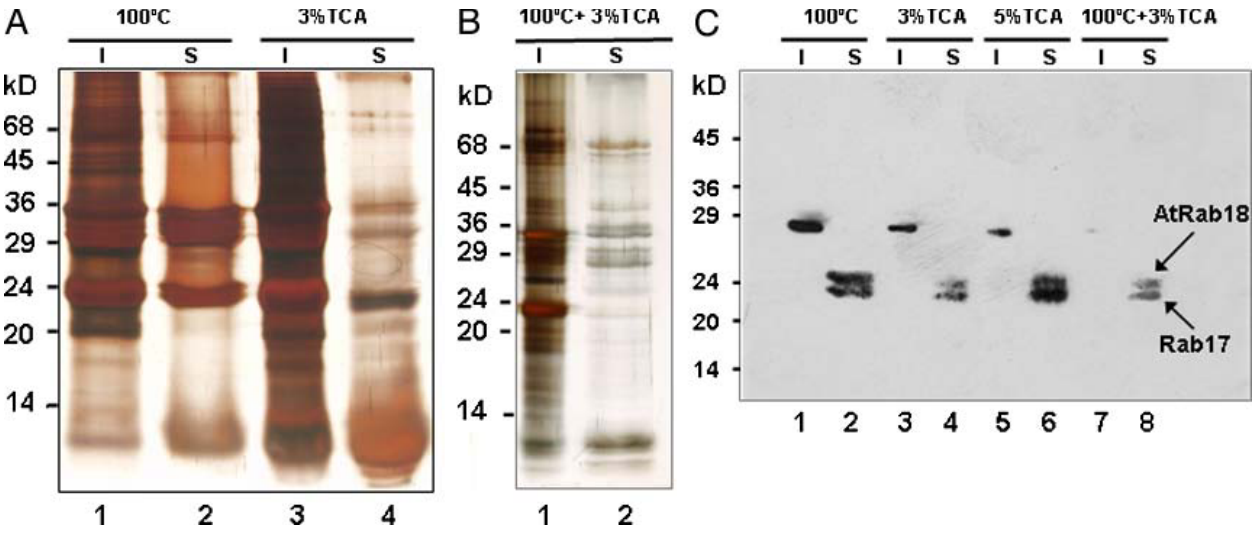


Figure 1: SDS-PAGE Electrophoresis of Arabidopsis seed proteins: heat and acid treatments.

(A) Silver stained protein extracts from heat or TCA treated samples: heat insoluble (lane1) and heat soluble (lane2) extracts; 3%TCA insoluble (lane 3) and 3% TCA soluble (lane 4). (B) Thermostable protein extracts treated with 3% TCA: acid-insoluble (lane1) and acid-soluble (lane 2) proteins. (C) Western blotting indicating the presence of dehydrin from maize Rab17 and AtRab18 from Arabidopsis. The arrows indicate the position of both dehydrins which are soluble by heating and by acid treatment. The antiserum identifies another protein band of approximately 30 kD which is heat-insoluble (Irar et al. 2006).

Since LEA-type proteins are in general thermostable proteins and natively unfolded (Tunnacliffe and Wise 2007) we reasoned that using this fractionation scheme they should be found preferentially in the thermostable-acid soluble fraction. In fact, in the isolation of the maize dehydrin DHN1 (Rab17), an acid-enriching step with 2% TCA was used prior to HPLC separation(Jiang and Wang 2004). In our system, we confirmed the degree of acid solubility of maize Rab17 analyzing by immunobloting *Arabidopsis* transgenic seeds over-expressing the maize dehydrin Rab17 (Figueras et al. 2004). Using an antiserum specific for maize Rab17 which also recognizes AtRab18, the homologous protein in Arabidopsis, we observed that both dehydrins were thermostable and soluble at 3% and 5% TCA (Figure 1C). They only become insoluble at 7% TCA (result not shown).

In initial experiments, we used 8 volumes of cold acetone as precipitating agent for the thermostable/acid soluble protein extracts. We found that the efficiency of the recovery of total acid-soluble proteins was lower in comparison to that obtained using 15% TCA as precipitating agent. Although the unusual amino acid composition of unstructured proteins seems to introduce errors in the traditional assays for determining protein quantity (Szőllősi et al. 2007), we measured recovered protein by Bradford assay and observed that roughly 50% more protein was recovered by 15% TCA precipitation. Differences in efficiency of protein precipitation using TCA or PCA as precipitating agents have also been reported in the isolation of unstructured proteins from *E.coli* (Cortese et al. 2005). In our system, TCA concentrations above 15% did not improve the protein yield (data not shown) and thus, 15% TCA was established as standard precipitating method in all experiments.

The Coomassie stained pattern of thermostable proteins after fractionation into acid-soluble and acid-insoluble fractions is shown in Figure 2. The differences in the protein bands prompted us to the analysis by MS, first on a gel-based protocol and then in a global gel-free MS analysis. The presence and distribution of LEA proteins, storage reservoir and proteins from other families among the two fractions was the prior goal in order to validate the fractionation procedure as a suitable LEA enriching system.

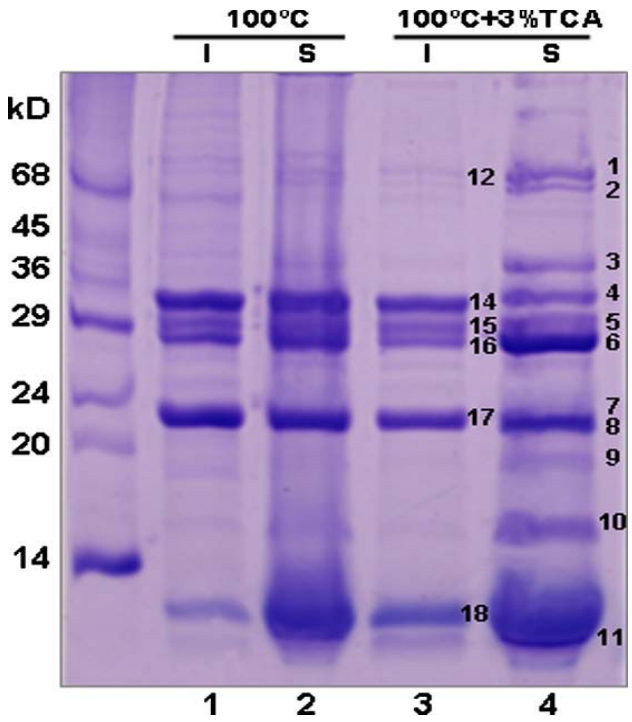


Figure 2: Coomassie staining of the fractionated proteins from Arabidopsis seeds after heat and 3% TCA treatment.

Total heat insoluble (lane 1) and heat soluble (lane 2) proteins. The heat soluble proteins from lane 2 are further fractionated into acid-insoluble (lane 3) and acid-soluble (lane 4). The arrows and numbers indicate the position of stained bands that were excised, trypsin digested and analyzed by MS. The band numbers correspond to the bands in Table 1. The molecular weight markers are shown on the left in kD. The amount of protein loaded in each lane corresponds to 2 mg (lanes 1 and 2) and 10 mg (lanes 3 and 4) of initial seed weight.

3.2 Analysis of SDS gel bands by LC-ESI-MSMS

In a first analytical approach proteins were separated in a SDS-PAGE gel, the stained bands were carefully excised from the gel as indicated by numbers in Figure 2 and protein composition was analyzed by LC-ESI-MSMS (Table 1). Protein identification shows a clear enrichment of LEA-type proteins in the acid-soluble fraction using this protocol since the major bands correspond mostly to LEA proteins.

In the acid-insoluble fraction, the central bands observed in the range of 20-35 kD (bands 14, 15, 16 and 17 in Figure 1, lane 3) correspond to 12S group of storage proteins while the lower molecular weight band of 12 kD (band 18 in Figure 2, lane 3) is composed essentially by 2S reservoir proteins. Except for two LEA proteins (LEA 76 homolog and an unknown LEA in bands 12 and 16, respectively) and a lipid transfer protein (band 18), no other polypeptides were simultaneously present in both fractions (Table 1).

Arabidopsis seeds contain two predominant types of seed storage proteins, 12S globulins and 2S albumins (Higashi et al. 2006). Four genes encode the 12S globulins which are synthesized as precursors and then cleaved post-translationally to α and β-subunits. They form the abundant spots in the 20-35 kD range in 2D gels, many spots being derived from the same subunit (Gallardo et al. 2002). The 2S albumins are encoded by five genes, four of them located in tandem in the same genomic region which encode for water soluble highly homologous proteins (Gallardo et al. 2002; Guerche et al. 1990; Higashi et al. 2006) running at the lower 10-12 kD region of the gels. Our results indicate that the major bands observed in the acid-insoluble extracts correspond basically to storage proteins.

Table 1: Proteins identified using gel-based protocol combined with LC-ESI-MSMS*.

Gel band	Protein Name	Group	Locus	N° Accession	Theoretical MW/pI	Mascot score	N° Peptides	% Coverage
1	unknown LEA	LEA 3	AT3G53040	gi|15231736	52.0/5.29	483	8	19
2	ATECP63	LEA 3	AT2G36640	gi|1526424	48.5/5.43	703	12	30
2	Rd29B/Lti65	-	AT5G52300	gi|22327778	65.9/5.07	127	3	5
3	LEA domain-containing	LEA 3	At5G44310	gi|15241469	37.1/8.62	160	4	10
3	LEA-like (similiar to DC-8)	LEA 3	AT4G36600	gi|30690704	36.41/6.99	40	1	3
4	LEA 32	LEA 4	AT3G17520	gi|4140257	31.4/5.21	254	4	18
4	ATEC31	LEA 5	AT3G22500	gi|15228770	26.8/5.34	106	2	14
5	LEA domain-containing	LEA 4	AT4G2102	gi|15233479	29.4/7.62	192	4	16
6	LEA76 homologue type2	LEA 3	AT3G15670	gi|1592677	24.1/8.62	404	7	34
6	nutrient reservoir	-	AT2G28490	gi|15226403	55.8/5.83	228	4	9
7/8	ATM1/GEA1	LEA 1	AT3G51810	gi|15231041	16.6/5.76	149	5	51
8	60S ribosomal protein L23a	-	ATrpL23A	gi|585876	17.46/10.20	92	1	8
8	LEA unknown (sim to LEA76)	LEA 3	AT1G52690	gi|15219071	18.09/6.99	58	1	7
9	LEA like	LEA 4	AT5G06760	gi|15240144	16.16/9.43	160	4	37
10	ATM6	LEA 1	AT2G40170	gi|15225645	9.9/6.75	187	4	69
11	nutrient reservoir	2S p3	AT4G27160	gi|15237012	19.3/8.03	205	4	25
11	nutrient reservoir	2S p2	AT4G27150	gi|15236995	19.9/6.9	189	4	24
11	nutrient reservoir	2S alb4	AT4G27170	gi|15237014	19.9/7.46	170	4	24
11	trypsin inhibitor putative	defense	AT1G47540	gi|15220249	11.4/6.23	94	2	25
11	nutrient reservoir	2S	AT5G54740	gi|15239728	19.3/7.53	61	2	10
11	LPT2	storage	AT2G38530	gi|15224898	12.5/9.32	38	1	8
12	unknown LEA	LEA 3	AT3G53040	gi|15231736	52.0/5.29	190	3	7
14	CRU3 (CRUCIFERIN 3)	12S	AT4G28520	gi|15235321	58.5/6.53	310	7	19
14	storage protein CRA1	12S	AT5G44120	gi|166676	52.9/7.68	166	4	11
15	storage protein CRA1	12S	AT5G44120	gi|166676	52.9/7.68	416	7	19
15	CRU3 (CRUCIFERIN 3)	12S	AT4G28520	gi|15235321	58.5/6.53	172	4	11
15	nutrient reservoir	cupin	AT3G22640	gi|18403467	55.2/6.64	121	6	12
16	CRU2 (CRUCIFERIN 2)	12S	AT1G03880	gi|15219582	50.86/6.52	284	5	15
16	12S storage protein CRA1	12S	AT5G44120	gi|166676	52.9/7.68	187	4	10
16	nutrient reservoir	cupin	AT3G22640	gi|18403467	50.86/6.52	126	4	10
16	LEA76 homologue type2	LEA 3	AT3G15670	gi|1592677	24.1/8.62	123	3	13
16	unknown protein	-	AT4G17440	gi|30684034	23.16/9.65	34	1	5
17	CRU3 (CRUCIFERIN 3)	12S	AT4G28520	gi|15235321	58.5/6.53	264	6	18
17	12S storage protein CRA1	12S	AT5G44120	gi|166676	52.9/7.68	213	3	10
17	CRU2 (CRUCIFERIN 2)	12S	AT1G03880	gi|15219582	50.8/6.52	152	3	7
17	beta-chain 6 (fragment)	storage	-	gi|1076411	1.86/4.53	51	1	100
18	nutrient reservoir	2S	AT4G27150	gi|15236995	19.9/6.79	161	3	20
18	nutrient reservoir	2S	AT4G27170	gi|15237014	19.9/7.46	95	2	15
18	nutrient reservoir	2S	AT5G54740	gi|15239728	19.3/7.53	76	2	10
18	LTP2	storage	AT2G38530	gi|15224898	12.5/9.32	57	1	8

* The *Arabidopsis* gene locus, protein accession number according to the NCBI Gene Bank and theoretical molecular weight and isoelectric point are indicated.

On the contrary, the analysis of the acid-soluble bands shows that most of the identified proteins correspond to LEA-type proteins. Ten different Coomassie stained bands covering the whole length of the gel (roughly from 16 kD to 70 kD) were excised (Figure 2, lane 4) and thirteen different LEA proteins were identified (Table 1). Only one storage protein and one unrelated ribosomal protein were identified in the acid-soluble fraction. In contrast, the analysis of the lower 12 kD molecular weight band showed that, except for a putative trypsin inhibitor defense protein, the rest of the polypeptides corresponded to 2S storage members also present in the same size band in the acid-insoluble extracts (band 18 in Lane 3, Figure 2). The presence of 2S storage proteins in this band indicates that these highly homologous reservoir polypeptides, in contrast to the 12S group, may be at least partially soluble in acid conditions.

In summary, the results obtained in the gel-based approach indicate that it is feasible to isolate a protein subpopulation enriched in Lea-type proteins, in sufficient amounts to be visualized in Coomassie stained gels, and essentially free of 12S globulins. The presence 2S storage proteins is concentrated in a single band of approximately 12kD, well below the rest of the Lea-containing bands, and does not interfere with the MS identification of upper gel bands.

3.3 Analysis of soluble and insoluble acid fractions by LC-MALDI-MSMS.

In this approach, both soluble and insoluble fractions were in-solution digested with trypsin and analyzed by mass spectrometry. 10 μL of digested peptides were injected to the HPLC and fractions of 25 s were collected on a MALDI plate using a microfraction collector (Figure 3). Spots were shot using an Interpretation Mode which allowed fragmentation of the 15 molecules per spot with the most intense signal. In the above gel-based system we analyzed the more abundant proteins visible by Coomasie staining.

Table 2: Proteins identified by the gel-free approach combined with LC-MALDI-MSMS*.

Protein name	Group	Locus	N° Accession	Theoretical MW/pI	Mascot score	N° Peptides	% Coverage
			Acid Soluble Fraction				
LEA-Type proteins							
Glycine-rich, unknown	-	AT2G05580	gi|4581166	26.1/9.9	215	5	38
LEA76 homologue type2	LEA 3	AT3G15670	gi|1592677	24.1/8.62	78	2	16
Rab18, glycine rich	LEA 2	AT5G66400	gi|15239373	18.5/7.1	95	1	11
unknown	LEA 4	AT2G35300	gi|15226951	10.4/9.6	33	2	30
AtM6	LEA 1	AT2G40170	gi|15225645	9.9/6.7	351	5	75
LEA D113 homologue type1	LEA 4	AT5G06760	gi|1592679	16.64/9.43	143	1	15
LEA unknown (LEA 76 homolgy)	LEA 3	AT1G52690	gi|15219071	18.09/6.99	118	1	10
LEA 32-like	LEA 4	AT3G17520	gi|4140257	31.4/5.21	54	1	8
LEA domain containing	LEA 3	AT2g42560	gi|15227965	67.1/5.78	37	1	2
Seed storage proteins							
nutrient reservoir	2S p3	AT4G27160	gi|15237012	19.3/8.03	212	4	17
nutrient reservoir	2S alb4	AT4G27170	gi|15237014	19.9/7.46	194	4	24
nutrient reservoir	storage	AT5G54740	gi|15239728	19.3/7.53	270	4	30
nutrient reservoir	2S	AT4G27140	gi|15236992	19.5/5.6	137	5	20
nutrient reservoir	2S p2	AT4G27150	gi|15236995	19.9/6.9	237	4	17
Other proteins							
trypsin inhibitor putative	defense	AT1G47540	gi|15220249	11.4/6.23	117	2	17
			Acid Insoluble Fraction				
LEA-Type proteins							
LEA domain containing	LEA 3	AT2G42560	gi|15227965	67.1/5.78	260	6	18
ATM1/GEA1	LEA 1	At3G51810	gi|15231041	16.6/5.76	312	3	31
unknown LEA	LEA 3	AT3G53040	gi|15231736	52.0/5.29	36	2	7
ATEC31	LEA 5	AT3G22500	gi|15228770	26.8/5.34	73	2	10
Probable dehydrin	LEA 2	AT2G21490	gi|15226568	19.2/6.3	45	1	12
LEA76 homologue type2	LEA 3	AT3G15670	gi|1592677	24.1/8.62	32	2	12
LEA 32-like	LEA 4	AT3G17520	gi|4140257	31.4/5.21	143	1	8
ATECP63	LEA 3	AT2G36640	gi|1526424	48.5/5.43	112	3	7
ATM6/GEA6	LEA 1	AT2G40170	gi|15225645	9.9/6.754	74	2	43
Seed storage proteins							
CRU3 (CRUCIFERIN 3)	12S	AT4G28520	gi|15235321	58.5/6.53	1400	20	52
CRA1	12S	AT5G44120	gi|166676	52.9/7.68	1030	16	45
CRU2 (CRUCIFERIN 2)	12S	AT1G03880	gi|15219582	50.86/6.52	1048	15	43
nutrient reservoir	2S p2	AT4G27150	gi|15236995	19.9/6.79	373	4	22
nutrient reservoir	2S alb4	AT4G27170	gi|15237014	19.9/7.46	171	4	24
nutrient reservoir	2S	AT4G27140	gi|15236992	19.9/5.6	176	3	18
unknown	lectin	AT3G21380	gi|34222076	49.2/6.0	121	4	12
nutrient reservoir	cupin	AT3G22640	gi|18403467	50.86/6.52	91	4	13
nutrient reservoir	2S	AT5G54740	gi|15239728	19.3/7.53	86	1	10
beta-chain 6 (fragment)	-	-	gi|1076411	1.86/4.53	75	1	100
beta-chain (fragment)	-	-	gi|1076408	2.12/4.17	61	1	52
Other proteins							
trypsin inhibitor putative	defense	AT1G47540	gi|15220249	11.4/6.23	290	3	18
unknown	-	AT5G15580	gi|15242342	104.6/9.13	50	1	1
ubiquitin homology	ribosomal	AT1G23410	gi|15220742	17.88/9.77	49	1	5
ROC1	-	AT4G38740	gi|15234781	18.5/7.6	29	1	6
PRLI-interacting factor	-	AF3157371	gi|11139270	23.4/5.6	32	2	9

* The identified proteins in the acid soluble and insoluble fractions are numbered and grouped as LEA-type, storage and other proteins. The *Arabidopsis* gene locus, protein accession number according to the NCBI Gene Bank and theoretical molecular weight and isoelectric point are indicated.

Now the objective was to compare the global protein content and to identify minor proteins below the detection limits of Coomassie staining and those proteins not identified using LC-ESI procedure. Again, most of the proteins identified using this system belong to the LEA and storage protein families (Table 2).

With LC-MALDI a total of fourteen LEA proteins were identified. Nine of them were also identified in the gel-based approach (Table 3). Nevertheless, five were exclusively identified by the MALDI system. And of these, four were TCA-soluble and only one was found in the TCA-insoluble fraction. This is the case of AtRab18 dehydrin, a LEA group 2 protein, highly soluble in acid conditions as assessed by Western blotting (Figure 1 C). The same occurs with other LEA-proteins (number 6, 8 and 13) as showed in Table 3. Interestingly, all these proteins are highly hydrophilic producing extremely hydrophilic tryptic peptides that could be lost in the chromatographic separation when running LC-ESI-MSMS (Figure3).

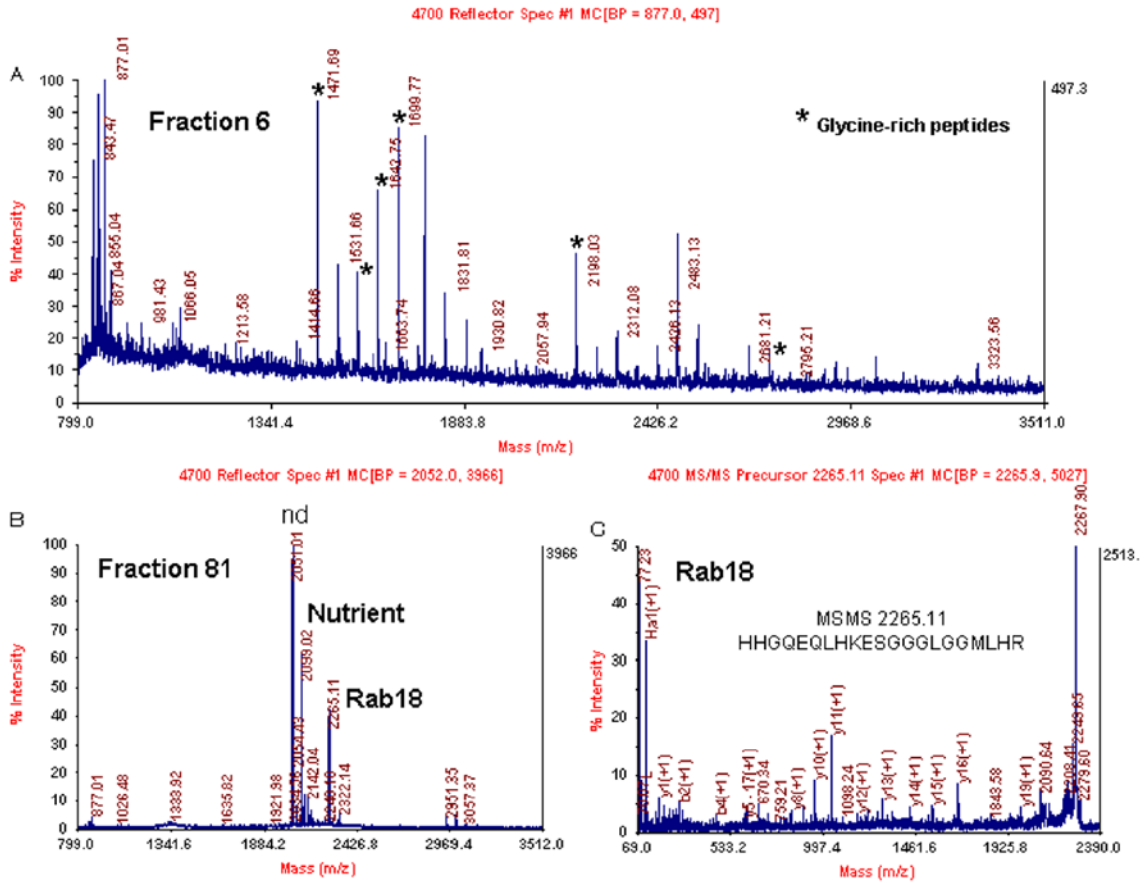


Figure 3: MS and MSMS spectra of fractions 6 and 81 obtained via LC-MALDI analysis of the acid-soluble fraction.

(A) MS spectrum of peptide containing in fraction 6. Most peptides correspond to glycine-rich protein (number 6, Table 3). (B) MS spectrum of fraction 81 and (C) MSMS spectrum of the parent ion $[M+H]^+$ = 2265.11 Da of Rab18, protein number 7 (Table 3). Peptides were separated by reverse phase chromatography using the following conditions: linear gradient of 5 min at 5 %B; 5 to 70 %B in 95 min; 70 to 95 %B in 10 min (A: 2% ACN/ 0.1% FA and B: 90% ACN/ 0.1% FA). The flow rate was 200 nL/min and fractions were recovered at each 25 seconds on a MALDI plate**.**

The seed storage protein family was again found to be mostly insoluble by TCA treatment (Table 2). The acid insoluble fraction contains the major storage proteins of higher theoretical molecular weight, corresponding to the 12S globulin family, the reservoir 2S albumin family plus some minor nutrient protein fragments. In contrast, the acid soluble fraction is devoid of the high molecular weight storage proteins and only contains the high homology group of 2S reservoir proteins. This result is in accordance to those presented above indicating that the 12S storage proteins become insoluble by 3%TCA treatment (Table 1). Only two storage proteins were exclusively identified by LC-MALDI system (22 and 24, Table 3). They were not detected by the LC-ESI system because protein 22 was probably not excised from the gel, since it was expected to run between bands 12 and 14 (Figure 2) and polypeptide 24 corresponds to a beta chain fragment of reduced size which runs with the gel front.

Table 3: List of the total identified proteins in each soluble and insoluble fractions*.

N° Protein		N° Accession	LC-MALDI[1]		Gel-LCMS[2]		Acid/Basic/ Non-charged amino acids[3]	IUPred[4]
	LEA-Type proteins							
1	LEA domain containing	gi|15227965	I	S	-	-	18/18/46	glob
2	LEA76 homologue type2	gi|1592677	I	S	6	16	13/18/59	A
3	LEA 32-like	gi|4140257	I	S	4	-	22/20/40	A
4	ATM6/GEA6	gi|15225645	I	S	9-10	-	17/20/47	A
5	LEA unknown (LEA 76 hom.)	gi|15219071	-	S	7-8	-	15/18/57	A
6	Glycine-rich, unknown	gi|4581166	-	S	-	-	5/12/73	A
7	Rab18, glycine-rich	gi|15239373	-	S	-	-	9/13/60	A
8	unknown	gi|15226951	-	S	-	-	11/25/46	A
9	LEA D113 homologue type 1	gi|1592679	-	S	9-10	-	9/16/47	A
10	ATM1/GEA1	gi|15231041	I	-	7-8	-	20/22/41	A
11	unknown LEA	gi|15231736	I	-	1	12	23/21/41	A
12	ATEC31	gi|15228770	I	-	4	-	14/13/45	glob
13	Probable dehydrin	gi|15226568	I	-	-	-	13/18/49	A
14	ATECP63	gi|1526424	I	-	2	-	21/20/40	A
15	Rd29B/Lti65	gi|22327778	-	-	2	-	18/16/40	A
16	LEA domain-containing	gi|15241469	-	-	3	-	22/25/36	A
17	LEA-like (similiar to DC-8)	gi|30690704	-	-	3	-	17/19/43	A
18	LEA domain-containing	gi|15233479	-	-	5	-	21/22/40	A
	Seed storage proteins							
19	CRU3 (CRUCIFERIN 3)	gi|15235321	I	-	-	14-15	9/10/46	glob
20	CRA1	gi|166676	I	-	-	14-15	10/12/43	glob
21	CRU2 (CRUCIFERIN 2)	gi|15219582	I	-	-	16-17	9/11/42	glob
22	unknown	gi|34222076	I	-	-	-	14/15/39	glob
23	beta-chain 6 (fragment)	gi|1076411	I	-	-	17	20/10/40	glob
24	beta-chain (fragment)	gi|1076408	I	-	-	-	26/11/32	glob
25	nutrient reservoir	gi|18403467	I	-	-	15-16	11/14/37	glob
26	nutrient reservoir	gi|15236995	I	S	11	18	11/12/36	glob
27	nutrient reservoir	gi|15237014	I	S	11	18	11/13/33	glob
28	nutrient reservoir	gi|15236992	I	S	-	-	13/13/34	glob
29	nutrient reservoir	gi|15239728	I	S	11	18	12/14/35	glob
30	nutrient reservoir	gi|15237012	-	S	11	-	9/12/37	glob
31	nutrient reservoir	gi|15226403	-	-	6	-	11/11/40	glob
32	LTP2	gi|15224898	-	-	11	18	1/9/51	glob
	Other proteins							
33	trypsin inhibitor putative	gi|15220249	I	S	11	-	12/12/33	glob
34	PRLI-interacting factor	gi|11139270	I	-	-	-	17/18/33	glob
35	unknown	gi|15242342	I	-	-	-	13/16/38	glob
36	ubiquitin homology	gi|15220742	I	-	-	-	11/24/31	glob
37	ROC1	gi|15234781	I	-	-	-	7/10/51	glob
38	60S ribosomal protein L23a	gi|585876	-	-	7-8	-	10/26/31	glob

* The proteins are numbered and grouped as LEA-type, storage and other proteins and their NCBI accession number is indicated. 1, LC-MALDI system; I stands for acid-insoluble fraction and S for acid-soluble fraction; 2, Gel-LCMS system, the number of the excised band from the SDS-PAGE gels is indicated; 3, percentage of acidic (Glu, Asp) and basic (Arg, Lys, His) charged amino acids and polar non-charged amino acids (Ser, Thr, Gln, Asn, Gly, Ala); 4, Absence (A) or presence (glob) of protein globular domains according to the IUPred prediction site (Guerche et al. 1990)

In addition five other proteins, not belonging to the LEA or storage families, were identified as heat soluble but TCA insoluble proteins. One of them, a trypsin inhibitor, was shared by both fractions, and had been also detected in the lower molecular weight soluble band from the SDS-PAGE gels. Overall, the results obtained with the MALDI system (Table 2) reinforce the inclusion of the TCA-treatment step in the isolation of LEA-type proteins from *Arabidopsis* seed, since it reduces significantly the presence of storage proteins.

4. Discussion

4.1 General comments on the analytical MS methods

A complete list of all identified proteins is illustrated in Table 3, indicating the MS approach by which they were identified. From a total of 38 proteins, 26 were identified by in gel-based LC-ESI system and 31 were detected by the gel-free system coupled to LC-MALDI. Up to 19 were simultaneously identified by both mass approaches. Of the twelve proteins identified exclusively by LC-MALDI, five were LEA-type. The physicochemical characteristics of this group of proteins could explain the observed results. Different examples of MS and a MSMS spectra obtained by the LC-MALDI technique are shown in Figure 3. The first mass spectrum corresponds to the fraction containing peptides eluted at 4.5 %ACN in 2.5 min of retention time, in the gradient used (Figure 3A). Most of those peptides matched to a glycine-rich protein (number 6, Table 3) generating a PMF-like spectrum, although a reverse phase chromatographic separation of the tryptic peptides had been performed. These data show that the digested peptides from a glycine-rich protein elute very early from the HPLC column and can be easily lost in the chromatographic step. In the LC-MALDI system, fractions were acquired from time 0 to 105 minutes while LC-ESI was set to start data acquisition after 5 minutes of HPLC run. A number of LEA-type proteins, especially those classified as belonging to groups 1 and 2, have a high content in glycine residues (Wise 2003).

In other instances, few tryptic peptides within the suitable mass range were produced, such as occurred with proteins number 7, 8 and 13 in Table 3. Figure 3B shows the MS spectrum

of fraction 81 where Rab18 (number 7 in Table 3) was identified and Figure 3C illustrates the MSMS spectrum of the single parent ion matched to this protein. These specific physicochemical features together with the fact that, in the LC-ESI system, the excised bands of the gel did not cover the complete length of the gel could explain why these proteins were undetected in the gel-based system. Moreover, the ionization behavior of peptides is sequence and methodology dependent, meaning that the same peptide can ionize via MALDI but not via ESI and *vice versa*. In fact, the comparison of a peptide list matching to the same protein identified via MALDI-MS or ESI-MS usually includes one or more different ions (data not shown). Thus, protein identification with a high coverage can be better achieved by combining the results obtained using different ionization approaches and/or employing very sensitive mass spectrometers with high velocity of peptide fragmentation.

Our results indicate that the gel-free approach seems to perform better for our samples, considering that many LEA proteins are highly hydrophilic and produce digested peptides with abnormal amino acid sequences. Nevertheless, if the gel-based system analysis had been performed along the entire length of the SDS gel, other minor proteins below the sensitivity level of the Coomassie staining might have also been identified. The fact that four LEA proteins were identified only by gel-based system, reinforces the complementarity of the two systems.

4.2 Physicochemical properties of the proteins

In order to understand better the solubility properties, we analyzed the physicochemical characteristics and theoretical structural prediction of the identified proteins (Table 3 and Figure 4). Low average hydrophobicity is characteristic of thermostable proteins and, in fact, the group of proteins isolated by heat treatment from *Arabidopsis* seed extracts show an overall hydrophilic nature as illustrated in Figure 4A. The only exception corresponds to protein LTP2 (protein 31 in Table 3) with a grand average hydrophobicity (GRAVY) of 0.354, but positive +9 net charge which may be relevant for its thermal solubility. Interestingly, and comparing to the rest, the LEA group clearly exhibits a notably higher hydrophilicity. High hydrophilicity is characteristic of LEA-type proteins and of

unstructured proteins in general (Cortese et al. 2005; Csizmók et al. 2006; Tunnacliffe and Wise 2007).

LEA proteins contain higher percentage of charged residues in comparison to the storage group (Table 3 and Figure 4B). The increased presence of charged residues in unstructured proteins may account for their stability at low pH (Cortese et al. 2005). In fact, all LEA proteins listed in Table 3, except for one, were acid-soluble. However, some were also present in the acid-insoluble fraction, possibly reflecting different degrees of acid solubility in the LEA family.

The prediction of structural disorder for each individual protein using the bioinformatics program IUPred (Dosztányi et al.) is shown in Table 3. According to this analysis, 16 of the 18 LEA proteins identified are predicted to lack globular domains completely, and two have a partial globular prediction. In contrast, proteins from the storage group and other unrelated proteins show a globular structure prediction and very rarely partially globular, irrespectively of their acid solubility properties. Disordered proteins typically have a biased amino composition having higher levels of Pro, Glu, Lys, Ser, Gly and Gln amino acids and lower levels of Trp, Tyr, Phe, Cys, Ile, Leu and Val amino acids (Romero et al. 2001). The LEA proteins in this work are examples of biased amino acid composition with a lower relation of order/disorder amino acids (Figure 4C).

In a large scale analysis of the unstructured mammalian proteome a higher proportion of proteins possessed acidic pI values (Galea et al. 2006). The examination of our data indicates that the theoretical pIs in the LEA group ranges from 5 to 10 and in the storage group from 4 to 9 (Tables 1 and 2). By comparing the presence or absence of each protein in the acid soluble and acid insoluble fractions we could not establish a correlation between pI values and acid solubility for the LEA proteins identified. Moreover, theoretical pIs are altered by posttranslational modifications and at least eight of the total eighteen LEA proteins identified are known to be *in vivo* phosphorylated in the Arabidopsis seed (Irar et al. 2006).

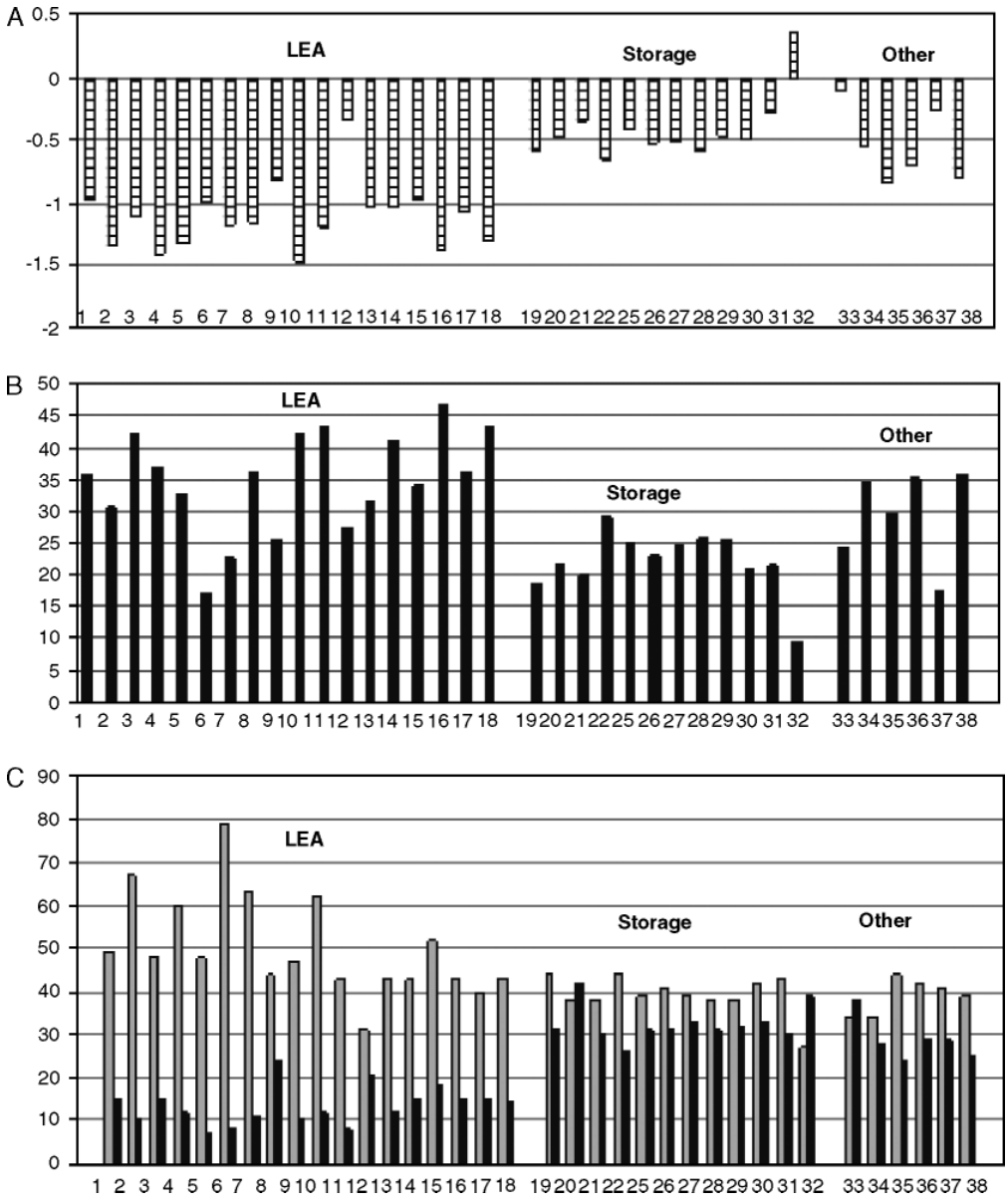


Figure 4: Biochemical properties of the proteins grouped as LEA-type, storage and other-identified proteins.

A. Hydrophobicity index: GRAVY (Grand Average Hydropathy) value for a protein is calculated as the sum of hydrophathy values of the amino acids, divided by the number of residues in the sequence. Less that 0 corresponds to a hydrophilic protein (http://www.expasy.org/tools/protparam.html). **B**. Percentage of charged amino acid residues (Glu/Asp/Arg/Lys/His). **C**. Deviations in amino acid composition. The relative percentage of disorder-promoting, Pro, Glu, Lys, Ser,Gly and Gln amino acids (grey bars) and order-promoting, Trp, Tyr, Phe, Cys, Ile, Leu and Val amino acids (black bars) are illustrated. The protein numbers below correspond to the proteins listed in Table 3.

5. CONCLUSION

In conclusion, in the present work we apply two LC-MSMS approaches suitable to identify proteins in moderately complex samples as those obtained from *Arabidopsis* seeds. We use a two step fractionation procedure (heat followed by acid treatment) which has proved to be an efficient method to isolate a protein fraction highly enriched in LEA-type proteins. The procedure relies on the fact that the majority of 12S globulins, otherwise ever present storage family proteins, are insoluble in 3% TCA whereas a great part of the LEA proteins identified remain soluble. More interestingly, the amount of LEA proteins in the acid soluble fraction is high enough to be visualized by Coomassie staining in SDS-PAGE gels. This opens the possibility for large scale analysis of LEA-type proteins, the search for their *in vivo* post-translational modifications and the comparison of total LEA protein content in stress resistant versus sensitive seed varieties. In addition, the system can be adapted to analyze and compare LEA protein production in vegetative tissues subjected to different environmental stress conditions.

CHAPTER 2

IV. Chapter 2

Chapter 2

Insights in maize LEA Proteins: from Proteomics to Functional Approaches

Amara, I., Odena, A., Oliveira, E., Moreno, A., Masmoudi, K., Pages, M. and Goday, A.

Plant Cell Physiol (2012) 53: 312-329

Abstract:

LEA (late embryogenesis abundant) proteins participate in plant stress tolerance responses but the mechanisms by which protection occurs are not fully understood. In the present work the unfolded proteins from maize dry embryos were analyzed by mass-spectrometry. Twenty embryo proteins were identified and among them thirteen corresponded to LEA-type proteins. We selected three major LEA proteins, Emb564, Rab17 and Mlg3 belonging to groups 1, 2 and 3 respectively and we undertook a comparative study in order to highly differences among them. The posttranslational modifications of native proteins were analyzed and the antiaggregation properties of recombinant Emb564, Rab17 and Mgl3 proteins were evaluated in vitro. In addition, the protective effect of the LEA proteins were assessed in living cells under stress in *E.coli* cells and in *N.bentamiana* leaves agroinfiltrated with fluorescent LEA-GFP fusions. Protein visualization by confocal microscopy indicated that cells expressing Mg3-GFP protein showed reduced cell shrinkage effects during dehydration and that Rab17-GFP protein co-localized to leaf oil bodies after heat shock. Overall, the results evidence differences and suggest functional diversity among maize LEA groups.

Keywords: abiotic stress, aggregation, Lea protein, maize, proteomic, oil body.

1. Introduction

Late embryogenesis abundant (LEA) proteins constitute a set of proteins that participate in plant stress responses. In general, LEA proteins accumulate naturally in desiccation tolerant structures such as seed or pollen grains and in plant vegetative tissues during exposure to abiotic challenges. Few show a constitutive expression. They are also found in anydrobiotic organisms, which can withstand complete loss of water without loss of viability, in resurrection plants, some invertebrates and microorganisms. The presence of LEA proteins correlates with the acquisition of desiccation tolerance, and hence a role in plant cell stress adaptative responses was early proposed (Bartels and Salamini 2001; Chakrabortee et al. 2010).

LEA proteins have been classified into several groups on the basis of specific domains and amino acid sequences, similarities in peptide composition or distinctive motifs (Battaglia et al. 2008; Bies-Etheve et al. 2008; Wise 2003); in general, groups 1, 2 and 3 are considered as the three major groups containing most members of the protein family. Group 1 LEA proteins (PFAM 00477) are mostly present in plants and they contain a 20 amino acid motif (GGQTRREQLGEEGYSQMGRK) at least in one copy. Group 2 LEA proteins or dehydrins (PFAM 00257) are also found in algae and share a common K-segment (EKKGIMDKIKEKLPG) present in one or several copies; many dehydrins also hold a S-segment (poly serine stretch) that can undergo extensive phosphorylation and a Y-domain (DEYGNP) similar to the nucleotide-binding site of plant and bacterial chaperones. Group 3 LEA proteins (PFAM 02987), also found in nematodes and prokaryotes, are characterized by holding at least one copy of the 11 amino acid TAQAAKEKAXE motif.

One feature of LEA proteins is their natively unfolded nature, hydrophyllicity and high proportion of charged amino acids. These biochemical properties contribute to their heat and acid stability and have been used to purify LEA enriched populations (Oliveira et al. 2007). Although LEA proteins are intrinsically disordered and show mostly random coil structure in solution, some members can adopt a degree of conformation during drying or in the presence of α-helix-promoting agents (Hundertmark and Hincha 2008; Liu et al. 2010). At the subcellular level, different LEA proteins have been found in most cell compartments including cytoplasm, nucleus, nucleolus, mitochondria, chloroplast, vacuole, endoplasmic

reticulum, perixosome and plasma membrane where they may exert their protective functions (Chakrabortee et al. 2007; Hundertmark and Hincha 2008). In relation to LEA protein posttranslational modifications (PMTs), most reports have focussed essentially on phosphorylation. In fact, phosphorylation is not uncommon in Arabidopsis seed LEA subproteome (Irar et al. 2006) and several LEA proteins from different species have been shown to be *in vivo* phosphorylated (Alsheikh et al. 2003; Goday et al. 1994; Heyen et al. 2002; Jiang and Wang 2004; Plana et al. 1991; Rohrig et al. 2006). Except for deamination and oxidation in LEAM, a mitochondrial pea LEA protein, so far other LEA protein modifications have not been reported (Tolleter et al. 2007).

There is extended information on the protective effects of LEA proteins in different *in vitro* assays and in *in vivo* approaches in which LEA proteins expressed in transgenic plants or heterologous expression in yeast and bacteria, produce tolerant phenotypes under drought, cold or high salinity stress. In general, LEA proteins are considered to contribute to the protection and stabilization of macromolecules and/or cellular structures during plant stress adaptative responses but their specific functions are not fully understood. At the molecular level, the protective effects of LEA proteins from groups 1, 2 and 3 have been proposed to occur through diverse mechanisms: binding or replacement of water, ion sequestration, stabilization and renaturation of unfolded proteins, membrane maintenance, alone or in combination with sugars. Because of redundancy and overlap among roles and groups and lack of systematic comparison of activities of different LEA groups, at present different roles for LEA proteins from different groups cannot be established. (Wise 2003).

In the current work we analyze the disordered subproteome of mature maize embryos, identify the LEA protein content and select three LEA proteins, Emb564, Rab17 and Mlg3, as representatives of groups 1, 2 and 3 for further analysis. Rab17 from group 2 has been previously characterized by our lab as a stress-induced protein, highly phosphorylated in the S-segment which is distributed in the cytoplasm and nucleus of seed scutellar cells; its expression in transgenic Arabidopsis plants increases the osmotic stress tolerance of the plants (Riera et al. 2004). To improve our understanding on LEA proteins, here we analyze comparatively different aspects of maize Emb564, Rab17 and Mlg3 proteins, including PTMs, subcellular localization and their properties in *in vitro* and *in vivo* assays. We

evidence differences among the three LEA proteins that might be relevant for their functionality.

2. Materials and methods

2.1 Heat and acid soluble proteins from maize embryos

Embryos from *Zea mays*, inbred line W64A, were manually dissected and the thermostable and acid soluble subproteome was obtained as described (Oliveira et al. 2007). Briefly, proteins soluble in salt buffer (20mM TES-KOH, pH 8.0, 0.5M NaCl) were heated at 100°C for 10 min, cooled on ice for 30 min and centrifuged at 15.000g for 10 min at 4 °C. Trichloroacetic acid (TCA) soluble proteins were obtained by treatment with increasing acid concentrations (3%, 7% and 10% TCA) for 15 min at 4ºC. TCA insoluble fractions were removed at 15.000g for 15 min and acid-soluble proteins were precipitated with 15% TCA. Pellets were resolved by 1-DE or 2-DE. Immunoblottings were developed using the ECL detection system (Amersham Pharmacia Biotech).

2.2 Expression of recombinant LEA proteins in *E.coli*

Mlg3 cDNA sequence was amplified by PCR using oligonucleotide primers (table of primer). Rab17 cDNAs have been described previously (Jensen et al. 1998). Sequences were cloned into pCR2.1-TOPO vector, cut out using the above enzymes and ligated into pET28a vector (Novagen, Madison, WI, U.S.A). Recombinant proteins containing His-tag were purified by HisTrapTM HP columns according to the manufacturer (GE Healthcare). Purified Emb564 and Mgl3 LEA proteins were used to raise polyclonal antibodies in rabbits. Rab17 antibodies had been obtained previously (Goday et al. 1994).

2.3 *In vitro* protein aggregation assay

Total salt soluble proteins (SP) from IPTG-induced *E. coli* cultures were obtained by sonication at 4ºC in salt buffer (20mM Tris-HCl, pH 7.9, 0.5M NaCl), centrifugation at 15.000g for 20 min and supernatant turbidity measured at A340 nm (Boucher et al. 2010; Chakrabortee et al. 2007). In desiccation assays, 500µl SP samples (0.15 mg) were vacuum dried for two hours and rehydrated in equal water volume. In freezing assays, samples were

frozen at -20ºC for two hours and in heating assays samples were heated at 48ºC for 45 min. Aggregation was measured at A340. In desiccation assays, soluble and aggregated proteomes were further separated at 15.000g for 15 min and analyzed in SDS acrylamide gels. Statistical relevance was determined by GraphPad Software,San Diego, CA.

2.4 *In vivo* assay of *E. coli* stress tolerance

E. coli cells expressing recombinant Emb564, Rab17, Mlg3 and empty pET28a vector were grown to OD_{600}=0.6 prior to IPTG addition. Cells were further grown for 5 hours at 37 ºC to OD_{600}=0.8. For osmotic stress treatment, cells diluted to a OD_{600}=0.6 were grown in 10% PEG containing medium and samples harvested every two hours for growth measurements (Gupta et al. 2010). For temperature stress assays, IPTG-induced cultures (OD_{600}=0.8) were exposed to cold stress (-20ºC for 2h) or heat stress (48ºC for 45min) (Dalal et al. 2009). After dilution to OD_{600}=0.6, 10 µl were spotted on agar plates and incubated at 37ºC for 16h. Cell viability was determined by counting colony forming units (CFU).

2.5 Maize stress treatments

Surface sterilized seeds were germinated in moistened filter papers, in the dark at 28ºC. To grow plants, germinated seeds were transferred to autoclaved paper towels, rolled in groups and placed vertically into 1:4 (v/v) MS medium in a growth chamber for two weeks at day/night temperature of 28/26 ºC, 16-h photoperiod, 60% relative humidity. Stress treatments with 15%PEG, dehydration at constant relative humidity, 200 mM NaCl ,100 µM ABA and 4ºC growth chamber, were performed during 3 days on two weeks old plants. Heat shock at 48ºC was performed during 45 min. For ABA treatment 12 dap (days after germination) manually dissected embryos were incubated with 100uM ABA during 4 hours.

2.6 RNA extraction and semi-quantitative RT-PCR analysis

Total RNA was extracted from mature embryos (60dap), leaves and roots from maize were using TRIZOL reagent (Invitrogen) according to the manufacturer's instructions (Gibco BRL, USA). Quantitative reverse transcription (Qiagen) was used for CDNA synthesis. For semi-quantitative RT–PCR, gene-specific primers were used to amplify Mlg3, Rab17 and

Emb564 (table of primer). Control reactions with *Z.mays* actin gene fragment were performed to ensure that equal amounts of RNA were used in each set of reactions. The actin gene was amplified with specific primers (table of primer). Samples were denatured for 5 min at 94ºC and then run for 25 cycles of 30 s at 94ºC, 45 s at 55ºC and 2 min at 72ºC with a final extension of 5 min at 72ºC. PCR products were separated by agarose gel electrophoresis.

2.7 Transient expression in *N. benthamiana* leaves and laser-scanning confocal microscopy. Transgenic Arabidopsis plants.

Full-length cDNAs for Mlg3 (640bp), Rab17 (630bp) and Emb564 (283 bp) were isolated from maize embryo RNA by RT-PCR using specific primers (table of primer). For Mlg03, PCR products with *BgL*II and *Spe*I ends were cloned into pCR2.1-TOPO. Mlg3 fusion to the GFP C-terminus was generated with *BgL*II and *Spe*I restriction enzymes and ligated into binary vectors PC1302 (Clontech) under the promoter 35S. For Rab17 and Emb564, *NCO*I and *Spe*I end sites were used. Leaves from 5 weeks old tobacco plants were transiently transfected with *Agrobacterium tumefaciens* GV3101/pMP90 transformed with the 35S::Lea protein-GFP constructs and the silencing suppressor HcPro (Bracha-Drori et al. 2004). After two days, infiltrated areas were examined by confocal microscopy at excitation 488nm/emission 500–600nm (Olympus FV 1000). For stress treatments, two-days infiltrated plants were irrigated with 15% PEG or transferred to 4ºC for one day, or heat-shocked at 48ºC for 45 minutes. Fluorescent cell perimeter measurements were performed using confocal software FV10-ASW2.1. Cellular fractionation of agroinfiltrated leaf disks into soluble, nuclear and insoluble fractions was performed as described (Sano et al. 2007). *Arabidopsis thaliana* transgenic lines were produced by *Agrobacterium*-mediated transformation via floral dipping method (Clough and Bent 1998). cDNA Rab17, mRab17 were cloned in pCAMBIA 1302 under the control of a CamV 35S promoter and fused in the 35 region with the *GFP* (Riera et al. 2004). Transgenic seedlings were selected by kanamycin 50mgml^{-1} and hygromycin 40mgml^{-1}.

Table of primer

Oligonucleotides used in this study		
Primer name	Sequence 1	Use (restriction site introduced)
Mlg3 f pET28a	5-ATGGCTTCCCACCAGG-3	Forward primer for *EXOP*I
Mlg3 r pET28a	5-CCTAGTGATCCCTGGTGATGG-3	Reverse primer for *EXOP*I
Emb564 f pET28a	5-ATGGCTTCCCACCAGG-3	Forward primer for *EXOP*I
Emb564 r pET28a	5-CTACTTGGTGAACTTGGACTCGTGCATCGTG-3	Reverse primer for *Hind* III
Mlg3 f	5-AAGGCGTCCGACACCGGCAGCT-3	Forward primer of RT-PCR
Mlg3 r	5-TCGTCTCGCCGGCCTTCTGCT-3	Reverse primer of RT-PCR
Rab17 f	5-GAAAGGAGGAAGGGAATC-3	Forward primer of RT-PCR
Rab17 r	5-AAGGCGTCCGACACCGGCAGCT-3	Reverse primer of RT-PCR
Emb564 f	5-AAGGCTCGCGAGGGCGAGACC-3	Forward primer of RT-PCR
Emb564 r	5-ACTCGTCGATCGATCGTGacgCCCTC-3	Reverse primer of RT-PCR
Actin f	5-CCAAGGCCAACAGAGAGAAA-3	Forward primer of RT-PCR
Actin r	5-GATACCGGAGCCGGACTCCTCCGCCTC-3	Reverse primer of RT-PCR

2.8 Cell stains

Cell viability of infiltrated tobacco leaf disks was assessed by 1% (w/v) Evans Blue stain for 30 min in distilled water (Jacyn Baker and Mock 1994) and examined under epifluorescence (Axiophot Zeiss microscope). Oil bodies were stained with 0.1mg ml^{-1} Nile red (Sigma-Aldrich) and fluorescence observed at excitation 552/emission 636 nm (Fowler and Greenspan 1985). Plasmatic membranes were labeled with FM1-43 dye and observed at excitation488 /emission 580 nm (Conde et al. 2007).

2.9 Lipid body isolation and Oleosin Coimmunoprecipitation

Four days germinated Arabidopsis seedlings were homogenized in 50 mM HEPES pH 7.5 extract buffer containing 10% w/w sucrose and standard protease inhibitors and lipid bodies were purified using a two-layer flotation method (Tzen et al. 1997). Integrity of purified lipid bodies was tested with Nile red stain. Presence of oleosin protein in lipid body samples was detected by immunoblotting with antibodies against *Oryza sativa* oleosin (provided by L. Montesinos and M. Coca, unpublished). Experiments of coimmunoprecipitation were performed as described (Zienkiewicz et al. 2010). Briefly, proteins from four-days germinated Arabidopsis seedlings were extracted in the above buffer and incubated with anti GFP antibodies (Invitogen) for 4 hours at 4ºC under gentle

agitation, followed by 2 hours incubation with 40 µl of packed protein A-Sepharose beads. The immuno complexes were collected by centrifugation, washed and resuspended in electrophoresis sample buffer.

2.10 Phosphorylation assays

In vivo phosphate labeling was performed on excised mature maize embryos by incubation with 5 µCi ^{32}P-ATP for 4 hours at 20ºC (Riera et al. 2001). For immunoprecipitation experiments, labeled extracts were incubated with Emb564 antibodies and protein A-Sepharose beads, as described above. *In vitro* phosphorylation assays with protein kinase CK2 were performed as described (Riera et al. 2004).

2.11 Mass Spectrometry (LC-MS/MS)

Proteins were in-gel digested with trypsin and analyzed by LC-MS/MS (Oliveira et al. 2007). Briefly, proteins were reduced with 10 mM dithiothreitol and alkylated with 55 mM of iodine acetamide. After washing, proteins were digested overnight at 37ºC and tryptic digested peptides were analyzed by liquid chromatography tandem mass spectrometry (CapLC-nano-ESI-Q-TOF). Samples resuspended in 1% formic acid were injected for chromatographic separation in a reverse-phase capillary C18 column (75 µm ID and 15 cm length, PepMap-LC Packings). Data, in PKL file format, were submitted for database search using MASCOT search engine. The following parameters were used: Database: NCBInr; enzyme: trypsin; 1 missed cleavage; fixed modifications: carbamidomethylation of Cys; variable modifications: oxidation of Met; peptide mass tolerance: 100 ppm; fragment tolerance: 0.25 Da**.**

2.12 PTM analysis by mass spectrometry

For phosphorylation determinations, trypsin digested protein bands excised from SDS acrylamide gels were analyzed by LC-MS/MS in a NanoAcquity (Waters) HPLC coupled to a LTQ-OrbitrapVelos mass spectrometer (Thermo Scientific). Tryptic extracts, previously enriched in phosphopeptides with TiO_2, were analyzed with phosphorylation-directed data dependent acquisition methods in the mass spectrometer. A TiO2 micro-column was built by stamping out a small C8 material plug from an extraction disk (3M

EmporeTM, Bioanalytical Technologies). The plug was placed at the constricted end of a P20 GeLoader tip (Eppendorf). The micro-column was packed with the TiO_2 beads (suspended in 100% acetonitrile) on top of the stamped plug, by applying air pressure. The micro-columns obtained were approximately 3mm long. The tryptic dried down extract was resuspended in loading buffer (1 M glycolic acid) in 80% acetonitrile, 5% TFA and loaded onto the packed TiO2 micro-columns. Flow through was also collected for analysis. The TiO_2 micro-column was washed first with 25 µl loading buffer and then with 30 µl 80% acetonitrile, 5% TFA. These washing volumes were also collected and pooled with previous flow through. The phosphopeptides bound to the TiO_2 micro-columns were eluted using 60µl ammonium water (pH 11), followed by 2 µL 50% acetonitrile to elute phosphopeptides bound to the C8 plug. The eluent was acidified by adding 5 µl 100% formic acid and analyzed by mass spectrometry.

The flow-through and washing buffers (FT fraction) were collected and pooled to be further analyzed. They were then dried down in a speed vacuum centrifuge, resuspended in 50µL 0.1 % TFA and 25 µL of those 50 were desalted with a microcolumn made by stamping out C18 material plugs from an extraction disk (3M EmporeTM, Bioanalytical Technologies). The sample was washed twice with 25µL 0.1% TFA and eluted from thecolumn with 2 µL 50% acetonitrile, 0.1% TFA . 48 µL formic acid 1% were added for MS analysis. For other PTMs, trypsin digested 2D spots were analyzed in a NanoAcquity coupled to an LTQ OrbitrapVelos mass spectrometer. Mass spectrometry was performed in a NanoAcquity (Waters) HPLC coupled to an LTQ OrbitrapVelos mass spectrometer (Thermo Scientific). 5µL samples were injected and separated in a C18 reverse phase column (75 µm Øi, 10 cm, nano Acquity, 1.7µm BEH column, Waters) using a gradient of 1 to 40 %B in 20 minutes followed by 40-60%B in 5 minutes at 250 nL/min flow rate (A: 0.1% formic acid and B: 0.1% formic acid in ACN). Eluted peptides were ionized in an emitter needle (PicoTipTM, New Objective). Spray voltage applied was 1700-2100V. Fractions were analyzed with phosphorylation-directed data dependent MS methods. Both fractions were analyzed using multistage activation MS method (MSA). Furthermore, the TiO2 fraction was also analyzed with MS3 acquisition (NL-MS3).

MSA method: peptide masses were measured in full scan in the Orbitrap at a resolution of 60,000 (*m/z:* 350-1800). Up the 5 most abundant peptides (minimum intensity of 500 counts) were selected from each MS scan and fragmented using CID (38% normalized collision energy) in the linear ion trap with helium as the collision gas. During fragmentation, if a phosphate neutral loss of 98, 49, 32.66 or 24.5 *m*/*z* below the precursor ion mass was detected, there was an additional activation of all four neutral loss *m*/*z* values.

Neutral loss method: peptide masses were measured in full scan in the Orbitrap at a resolution of 60,000 (*m/z:* 350-1800). Up the 5 most abundant peptides (minimum intensity of 500 counts) were selected from each MS scan and fragmented using CID (38% normalized collision energy) in the linear ion trap with helium as the collision gas. Loss of either 24.5, 32.6, 49 or 98 triggered a third MS event for the top 3 ions. For both acquisition methods, raw data were collected with Thermo Xcalibur. Data analysis was performed in Proteome Discoverer (v.1.2.0.208) using Sequest search engine. The parameters used to run the searches were: Uniprot-SwissProt (version January 2011) database; enzyme: trypsin; 2 missed cleavage; fixed modification: carbamidomethylation of Cys; variable modifications: Oxidation of Met, deamidation (N,Q), methyl (D, E, K, S), acetyl (K), phospho (S,T,Y) (only in case of the phosphopeptide enriched sample); peptide tolerance: 10 ppm and 0.6 Da (respectively for MS and MS/MS spectra) Fragment spectra were manually inspected to confirm phosphorylation sites and other PTM described here.

3. Results

3.1 The unfolded subproteome in maize embryos: enriching for LEA proteins

Maize embryo protein extracts were enriched for unfolded proteins by boiling followed by acid treatment. A number of thermostable proteins remain soluble at 3% TCA, or even at higher TCA concentrations (Figure 1A). The presence of LEA proteins was monitored using antibodies against group 2 Rab17 protein (Figure 1B). The protein fraction containing heat and 3%TCA soluble maize proteins was run in acrylamide gels and twelve visible Coomassie stained bands and six unstained gel zones in between these bands were excised for subsequent mass spectrometric analysis (LC-MSMS) (Figure 1C). Twenty different maize polypeptides were successfully identified as components of the unfolded

subproteome of maize embryo (Table 1). They are classified into three functional categories: LEA proteins, DNA binding proteins, and stress-related.

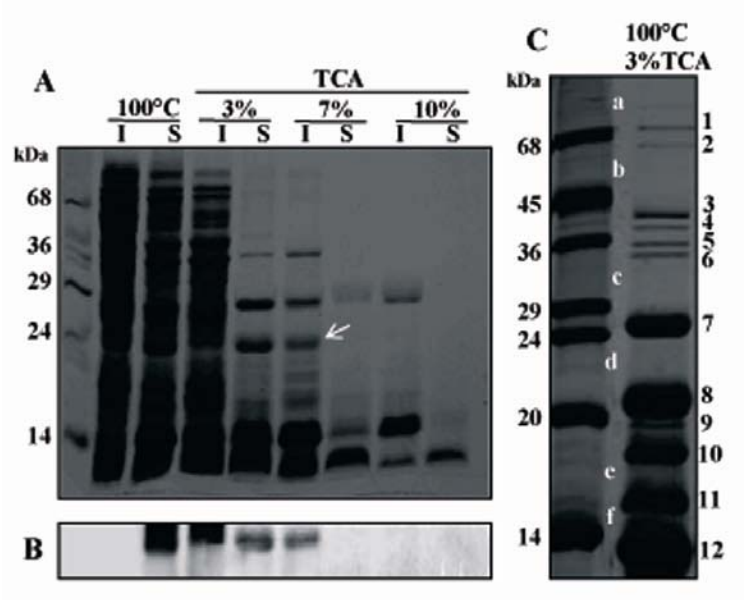


Figure 1. Maize embryo heat and acid soluble subproteome.

A. SDS-electrophoretic pattern of insoluble (I) and soluble (S) maize embryo proteins after heating at 100ºC. The heat-soluble fraction was treated with increasing trichloroacetic acid concentrations, 3%, 7% and 10%. Proteins are visualized with Coomassie G250 staining. The arrow indicates the position at 23 kDa of group 2 Rab17 LEA protein. B. Immunoblot of the above samples developed with polyclonal antiserum against group 2 Rab17. C. Subproteome containing maize embryo thermostable and 3%TCA soluble proteins. The isolated proteins were run in 18 cm long SDS acrylamide gel, stained with Coomassie G250 and the gel was cut along the entire length excising visible protein bands (1-12) and non-stained gel areas in between bands (a-f). After trypsin digestion samples were analyzed by LC-MSMS (Table 1). On the left, the molecular weight markers are indicated. The amount of protein loaded corresponds to 20 µg.

All share a common profile of totally unfolded proteins and they migrate at higher molecular masses than predicted. Exceptions to complete disorder are Rab28, a group 5 LEA protein which is only partially unfolded and two globular non-related proteins which were classified under the miscellaneous group (Table 1). The LEA subset comprises proteins from groups 1, 2, 3, 4 and 5. Two major group 1 LEA proteins (Pfam 00477) Emb5 and Emb564, were identified in bands 11 and 12 respectively (Figure 1C). They contain two and one copies respectively of the hydrophilic 20-amino acid motif characteristic of group 1 and show blast homology to cotton Em proteins and Atm1

(gi/15231041) and Atm6 (gi/15225645) from Arabidopsis. From group 2 (Pfam 00257), only Rab17 in band 8 was identified (Figure 1C). By contrast, group 3 (Pfam 02987) was found to be the largest LEA subgroup. Mlg3, an ABA and osmoticum induced maize protein (Rivin and Grudt 1991) was identified in band 7 (Figure 1C). This protein contains five copies of the 11 amino acid motif characteristic of group 3 and its ortholog in Arabidopsis corresponds to LEA 76 (gi/1592674) with 87% sequence coverage. Another major group 3 LEA protein is DC-8 identified in band 3 (Figure 1C) which holds a signal peptide in its N-terminal sequence and shows homology to DC-8 from *Daucus carota* (gi/18333) and to Arabidopsis LEA32 protein (gi/4140257). Other LEA proteins corresponded to group 4 (Pfam 03760) or related proteins (Figure 1C, bands 9, 10 and 11). Rab28 from group 5 (Niogret et al. 1996) was identified in a non-stained gel area (c in Figure 1C) indicating that although not completely disordered, is still partially soluble in acidic conditions. The second largest group of unfolded proteins, the DNA binding category, included three proteins harboring methyl-binding domains (Figure 1C, band 2 and a), two histones and one protein containing a histone linker domain (Figure 1C, bands 4, 5 and d). Nucleic acid binding proteins, histones, histone-binding proteins and high mobility group protein HMG-I/HMG-Y have been previously identified as part of the unfoldome of *E.coli* and mammalian cells (Cortese et al. 2005; Galea et al. 2006). In another functional category, superoxide dismutase (Cu-Zn) a stress-related protein involved in ROS detoxifying metabolism was also identified (Figure 1, band 10). Overall, we have identified twenty proteins as components of the disordered subproteome of dry maize embryo, thirteen belonging to the LEA family.

Table 1 Thermostable and acid-soluble subproteome of *Zea mays* embryos

	Position in the gel	Group domain/Pfam	NCBI	Mascot score	Experimental mol. wt.	Theoretical mol. wt./pI	No. of peptide matches/% sequence coverage	IUPred
LEA proteins								
Embryonic protein DC-8	3	LEA 3/LEA _4, PF02987	gi/226497424	311	40	33.552/6.16	5/15%	+
Rab28	c	LEA 5/SMP, PF04927	gi/22460	80	29	27.693/4.9	5/7%	Partial
Mlg3	7	LEA 3/LEA _4, PF02987	gi/7387829	254	28	22.74/8.8	5/19%	+
LOC100286330	7	LEA 3/LEA _4, PF02987	gi/226529441	298	28	23.528/8.96	4/20%	+
Lea group 3	7	LEA 3/LEA _4, PF02987	gi/195606062	155	28	21.77/8.98	3/11%	+
Unknown protein	9	Possible LEA 4	gi/226499732	195	20	13.14/5.62	4/27%	+
Rab17	8	LEA 2/Dehydrin, PF00257	gi/126632258	102	23	17.5/8.75	3/19%	+
Seed maturation protein SMP	10	LEA 4/LEA_1, PF03760	gi/126632258	42	21	16.08/110.14	1/12%	+
EMB5	11	LEA 1/LEA_5, PF00477	gi/195623094	401	17	12.551/11.43	2/17%	+
Embryonic DC-8-like protein	11	LEA 3/LEA _4, PF02987	gi/226507683	233	17	14.22/6.62	4/43%	+
LOC100286326	11	LEA 4/LEA_1, PF03760	gi/226499754	115	17	11.63/9.46	2/39%	+
Seed maturation protein PM41	f	LEA 3/LEA _4, PF02987	gi/226500792	112	16	8.52/4.96	2/25%	+
EMB564	12	LEA 1/LEA_5, PF00477	gi/162463828	266	12	9.676/6.6	8/42%	+
DNA binding								
Methyl-binding domain protein	a	MBD superfamily	gi/162461620	107	72	43.741/4.46	2/7%	+
MBD 109	2	Methyl-binding protein	gi/22135477	222	64	42.21/4.47	6/15%	+
Putative MBD115	2	MBD2 methyl-binding domain	gi/57791234	154	64	38.39/4.53	1/12%	+
Histone H1	4	Histone	gi/195658563	73	38	26.67/10.7	2/10%	+
Histone H1-like protein	5	Histone	gi/162464456	168	36	42.8/10	3/12%	+
MHGI/Y	d	H15 histone linker	gi/162462851	128	28	19.65/10.51	3/19%	+
Stress-related								
Superoxide dismutase (Cu–Zn)	10	ROS metabolism	gi/134598	189	18	15.17/5.64	4/27%	+
Other proteins								
60S ribosomal protein L36-2	e	Ribosomal	gi/56182368	75	17	12.55/11.43	1/12%	Globular
Unknown protein	f	–	gi/194700720	112	16	18.93/9.48	2/17%	Globular
Not identified	1,b,6	–	–	–	–	–	–	

List of proteins excised from SDS–acrylamide gels identified by mass spectrometry (LC-MS/MS). Proteins are classified under LEA family, DNA binding, stress-related and other non-related proteins. The protein position in the gel, protein domains indicative of functional category, *Z. mays* accession number according to the NCBI Gene bank, Mascot score, experimental and theoretical molecular weight, pI, number of peptides matched and percentage of protein sequence coverage are indicated. The unfolded (+), partial or globular nature of each protein is indicated according to the disorder IUPred prediction site.

3.2 Two dimensional analyses of Emb564, Rab17 and Mlg3 LEA proteins: posttranslational modifications.

Next, three major LEA proteins belonging to the main groups in the LEA family we selected: Emb564 from group 1, Rab17 from group 2 and Mlg3 from group 3. Using polyclonal antibodies developed against each LEA protein, the immunoexpression pattern of the three LEA proteins was examined by immunoblotting in enriched-LEA protein fractions (Figure 2A). Interestingly, absence of cross reactions among the three antisera suggested different functional antigenic profiles in 1, 2, and 3 LEA groups. The same samples were then analyzed in 2-DE gels (Figure 2B) and the immunopatterns of native Emb564, Rab17 and Mlg3 LEA proteins were compared to the corresponding recombinant proteins produced in *Escherichia coli* (Figure 2C). Since recombinant Rab17 is not phosphorylated by *E.coli* cells, Rab17 antibodies recognize a single basic spot in bacterial extracts, in contrast to the characteristic acidic multispot pattern recognized in native Rab17, which is due to the phosphorylation in the S-segment (Goday et al. 1994; Plana et al. 1991; Riera et al. 2004) (Figure 2C). Using antibodies directed to Mlg3, 2-DE immunoblots revealed one major spot of 30kD and 8.5 pI in both maize embryo and *E.coli* extracts, and two additional 30kD minor spots only present in maize extracts (Figure 2C). The nature of these neighbor spots was determined by LC-MSMS (spots 1, 2 and 3 in Figure 2B) and two different group 3 LEA proteins with similar sizes and isoelectric points,

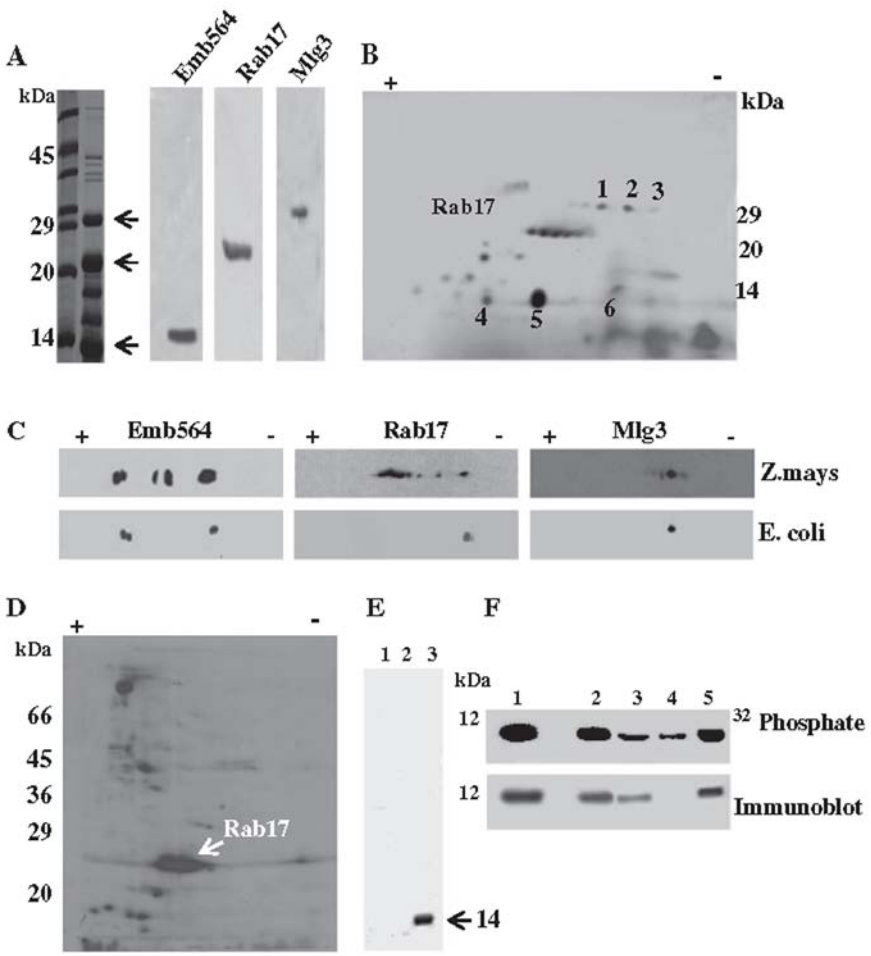


Figure2. Posttranslational modifications in group 1 Emb54, group 2 Rab17 and group 3 Mlg3 LEA proteins.

A. Electrophoresis and Coomassie staining of maize embryo enriched LEA subproteome containing heat-stable and 3% trichloroacetic acid soluble proteins (left). The arrows indicate the position of the three LEA proteins selected: group 1 Emb564 (12kDa), group 2 Rab17 (23kDa) and groups 3 Mlg3 (29kDa). Western blotting developed with rabbit polyclonal antisera raised against each LEA protein (right). **B**. 2-DE pattern of enriched LEA subproteome stained with Coomassie. Samples were separated by isoelelectric focusing using a 3.0-11.0 pH range. Numbers indicate the spots excised and analyzed by LC-MSMS. The position of Rab17 is indicated. **C**. Comparative 2-DE patterns of native proteins from maize (top) and recombinant proteins expressed in *E.coli* (below). Protein extracts were resolved by isoelelectric focusing in a 3.0-11.0 pH range. **D**. 2-DE map of *in vivo* phosphorylated maize embryo proteins. Proteins were incubated with ^{32}P-phosphate during the first four hours of imbibition. Group 2 Rab17 is one of the most abundant phosphorylated polypeptides in maize embryos (arrow). Proteins were focused in a 3.0-11.0 pH range. **E**. *In vitro* phosphorylation of recombinant Emb564 by CK2. Kinase assays with α1 catalitic subunit of CK2 (lane1), recombinant Emb564 (lane2) or Ck2α1 plus recombinant Emb564 as substrate (lane3). The arrow indicates the position of the phosphorylated Emb564 band. **F**. Emb564 is *in vivo* phosphorylated in mature maize embryos. Autorradiography of maize ^{32}P labeled extracts immunoprecipitated with anti Emb564 antibodies. Total ^{32}P labeled embryo extracts (lane1), extracts after 2 hour incubation with sepharose beads plus Emb564 antibodies (lane2), washes with incubation buffer (lanes 3, 4) and sepharose beads recovered after centrifugation (lane 5). Below, western blotting of the same samples developed with anti-Emb564 antibodies.

which were the expression of genes gi/226529441 and gi/195606062, were identified (Table 1).The high homology of these LEA proteins to Mlg3, 76% and 86% respectively, explains why Mlg3 antibodies detect three spots in maize 2DE gels and rules out the existence of posttranslational modifications in embryo native Mlg3 protein. The two dimensional pattern of group 1 Emb564 protein showed to be more complex. Emb564 antibodies recognize four major 14kD spots in the 6.6-8.7 pI range in embryo extracts, and only two spots in the recombinant Emb564 protein (Figure 2C). The spots were not artifacts produced during the heat and acid extraction procedure because identical spot pattern was obtained when proteins were extracted in standard conditions (not shown). Individual LC-MSMS analysis, showed that all spots corresponded to Emb564 (spots 4, 5 and 6 in Figure 2B), strongly suggesting the existence of posttranslational modifications in Emb564 protein. Emb564 contains several predicted CK2 phosphorylation sites (Figure S1) and thus, *in vivo* ^{32}P-labeled embryo extracts were analyzed by 2-DE electrophoresis (Figure 2D); immunoprecipitation with Emb564 antibodies confirmed that Emb564 was effectively *in vivo* phosphorylated (Figure 2E).

A

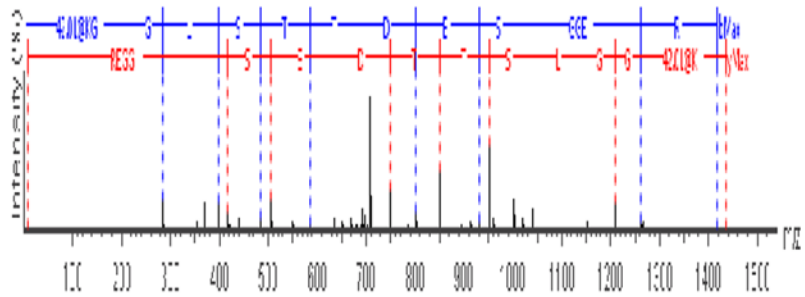
Emb564 : gi 1169520

masgqesrke ldrkareget vvpggtggks
veaqehlaeg rsrggqtrre qlgqqgysem
gkkgglsttd eSggeraare gvtideskft k

B

Site	Kinase Motifs
Ser3	CK2
Ser7	CK2, PKC
Thr47	CK2, PKC
Thr68	CK2
Ser72	CK2
Thr83	CK2

C



D

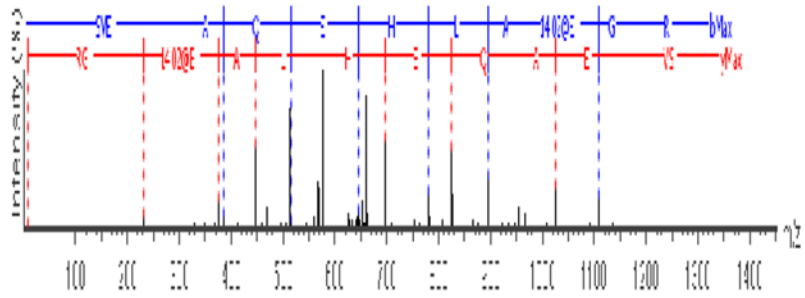


Figure S1. Identification of phosphorylation, methylation and deamidation sites in Emb564 LEA protein by mass spectrometry.

A. Sequence of Emb564 protein with the putative phosphorylation sites market in green (threonine) and red (serine). B. Table listing the putative phosphorylation sites according to http://expasy.org/cgi-bin/prosite/ScanView. Abbreviations: CK2 casein kinase 2, PKC protein kinase. C. Fragmentation spectrum showing acetylation in K-63 in Emb564 protein. D. Fragmentation spectrum of the peptide SVEAQEHLAmethylEGR showing the methylation in E-39 in Emb564 protein.

Moreover, CK2 was able to *in vitro* phosphorylate Emb564 (Figure 2F) and Serine 72 was identified as the site of phosphorylation by mass spectrometry (Table 2A). However, a single phosphorylation would hardly explain the complex 2DE pattern observed; in fact, in alkaline phosphatase-treated samples no appreciable changes in the position of Emb564 spots were observed (not shown). The detailed analysis of the mass spectra generated from spots 4, 5 and 6 (Figure 2B) showed the presence of several variants of trypsinized peptides: acetylation in lysine residue K-63 in spot 6 and several methylations in glutamic and aspartic acid residues E-19, E-32, E-59, E-71 and E-75 in spot 5 and E-49, E-32, E-35, E-39, D-70, E-71 and E-75 in spot 6 (Table 2B). Chemical modifications such as deamidations in glutamine residues Q-34 were also detected in spot 6. Although deamidation and lysine acetylation introduce acidic balance of charges to the protein, methylation modifications in acidic residues confer a more basic charge balance as observed in the 2D pattern (Figure 2B). These protein isoforms generate a mixed population which exhibits peptide sequences containing different modifications (Table 2B). The combination of these posttranslational modifications might explain the complex 2D map observed in embryo native Emb564 protein.

Table 2. Identification of phosphopeptides and other PTMs in Emb564 by mass spectrometry.

A

Peptide	Sequence	Score	*m/z*	Charge	Activation type
S72	KGGLSTTDEsGGER	116	737.31	2^+	CID

B

			Spot 5 (modified peptides)		Spot 6 (modified peptides)		
Sequence	From–to	Molecular mass	Methylation	Deamidation	Methylation	Dimethylation	Acetylation
EGTVVPGGK	17–29	1,186.58	120.6		1,200.6		
EQLGQQGYSEMGK	49–55	1,453.68	1,467.7	1,454.66			
GGLSTTDESGGER	64–72	1,264.55	1,278.57		1,278.57	1,292.59	
KGGLSTTDESGGER	72–76	1,392.68	1,406.7		1,406.7		1,434.69
REQLGQQGYSEMGK	49–57	1,609.78	1,623.8				
SVEAQEHLAEGR	30–41	1,324.66	1,338.68	1,325.64	1,338.68		

A. Phosphorylated site, serine-72, in Emb564 protein was confidentially determined from the fragmentation mass spectrum. **B**. Methylation, deamination and acetylation in Emb564 protein (spots 5 and 6, Figure 2B). Sequence and molecular mass of trypsinized peptides were detected by LC-MSMS. From-to represents the amino acid positions in the Emb564 protein sequence.

3.3 Recombinant Emb564, Rab17 and Mlg3 LEA proteins: in vitro antiaggregation effects and in vivo protection.

To test the antiaggregation activity of maize LEA proteins *in vitro*, *E.coli* salt-soluble proteomes (SP) containing recombinant LEA proteins were induced to aggregate by desiccation, heat-shock or freezing. The three LEA proteins were produced in similar concentrations by bacteria after IPTG induction, approximately at a mass ratio of 0.5 LEA protein: 1 *E.coli* SP (Figure 3A). In these conditions Emb564, Rab17 and Mlg3 proteins conferred different levels of antiaggregation protection as measured by a light scattering assay (Figure 3C). A control with added BSA (mass ratio BSA: SP 1:1) was also included to compare the antiaggregation effects of a non-related globular protein. Rab17 behaved as a very potent antiaggregation agent under desiccation, heating and freezing, providing almost complete protection to the formation of protein aggregates. Emb564 was mildly efficient, only partially avoiding the formation of protein aggregates. The protection conferred by Rab17 or Emb564 was independent on the form of stress applied, ie, drying, heating or freezing. By contrast, the protection efficiency of group 3 Mlg3 LEA protein was stress-dependent; under dehydration, Mlg3 provided a strong protective effect similarly to Rab17, but was mildly efficient when the aggregation of the *E.coli* proteome was induced by heat or freezing (Figure 3C). The composition of the aggregated proteins in desiccation

trials was further analyzed by one-dimensional acrylamide gels (Figure 3B). After rehydration of desiccated samples, proteins that remain soluble and those that become insoluble were separated by centrifugation. In control *E.coli* cells, desiccation produces a general aggregation process and bacterial proteins become insoluble (Figure 3B). In the presence of Rab17 or Mlg3 the solubility of the bacterial proteome upon drying and subsequent rehydration is globally preserved; this discards the possibility of specific affinities for a subset of proteins (Figure 3B). The weaker Emb564 antiaggregant effect was evidenced by the distribution of bacterial proteins in soluble and insoluble protein fractions; in fact, the Emb564 protein itself is prone to certain degree of aggregation as indicated by its presence in the sediment (Figure 3B).

To investigate the protective properties of recombinant LEA proteins *in vivo*, the growth of *E.coli* cells over-expressing Emb564, Rab17 and Mlg3 LEA proteins was measured under stress (Figure 3D-G). The effects of temperature were assessed by the reduction of colony formation after heat shock or freezing treatments in IPTG-induced cultures. In general, cells producing LEA proteins showed higher survival rates than controls (Figure 3D and 3E). Cells over-expressing Rab17 exhibited the highest tolerance to heat-shock and freezing temperature, producing a number of colonies 9-fold that of controls. Over-expression of Emb564 or Mlg3 contributed to intermediate tolerances to temperature stress, with 4.5-fold colony number compared to control cells (Figure 3E). Dehydration stress was mimicked by adding 10% PEG to the culture medium. Again, cells producing LEA proteins showed higher growth rates compared to control cells (Figure 3F). The protection effect was evidenced by the shorter lag phase and higher cell concentration achieved. Interestingly, of the three LEA proteins tested, Mlg3 provided the highest protection against the osmotic stress imposed by PEG (Figure 3F). In non-stress conditions, the growth rate was similar in all *E.coli* cells (Figure 3G). In summary, the highest general protective efficiencies corresponded invariably to group 2 Rab17 under any stress and to group 3 Mlg3 under dehydration. Therefore, a correlation can be established between the *in vitro* antiaggregation properties of recombinant Emb564, Rab17 and Mlg3 LEA proteins and their *in vivo* protection efficiencies.

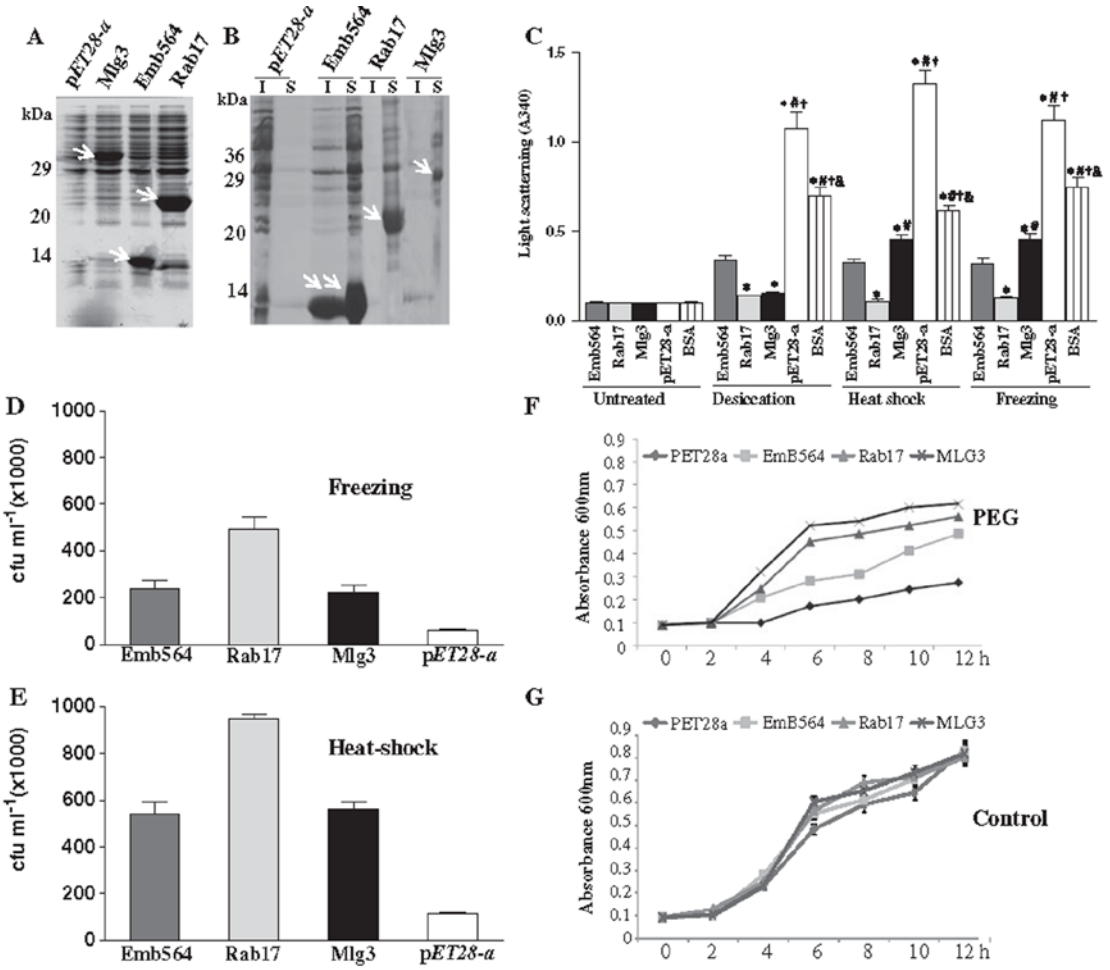


Figure 3. Antiaggregation properties of Emb564, Rab17 and Mlg3 LEA proteins and protective effects on *E.coli* growth under stress.

A. Expression of recombinant Mlg3, Rab17and Emb564 LEA proteins in *E.coli.* Coomassie-stained polyacrylamide gel indicating the presence of the different LEA proteins in bacterial extracts five hours after IPTG induction. B. Solubility of *E.coli* salt-soluble proteome after desiccation and subsequent rehydration. Insoluble (P) and soluble (S) proteins were separated by centrifugation. The arrows indicate the position of recombinant LEA proteins. C. Anti-aggregation activities of Emb564, Rab17 and Mlg3 recombinant proteins. Light scattering of *E.coli* salt-soluble proteomes expressing LEA proteins or control pET28a measured as absorbance at 340nm after desiccation, heat shock, freezing and untreated controls. BSA, bovine serum albumin, was included as globular protein control. Data are the average of three independent measurements. Variance analysis followed the Student-Newman-Keuls test, *$p<0.05$ versus Emb564,# $p<0.05$ versus Rab17and† $p<0.05$ versus Mlg3.& $p<0.05$ versus BSA. D and E. Freeze and heat survival of IPTG-induced *E.coli* cultures expressing Emb564, Rab17 or Mlg3 LEA proteins or control pET28a. Effects of freezing (D) and heat shock (E). Number of colonies (cfu ml^{-1}) was measured after 16 hours of culture. Data reproduce three independent duplicate experiments. SD bars are indicated. F and G. Growth kinetics under osmotic challenge. Growth of IPTG-induced *E.coli* cultures producing Emb564, Rab17, Mlg3 LEA proteins or control pET28a in medium supplemented with 10% PEG (F) or in standard medium (G). The increasing density of the liquid cultures was measured at 600nm absorbance.

3.4 LEA-GFP fusions in tobacco epidermal cells: group 3 Mlg3 LEA protein reduces dehydration-induced cell shrinkage effects

In maize vegetative tissues, Rab17 and Mlg3 LEA transcripts and proteins are induced by dehydration, low temperature, heat-shock, salinity or ABA in leaves and roots, whereas Emb564 expression is restricted to embryonic tissues (Figure S2). To investigate *in vivo* the subcellular localization and the potential protective effects of the different LEA proteins in plant vegetative tissues, transient expressions of GFP-fused LEA proteins were analyzed in agroinfiltrated tobacco leaves. Under non-stress conditions, confocal fluorescence imaging revealed that Emb564, Rab17 and Mlg3 GFP-fused proteins were distributed into cytosol and nuclear compartments (Figure 4A, Figure S3). Integrity of the fusion proteins in agroinfiltrated leaves was verified by immunoblotting (Figure S3), confirming that the fluorescence observed under the microscope corresponded to intact GFP fusions and not to *in vivo* breakdown products. In the cytoplasm, florescence from Emb564, Rab17 or Mlg3 GFP fusion proteins was excluded from the vacuoles and the cell wall, and decorated the entire cytosol (Figure 4A). The cytoplasm in tobacco epidermal cells is pressed between vacuoles and cell walls, and flows between these structures showing a channel-like appearance. In the nucleus the fluorescence was fainter, showing a diffuse uniform pattern with no defined structures and absence of stained nucleoli suggesting a soluble non-structure associated localization.

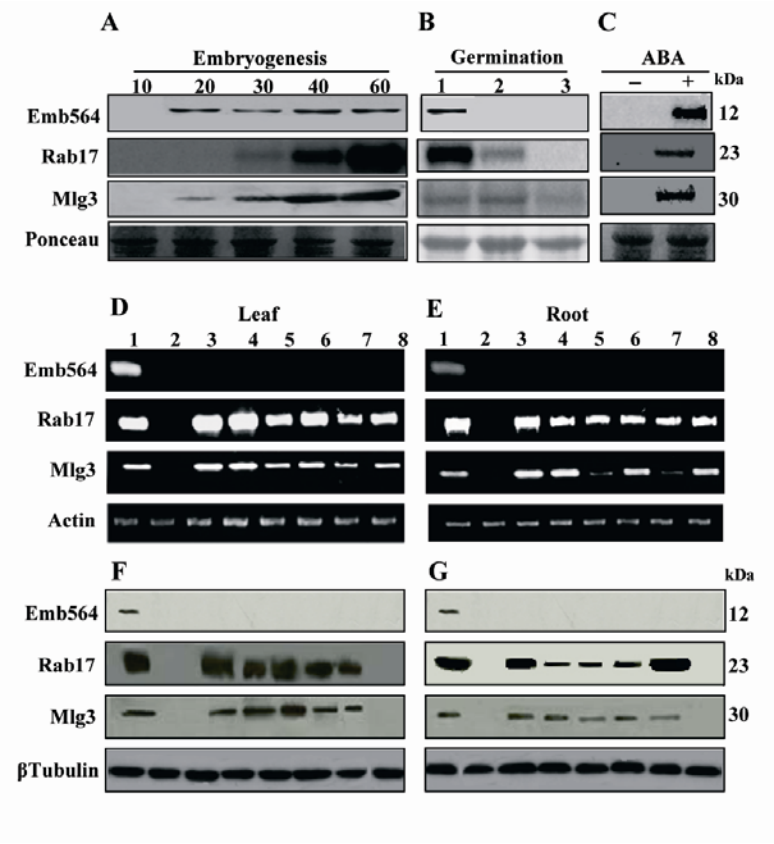
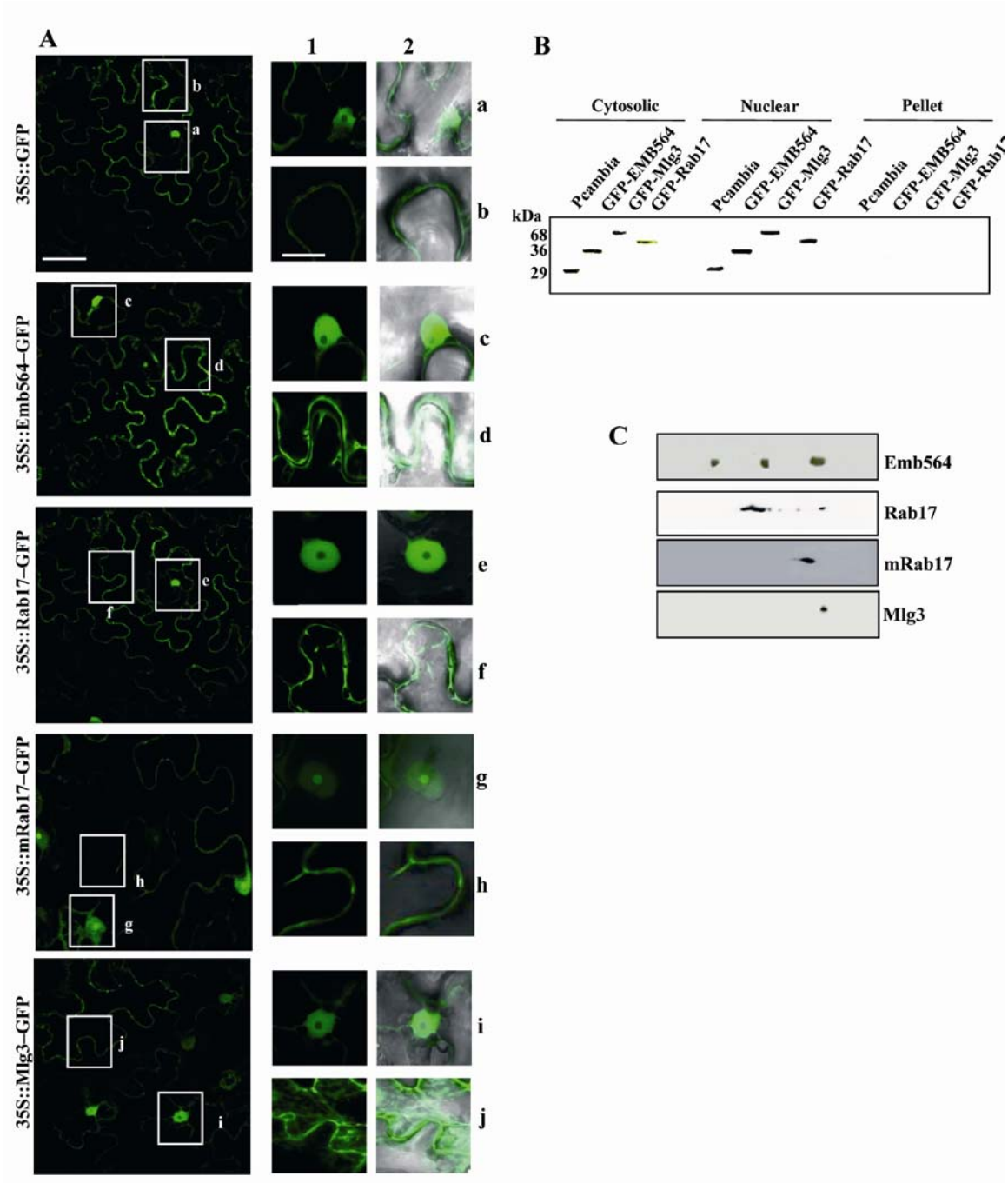


Figure S2. Expression pattern of group 1 Emb54, group 2 Rab17 and group 3 Mlg3 LEA proteins during embryogenesis and in vegetative tissues under stress.

A. Accumulation LEA proteins along the embryogenesis. Western blotting of protein extracts (15µg) from immature to mature maize embryo at the indicated days after pollination. LEA proteins were identified by hybridization with the antibodies against Emb564, Rab17 and Mlg3 LEA proteins. Ponceau Red staining is shown as loading reference. **B**. Monitoring LEA protein disappearance during germination. Western blotting of protein extracts (15 µg) from maize embryos in the first, second and third days after seed imbibition. **C**. Induction of LEA proteins by exogenous abscisic acid (ABA) in early embryogenesis. Immature 10 dap embryos were excised and incubated for 3 hours with water (-) or 100µM ABA (+) before protein extraction and immunodetection of LEA protein content. **D** and **E**. Expression level of the three *Lea* transcripts in leaves (D) and roots (E) from 2 week-old plants under stress, monitored by RT-PCR. Each lane was loaded with total RNA (2µg) from non-treated (lane 2), dehydration (lane 3), 25% PEG (lane 4), 200mM NaCl (lane 5), 4ºC (lane 6), 100µM abscisic acid (lane 7) and heat-shock (lane 8) treated plants. Transcript levels in mature (60 dap) embryo were included as controls (lane 1). Actin transcript was used as loading reference. **F** and **G**. Accumulation of LEA proteins in leaves (F) and roots (G) in stress treated plants. Immunoblot of leaf and root protein extracts (15µg) developed with LEA polyclonal antibodies. β tubulin was used as loading reference.



S 3 **A**. Subcellular localization by confocal microscopy of Emb564, Rab17, mRab17 and Mlg3-GFP fusion proteins in *Agrobacterium*-infiltrated *N.benthamiana* leaves. Left, expression pattern of fluorescent epidermal cells two days after agroinfiltration. Bar corresponds to 10 μm. Right, details of nucleus (a,c,e,g,i) and cytosol (b,d,f,g,j) showing GFP fluorescence (lane1) and GFP fluorescence plus bright field merged images (lane 2). Bar corresponds to 20 μm. **B**. Integrity of LEA-GFP protein fusions in *N.benthamiana* leaves two days after agroinfiltration. Cytosolic and nuclear-enriched protein fractions from agroinfiltrated tobacco discs were immunoblotted and developed with antiGFP antibodies**. C**. 2-DE western blot analysis of Emb564, Rab17, mRab17 and Mlg3-GFP fusion proteins in tobacco leaves two days after agroinfiltration. Protein extracts from agroinfiltrated tobacco discs were analyzed in 2-DE gels using a 3-11 pH gradient and western blots were develop with the corresponding LEA antibodies.

The subcellular locations of the fluorescent LEA fusion proteins were next examined under dehydration conditions. Agroinfiltrated tobacco plants expressing Emb564, Rab17, Mlg3 GFP-fusions or GFP were irrigated with 15% PEG which resulted in general mild stress symptoms in all plants. At the confocal microscope the distribution of fluorescence remained essentially unchanged compared to non-treated plants. (Figure 4 A, B). A general size reduction was evident in all fluorescent epidermal cells observed but, surprisingly, fluorescent Mlg3-GFP cells were considerably bigger than cells expressing Emb564, Rab17 fusion proteins or control GFP (Figure 4B). A similar reduction in size was observed in agroinfiltrated tobacco leafs detached and allowed to dehydrate overnight at room temperature but never during cold (not shown) or heat shock treatments (Figure 5). During dehydration cells tend to lose water and hence their volume is reduced. The distance between neighbor cells after PEG-induced dehydration was visualized using FM1-43, a fluorescent endocytosis marker that is taken up by intact walled cells and stains promptly the plasma membrane (Figure 4C). After PEG treatment, staining with FM1-43 revealed that epidermal cells over-expressing Mlg3-GFP remained closer compared to cells expressing Rab17-GFP or GFP, suggesting different degrees of cell size reduction or cell shrinkage upon water loss. Fluorescent cell perimeters were quantified by measuring the fluorescent contour of cell projections (Figure 4D). The measurements proved that after PEG challenge a general reduction in the cell fluorescent perimeter occurred; however, cells expressing Mlg3-GFP protein reduced their cell contour by 40% compared to fully hydrated cells, whereas cells expressing Emb564-GFP, Rab17-GFP or GFP proteins experimented perimeter reductions of 65%, 68% and 70% respectively (Figure 4D). Overall these results show that group 3 Mlg3 LEA protein, but not group 1 Emb564 or group 2 Rab17 proteins, has the ability to minimize the *in vivo* cell shrinkage effects due to water losses.

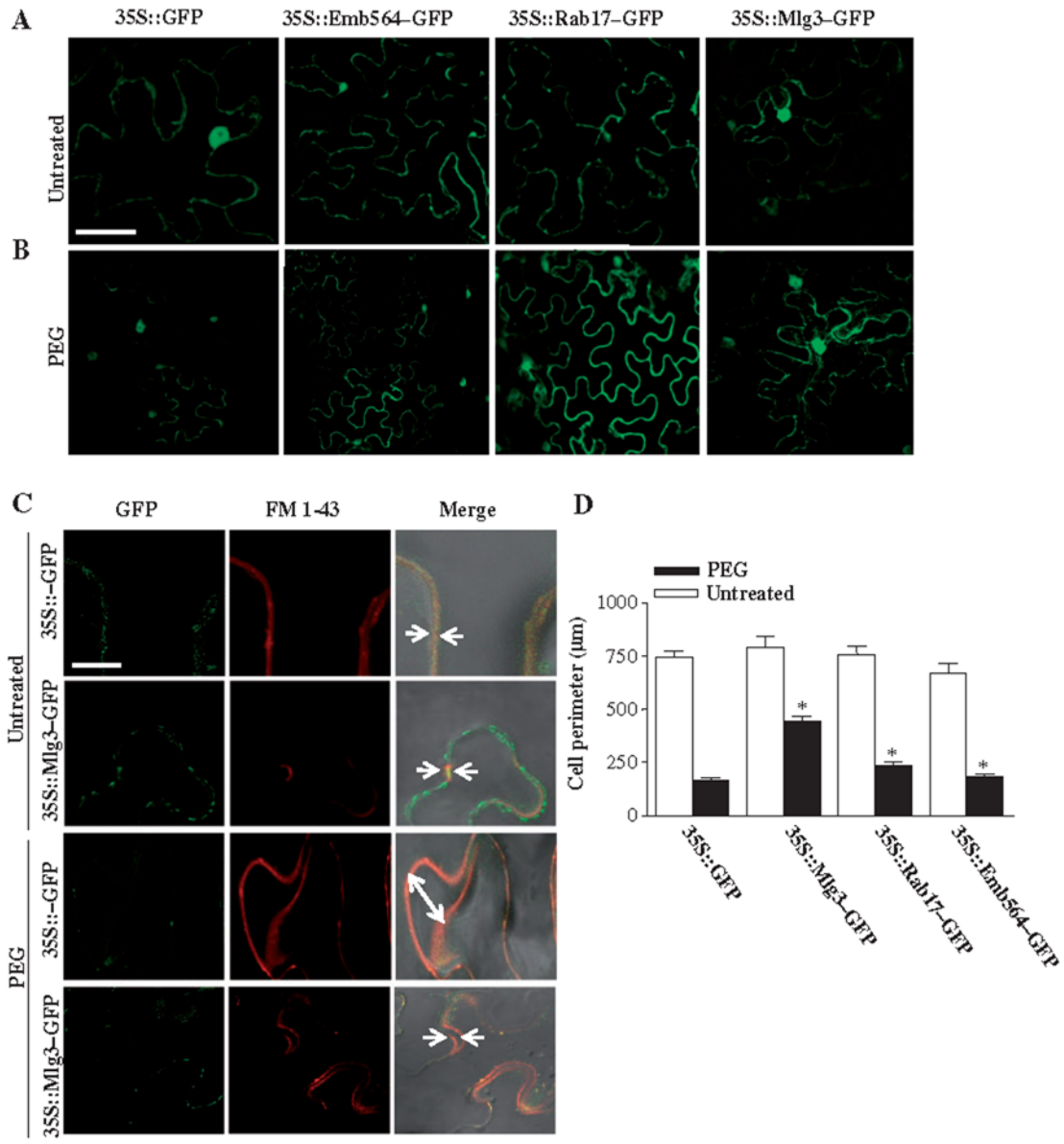


Figure 4. Mlg3-GFP fusion protein reduces osmotic shrinkage effects in *N.benthamiana* cells.

A. Subcellular distribution of Emb564, Rab17 and Mlg3 GFP fusion proteins and control GFP in agroinfiltrated *N.benthamiana* leaves from untreated plants. Fluorescence in the tobacco epidermal cells was visualized in cytosol and nucleus two days after *Agrobacterium tumefaciens* infiltration. Bar corresponds to 10 µm. **B**. Confocal imaging of Emb564, Rab17and Mlg3 GFP fusion proteins in agroinfiltrated *N.benthamiana* plants subjected to osmotic treatment. Tobacco plants two days after *Agrobacterium* infiltration were irrigated overnight with 15% PEG. **C**. Details showing cytosolic green GFP and membrane red FM1-43 fluorescences and merged images in Mlg3-GFP and control GFP agroinfiltrated epidermal cells. Untreated (top) and 15% PEG stressed cells (bottom). Arrows indicate the distance between neighbor cells before and after osmotic treatment. Bar corresponds to 20 µm. **D**. Quantitative analysis of cell perimeter before and after 15% PEG treatment. Measurements were performed using confocal software FV10-ASW2.1. Data were obtained from three independent experiments with n=30 cell perimeter measurements. Variance analysis followed Student-Newman-Keuls test, *P<0.05.

3.5 Group 2 Rab17 LEA protein preserves cell viability under heat shock and colocalizes to oil bodies.

Tobacco plants harboring Emb564, Rab17 or Mlg3 GFP-fusions were subjected to heat shock. The treatment, 48ºC for 45 minutes, produced general wilty symptoms in all plants. However, at the microscope level a reduction in the fluorescence intensity was evidenced in cells expressing Emb564-GFP, Mlg3-GFP proteins or GFP but not in cells expressing Rab17-GFP (Figure 5).

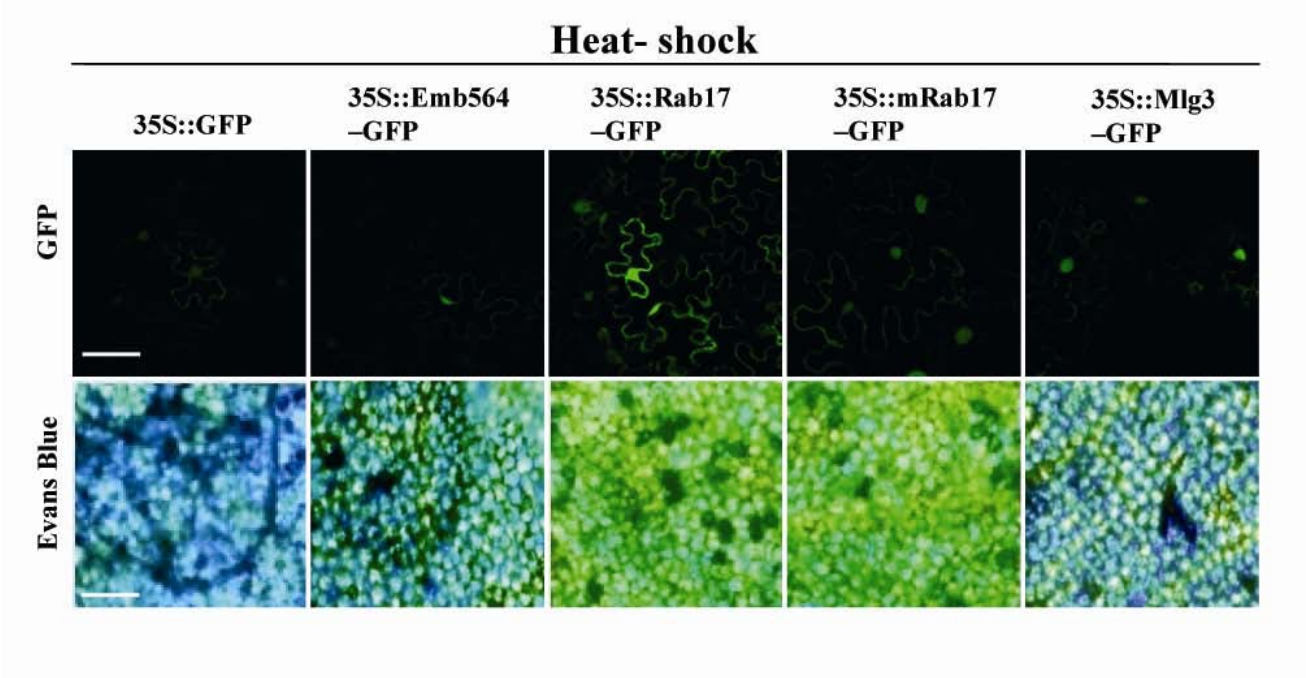


Figure 5. Cell viability in epidermal leaf cells after heat shock.

Top panels: Laser-scanning confocal micrographs of heat-shocked *N.benthamiana* epidermal cells expressing GFP or GFP-fused to Emb564, Rab17, mRab17 (the non-phosphorytable CK2 mutated version of Rab17) and Mlg3 LEA. Tobacco plants were subjected to heat shock in the second day after agroinfiltration. Fluorescence is distributed into cytosol and nucleus. Scale bar indicates 10 μm. Lower panels: Evans Blue stain. Following heat shock, leaf discs were immediately infiltrated with Evans Blue and cells visualized by bright field microscopy. Dead cells are permeable to Evans blue and stain in blue. Scale bar indicates 20 μm.

Cell viability was assayed using Evans Blue stain which is permeable only to dead cells. In agroinfltrated leaf discs overproducing Emb564-GFP, Mlg3-GFP or GFP, most cells after heat-shock treatment stained in blue due to dye uptake. Viability was especially low in the epidermal cells that expressed the fluorescent GFP control protein. In contrast, leaf areas over-expressing Rab17-GFP were mostly impermeable to the stain and remained essentially green indicating that cell viability was high and that cell membranes were intact (Figure 5). Because Rab17 is phosphorylated by maize tissues but not by *E.coli* cells, we examined the phosphorylation status of Rab17 in agroinfiltrated tobacco leaves. 2-DE electrophoresis confirmed that the fluorescence observed by confocal micrsocopy corresponds effectively to the phosphorylated forms of Rab17-GFP fusion protein (Figure S3). To evaluate the role of Rab17 phosphorylation, tobacco leaves were agroinfiltrated with mRab17, a mutant version of Rab17 LEA protein which carries an amino acid substitution in the CK2 consensus site that precludes its phosphorylation by the enzyme (Riera et al. 2004). Absence of phosphorylation in mRab17-GFP protein was confirmed by 2-DE (S3).

Interestingly, after heat shock tobacco leaves over-expressing mRab17-GFP were as resistant as those expressing wild type Rab17-GFP fusions as assessed by Evans Blue stain (Figure 5). These results indicate that over-expression of Rab17-GFP fusion protein, irrespectively of its phosphorylation status, preserved epidermal cell viability from the damaging effects of thermal shock. We believe that this effect relates to the high antiaggregation activity and protection conferred by the non-phosphorylated form of group 2 Rab17 to bacterial cells during heat shock, in contrast to group 1 Emb564 or group 3 Mlg3 LEA proteins (Figure 3).

After thermal shock, the general distribution of GFP fluorescence into cytosol and nucleus remained essentially unchanged (Figure 5). However, a closer observation at higher magnification evidenced the existence of bright small fluorescent organelles (0.5-1 μm diameter) close to the membranes in heat-shocked Rab17-GFP or mRab17-GFP expressing cells but not in Emb564-GFP, Mlg3-GFP or GFP expressing cells or in untreated controls (Figure 6A). These GFP-labeled structures co-localized to Nile red stained bodies. Nile red is a lipophyllic stain regularly used to stain oil bodies in seeds. In untreated epidermal cells

few globules stained with Nile red but upon thermal shock its number increased significantly (Figure 6A).

To confirm in another plant system the observed colocalization of Rab17 to leaf oil bodies, transgenic Arabidopsis plants over-expressing Rab17-GFP and its mutated version mRab17-GFP were generated. In these plants intracellular fluorescence is also distributed into cytosol and nucleus (Figure 6B). After heat shock, colocalization of Rab17-GFP/mRab17-GFP and Nile red fluorescences in leaves was evidenced (Figure 6B). Since oil bodies are especially abundant in cotyledons, we isolated enriched-lipid body suspensions from 4-days-old germinated seedlings subjected to heat-shock. In these suspensions, co-localization of Rab17-GFP /mRab17-GFP and Nile red fluorescences was again noticed (Figure 7A). Oleosins are structural proteins from seed oil bodies (Tzen et al. 1992). By using a polyclonal antibody against rice oleosins (provided by L. Montesinos and M. Coca, unpublished results) a 15kDa protein band corresponding to oleosin was immunodetected in Arabidopsis seed extracts (Figure 7B).

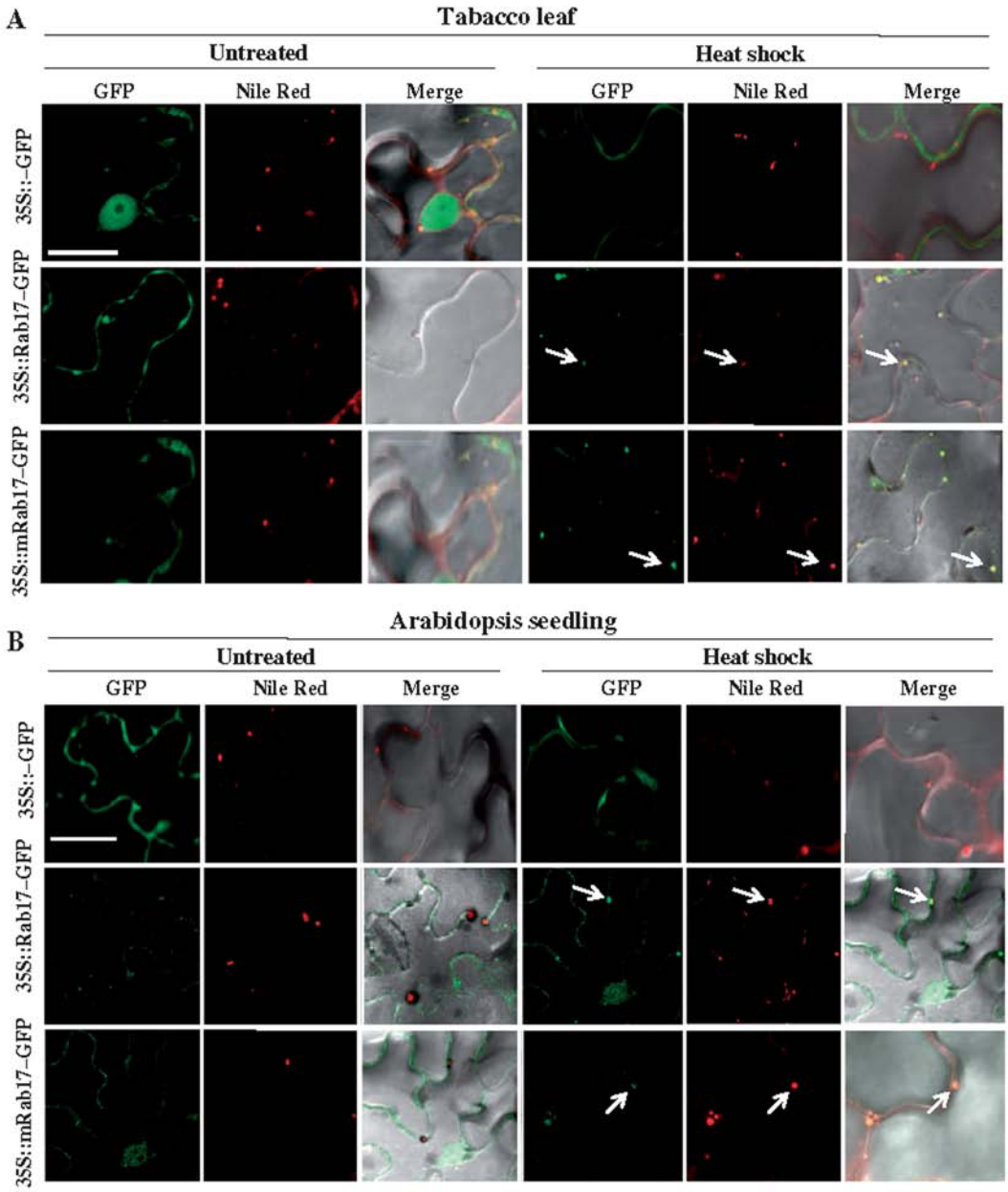


Figure 6. Rab17-GFP fusion protein colocalizes to oil bodies in *N.benthamiana* leaves and in Arabidopsis cotyledons under heat-shock.

A. Confocal images of control GFP, Rab17 and mRab17 fluorescent fusion proteins in tobacco epidermal cells under untreated conditions (left panels) and following heat-shock (right panels). Fluorescence in the green channel (GFP), red channel (Nile red) and overlay images (merge) show the colocalization of the dehydrin proteins to oil bodies after heat-shock. Arrows indicate the position of different oil bodies. Scale bar corresponds to 20 μm. **B**. Confocal images of control GFP, Rab17 and mRab17 fluorescent fusion proteins in transgenic Arabidopsis cotyledons from four days germinated seedlings. Untreated (left panels) and heat shocked (right panels) seedlings. Fluorescence in the green channel (GFP), red channel (Nile Red) and overlay images (merge) show the colocalization of Rab17 and mRab17 proteins to oil bodies following heat-shock. Arrows indicate the position of different oil bodies. Scale bar indicates 10 μm.

To assess the association of Rab17 protein to oil bodies, isolated oil body fractions were coimmunoprecipitated with GFP and oleosin antibodies. The 15 kDa oleosin band was only immunoprecipitated by GFP antibodies in oil body fractions obtained from heat-shocked Arabidopsis transgenic seedlings expressing Rab17-GFP/mRab17-GFP but never from control GFP expressing lines or untreated seedlings (Figure 7C). Overall, these results indicate that under heat-shock Rab17 protein, independently of its phosphorylation status, preserves epidermal tobacco cell viability and colocalizes to leaf and cotyledon oil bodies in tobacco and Arabidopsis plants.

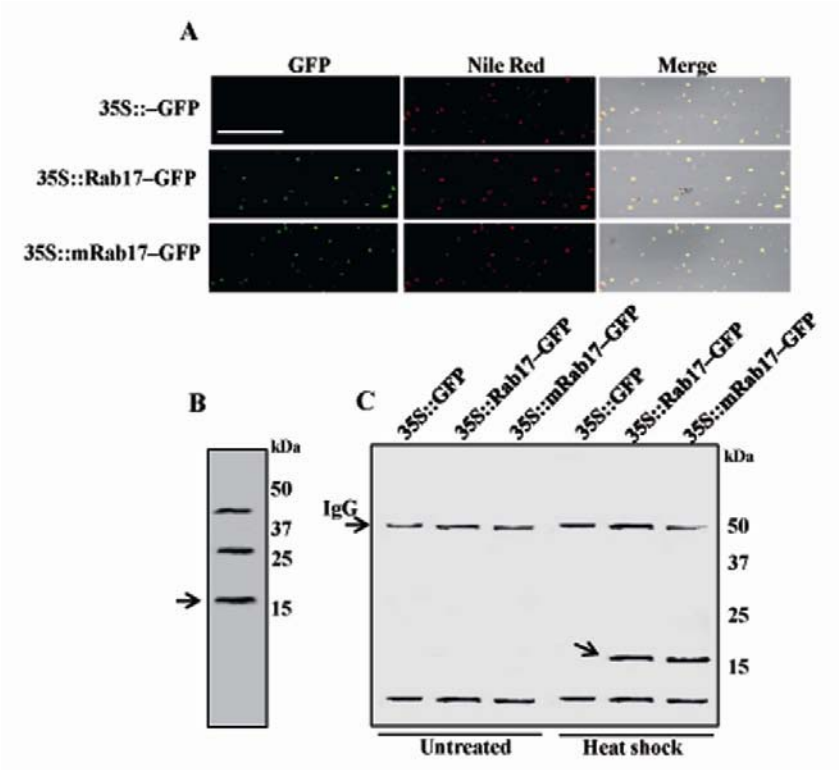


Figure 7. Isolation of oil bodies from Arabidopsis seedlings.

A. Oil body suspensions isolated from four days germinated transgenic Arabidopsis seeds expressing Rab17-GFP, mRab17-GFP and control GFP after heat-shock. GFP fluorescence (green), Nile Red fluorescence (red) and overlay images (merge) show the colocalization of the dehydrin constructs to oil bodies. Scale bar indicates 5 μm. **B**. Immunodetection of oleosin protein in oil body suspensions. Oleosin was detected by western blot analysis using antioleosin antibodies in oil bodies isolated from four days germinated Arabidopsis seedlings. The arrow indicates the position of the 17 kDa size band corresponding to Arabidopsis oleosin. **C**. Coimmunoprecipitation of Rab17-GFP and mRab17-GFP LEA proteins with oleosin protein. Oil body suspensions isolated from untreated and heat-shocked Arabidopsis seedlings overexpressing GFP, Rab17-GFP and mRab17-GFP were immunoprecipitated with GFP antibodies, transferred to nylon membranes and developed with oleosin antibodies. The arrows indicate the position of oleosin protein at 17 KDa and IgG heavy chain at 50 KDa.

4. Discussion

4.1 Maize embryo LEA proteome

During the last stages of seed development, high levels of LEA proteins accumulate as desiccation proceeds. This makes this plant structure especially suitable for LEA proteomic analysis. Here, on the basis of the unusual heat-stability and acid solubility characteristic of unfolded proteins (Cortese et al. 2005; Galea et al. 2006) we have identified twenty unfolded maize embryo proteins, thirteen belonging to the LEA family. Searches in maize databases http://www.maizesequence.org/blast) using the specific group LEA Pfam motifs have yielded an estimation of two group 1, nine group 2 and fourteen group 3 LEA proteins in the maize genome. On this basis, in the present approach we recover the totality of group 1, one member of group 2 and six from group 3 as components of maize embryo proteome. The two group 1 proteins identified, Emb5 and Emb564 were detected in high abundance in mature embryos. By contrast, Rab17 was the only member from group 2 found in maize embryo, although in other maize inbreds the presence of another dehydrin, DHN2, of 40kDa has also been reported in embryo tissues (Asghar et al. 1994). Moreover, in Arabidopsis several dehydrins are non-seed expressed (Hundertmark and Hincha 2008). Interestingly, the LEA group 3 is clearly dominant in maize embryo similarly to what previously found in Arabidopsis seed (Oliveira et al. 2007). The overrepresentation of group 3 in seeds might be indicative of the contribution of this LEA group in seed desiccation.

4.2 Search for posttranslational modifications

In the search for PTMs, the native forms of the selected LEA proteins from groups 1, 2 and 3 reveal different degrees of protein modifications. Group 3 Mlg3LEA is not post-translationally modified, group 2 Rab17 is phosphorylated as previously known (Plana et al. 1991) and group 1 Emb564 displays a complex 2D pattern which combines phosphorylation, acetylation, methylation and deamination. Emb564 phosphorylation is consistent with the previous isolation of group 1 Arabidopsis Atm1 and Atm6 LEA proteins by phosphoaffinity chromatography (Irar et al. 2006). Other group 1 LEA proteins from brine shrimp *Artemia franciscana*, bear multiple sites of high probability phophorylation,

although phosphoproteins were undetectable by Pro-Q Diamond fluorescent gel stain (Warner et al. 2010). In our hands, Pro-Q Diamond stained promptly Rab17 but not Emb564 protein and, in order to show Emb564 phosphorylation, increased sensitivity using ^{32}P labeling and immunoprecipitation was required. In addition to phosphorylation, Emb564 holds several methylations, acetylations and deaminations in different combinations. This is the first report for such PTMs in a group1 LEA protein. In LEAM, a pea mitochondrial LEA protein from group 3, the occurrence of deaminations and oxidations has been proposed to contribute to the functional conformation of the protein (Tolleter et al. 2007). In maize Emb564, several of the observed PTMs are located inside the 20 amino acid conserved motif characteristic of LEA group 1. Despite that in wheat Em protein, removal of group1 conserved repeat did not affect the ability of rEm protein to protect *in vitro* LDH activity from drying, other functional aspects were not evaluated (Gilles et al.). Recombinant proteins used in *in vitro* assays may not hold the PTMs present in native proteins and thus they may not display the complete functional potential of native proteins. PTMs in Emb564 might be critical factors in the protein conformation, localization, turnover and/or interactions with subcellular partners *in planta.*

4.3 Experiments in E.coli: differences in anti-aggregation and protective properties among LEA proteins.

Using the *E. coli* expression system, recombinant Emb564, Rab17 and Mlg3 LEA proteins provided different levels of *in vitro* anti-aggregation activity which correlated with different levels of *E.coli* growth improvement under stress. Differences in growth improvement were also found among three soybean LEA proteins from groups 1, 2 and 3 (Lan et al. 2005). In the present study, the correlation of *in vitro* and *in vivo* activities suggests that prevention of aggregation might be the underlying protective mechanism. In non-stress conditions, expression of maize Emb564, Rab17 or Mlg3 proteins, *per se*, is not deleterious for bacterial cell growth as found for other LEA proteins (Campos et al. 2006). Under stress trials, group 1 Emb564 showed a mild anti-aggregation activity, consistent with the reported lower protective efficiency of group 1 MtEM compared to group 5 PM25 from *Medicago trunculata* during freezing or drying (Boucher et al. 2010). By contrast, group 2 Rab17 was always highly efficient both in preventing denaturation of *E.coli* soluble

proteome and improving bacterial cell growth during desiccation, heat-shock or low temperature. These effects were not dependent on phosphorylation of the S-segment because recombinant Rab17 is not phosphorylated by *E.coli* cells. The general protective effects of Rab17 are consistent with the wide chaperone activity reported for group 2 ERD10 and ERD14 LEA proteins from Arabidopsis (Kovacs et al. 2008). Interestingly, group 3 Mlg3 displayed maximal protective activity under dehydration and osmotic stress but was poorly efficient under heat-shock and low temperature challenges. Different enzyme protective efficacies, depending on how water is removed, by water-loss or freezing, were also observed for Arabidopsis group 4 AtLEA 4-5 and yeast Sip18 proteins (Reyes et al. 2005). In fact, some LEA proteins gain structure during dehydration (Boudet et al. 2006a; Hand et al. 2011; Tolleter et al. 2007) and this gain was found to occur gradually in group 3 AavLEA1 LEA protein, from random coils in solution to mainly a-helical 3D conformation as water content diminished (Li and He 2009). Pea mitochondrial LEAM, also from group 3, adopts α-helical structure upon air drying but surprisingly not upon lyophillization; the authors conclude that the rapid freezing before dehydration prevents the transition to α-helix structure and that LEAM needs to adopt an helical conformation upon drying in order to interact with membranes (Tolleter et al. 2007). Therefore, increased conformation of maize Mlg3 during drying but not under rapid heat or freezing could explain why the high protective activity of Mlg3 was found exclusively associated to dehydration.

4.4 LEA-GFP fusion proteins in plant cells

The GFP-fusions of group 1, 2 and 3 maize LEA proteins were distributed into the soluble compartments of cytosol and nucleus suggesting that their functions take place in these compartments. Maize group 1 Emb564 protein, despite being seed specific, was stable when over-expressed in agroinfiltrated tobacco leaf cells. The expression of Emb564 only in maize embryonic tissues, is consistent with group 1 LEA proteins being preferentially accumulated in embryo development and not in vegetative tissues (Battaglia et al. 2008; Manfre et al. 2009); it has been shown that Atm1 plays a role in the dehydration of Arabidopsis seeds (Manfre et al. 2009). Because embryo native Emb564 protein holds such a complex combination of PMTs it is plausible that these modifications may be relevant for

interaction with seed specific partners. Uncovering the subcellular localization of Emb564 in maize embryogenic tissues will contribute to clarify the specific role of this LEA protein in the drying embryo.

Under osmotic stress, tobacco cells expressing group 3 Mlg3-GFP fusion proteins were larger compared to cells expressing group 1 Emb564-GFP or group 2 Rab17-GFP fusion proteins. This indicates that in the presence of Mlg3-GFP fusion protein cells are able to preserve better their cell volume and cell water content under dehydration conditions. During desiccation cells not only increase the molecular composition of the aqueous compartments, or crowding effect, but change their physical dimensions and shape (Mouillon et al. 2008). In cell shrinkage due to water loss, the rate at which the dehydration occurs in physiological systems is essential for cell adjustments. Interestingly, Mlg3 counteracts the shrinkage effects caused by osmotic treatment in an *in vivo* plant system, and protects bacterial growth during osmotic stress but not during rapid freezing or heat-shock. Using synthetic model peptides of group 3 LEA proteins from insects, plants and nematodes it has been shown the occurrence of reversible structural changes between random coils to α-helical coiled coils in response to water activity (Shimizu et al. 2009). In this line, desiccation-induced folding would be a requisite for Mlg3 functionality. Microscopic observation of fluorescence in Mlg3-GFP expressing cells under osmotic stress, showed a general soluble protein distribution into cytosol and nucleus similar to untreated cells, and did not reveal preference for a specific subcellular structure. Future work directed to identify specific intracellular targets for Mlg3 LEA protein in the process of dehydration under different experimental conditions will surely contribute to define how group 3 Mlg3 LEA protein minimizes the cell shrinkage effects due to water losses.

In relation to group 2 Rab17 protein, here we provide evidence for the *in vivo* colocalization of Rab17 protein to leaf oil bodies under stress. Colocalization was clearly observed in heat-shocked and to a lesser extend in osmotically stressed leaves, but never in non stressed leaves. Binding of Rab17 to leaf oil bodies during heat-shock makes this organelle one biological target for Rab17 LEA protein under stress. In higher plants, oil bodies are mostly found in seeds where they have a role as food reserves for germination and post-germination growth (Huang et al. 2009).They accumulate triglycerides which are

surrounded by a phospholipid monolayer overlaid with oil-body membrane proteins, oleosin, caleosin and stereosin. In Arabidopsis oleosins modulate the size and prevent the coalescence of oil bodies during seed desiccation and its suppression produces unusual oil bodies (Siloto et al. 2006). Oil bodies are also synthesized by leaves with similar morphology but different fatty acid composition suggesting different functions (Lersten et al. 2006), (Huang et al. 2009). In Arabidopsis the expression of early stress responsive gene RD20 from the oil body caleosin family increases in water stressed-leaves and with ABA, similarly to dehydrin Rab18. Since *rd20* knock-out plants exhibited reduced water deficit tolerance a role in drought tolerance was proposed (Aubert et al. 2010). Other authors have suggested alternative roles for oil bodies in adaptation to cold temperatures (Lersten et al. 2006). Hence, a link between oil bodies and associated proteins and plant stress responses is emerging. We never noticed the association of Rab17 protein to lipid bodies in maize embryos before. Possibly the use of organic solvents in microscopic sample preparation and routine discard of top oil layers in conventional seed protein extraction procedures may have contributed to missing the Rab17-oil body interaction. In fact, wheat WCOR414 and Arabidopsis LTi29 acidic dehydrins were immunodetected in the vicinity of the plasma membrane during cold acclimation in aerial tissues (Danyluk et al. 1998; Puhakainen et al. 2004). And *in vitro* studies support binding of Arabidopsis dehydrins ERD10 and ERD14, and maize DHN (Rab17) to acidic phospholipid vesicles (Kovacs et al. 2008); Koag et al. 2003); binding and gain in conformation were attributed to the K-segment (Koag et al. 2003; Koag et al. 2009). Recently, the contribution of dehydrin K-segments from group 2 Lti30 LEA protein and the relevance of the flanking His side chains as regulators of the interaction between the K-segments and membranes in a pH-dependent manner has been shown (Eriksson et al. 2011). In the present study, we show that the interaction dehydrin-lipid vesicle is triggered by stress in a living cell. The nature of the interaction is not as loose as might be expected since binding of Rab17 to oil bodies is strong enough to resist coimmunoprecipitation approaches. However, absence of colocalization to oil bodies in non-stressed conditions opens the question of whether the binding is restricted to new specific stress-induced oil bodies, to stress-induced conformational changes in Rab17 structure previous to the binding or it is dependent on a yet unidentified intermediate element.

LEA proteins, as unfolded proteins, are more likely to be multifunctional than globular proteins and may perform more than one function (Tompa and Kovacs 2010; Tunnacliffe et al. 2005). In addition to oil body binding, Rab17 is highly efficient in preventing protein aggregation by stress denaturation in *E.coli* and improving cell viability in tobacco epidermal cells under heat shock; these properties may contribute to the general stabilization of the cellular proteome. The functionality of phosphorylation is a recurring issue in dehydrins with S-segments. In dehydrins ERD10, ERD14 and COR47, phosphorylation determines *in vitro* ion binding capacity and chaperone calcium-dependent function (Alsheikh et al. 2003; Heyen et al. 2002). Here we observe that, *in vivo,* phosphorylation of Rab17 is not involved in the protein anti-aggregation properties, in consistency with *in vitro* reports (Koag et al. 2009; Kovacs et al. 2008) or in the interaction with lipid bodies We know from previous studies that phosphorylation of Rab17 promotes the arrest of seed germination under salinity (Riera et al. 2004) and that Rab17 phosphorylation enhances protein *in vitro* self-association (M.Figueras, unpublished results). A model for phosphorylation-controlled dimerization has been proposed for CDcT11-24, a dehydration and ABA responsive protein, as a system to stabilize coil-coil interactions with itself or other proteins, creating stabilized networks (Rohrig et al. 2006). Recently, phosphorylation of two dehydrins from *Thellungiella salsuginea* has been shown to modulate *in vitro* actin polymerization dynamics (Rahman et al. 2011). Thus, phosphorylation of Rab17, although not essential for oil body binding or anti-aggregation activity, may participate or modulate interactions with other intracellular partners, perhaps more specifically in seeds where the complete water loss endured by this tissue may enable protein conformations and cellular interactions which may never occur in transient dehydrations in vegetative tissues.

5. Conclusions

We present evidences which illustrate differences among the LEA representatives of group 1, 2 and 3 in maize. Group 1 Emb564 stands out for displaying a complex combination of different post-translationally modifications (PTMs), including phosphorylation, acetylation, methylation and deamination in the native protein, which may be relevant for its seed specific role. Group 3 Mlg3 shows high anti-aggregative properties specifically associated to dehydration conditions which, in agroinfiltrated tobacco epidermal cells, results in a reduction of cell shrinkage effects due to cell water losses. And finally, group 2 Rab17 which displays a general potent anti-aggregation activity that prevents protein denaturation under stress, also binds to leaf oil bodies under heat shock *in vivo*, supporting a role for dehydrins in membrane protection.

CHAPTER 3

V. Chapter 3

Chapter 3

Characterization of transgenic maize plants over-expressing *Rab28* LEA gene

Abstract

LEA proteins (Late Embryogenesis Abundant) participate in plant stress responses. In this report Rab28 LEA gene has been over-expressed in maize plants to test the stress tolerance of the transgenic lines under osmotic stress. Expression of Rab28 transcripts was driven under the constitutive maize *ubiquitin* promoter and led to the accumulation and stability of Rab28 protein in vegetative transgenic tissues. Transgenic lines were analyzed and resulted in sustained growth under PEG (Polyethyleneglycol)-mediated dehydration in relation to wild-type controls. Under osmotic stress transgenic seedlings showed increased leaf and root areas, higher Relative Water Content (RWC), reduced chlorophyll loss and lower Malondialdehyde (MDA) production compared to wild-type plants. Under water deficit transgenic seeds also exhibited higher germination rates than wild-type seeds. In wild-type maize seeds Rab28 protein was found distributed in all mature dried embryonic tissues and immunolocalized to the nucleolus of cells. Fractionation of isolated embryo nuclei indicated that the protein is preferentially recovered in the S2 extract or ribonucleoprotein fraction. In transgenic maize root cells, wholemount immunocytochemistry revealed that Rab28 also occurs in nucleoli in vegetative tissues. Overall, our results highlight the presence of Rab28 protein in nucleolar structures and point to the potential of group 5 LEA *Rab28* gene as candidate to enhance stress tolerance in maize plants.

1. INTRODUCTION

Environmental stresses affect most areas of the world and impose great limitations to crop productivity (Ahuja et al. 2010; Boyer 1982). The search for drought resistant crops has become a challenging task for crop scientists. Conventional breeding for drought resistance has been a basic approach for a long time and improvements have been achieved in several crops (Blum 2011; Yang et al. 2010). Genetic engineering has now the potential to further improve plant growth and crop productivity by introducing selectively genes involved in stress resistance traits. Although genetic differences exist among plant species regarding stress tolerance, crops are usually very sensitive to water stress. Maize (*Zea mays*) is an annual crop cultivated worldwide which productivity is largely affected by drought (Boyer and Westgate 2004; Campos et al. 2004). Water stress in plants occurs when transpiration demands exceed water absorption by roots. A first response to water stress is a decrease in leaf turgor and reduction of cell elongation, followed by a decrease in protein synthesis and cell division with inhibition of plant growth. Stomatal closure behind loss in turgor is the primary cause for reductions in photosynthetic capacity under water stress (Chaves 1991). At the molecular level, the extensive changes in gene expression induce alterations in the signal transduction pathways and in the proteomic machinery of the cell (Ahuja et al. 2010; Cramer et al. 2011; Qin et al. 2011; Yamaguchi-Shinozaki and Shinozaki 2006). One of the elements participating in the stress response are the Late Embryogenesis Abundant (LEA) proteins which accumulate in plant desiccation tolerant structures such as seeds and in vegetative tissues under environmental stresses like drought, freezing and salinity, or abscisic acid (ABA) treatment. The contribution of LEA proteins in stress protection and acquisition of desiccation tolerance is largely accepted although the precise molecular mechanisms are still not fully understood. Present data support their capacity to stabilize macromolecules and cellular structures, proteins and membranes during drying, especially in the presence of sugars like trehalose (Bartels and Salamini 2001; Chakrabortee et al. 2010). A unifying characteristic of LEA proteins is their high hydrophilicity, high content in glycine and largely unstructured conformation in the hydrated state. These properties contribute to their heat and acid solubility (Amara et al. 2012; Oliveira et al. 2007); However, under drying conditions they can acquire a degree of conformation (Hand et al

2011, Shih et al 2008). In plants LEA proteins are targeted to multiple cellular locations, where they might exert their protective functions (Tunnacliffe and Wise 2007; Hundertmark and Hincha 2008, Amara et al, 2012). Transgenic approaches have shown that over-expression of LEA proteins from different species in Arabidopsis, tobacco, rice, wheat, lettuce or cabbage produces improved abiotic stress resistant phenotypes (Leprince and Buitink 2010)

The LEA family comprises different groups classified on the basis of amino acid sequence similarities and conserved motifs (Battaglia et al. 2008; Bies-Etheve et al. 2008; Wise 2003). The LEA proteins from Group 5 (PFAM 04927 SMP) are, however, distinct from other LEA groups in that they contain a higher proportion of hydrophophic residues and have a rather amphipathic structure (reviewed in Tunnacliffe and Wise 2007, Shih et al. 2008, Leprince and Buitink 2010). This type of unusual LEA proteins was initially described as LEA D-34 in cotton seeds (Baker et al. 1988) and since then homologous proteins have been described in several plant species (Shih et al. 2008). In maize, the group 5 *lea* gene *Rab28* (gi/22459) is expressed along the embryogenesis and it is induced by dehydratation and ABA in vegetative tissues (Pla et al. 1991). It encodes for a 28kD protein which in the dry mature embryos is basically located in the nucleolus (Niogret et al. 1996).

In other to clarify/analyze the contribution of LEA 28 protein in relation to osmotic stress, we generated transgenic *Z.mays* plants over-expressing maize group 5 *lea* gene, *Rab28,* under the maize *ubiquitin* promoter. Accumulation of *Rab28* transcripts and protein, phenotype and stress tolerance traits of the transgenic plants were investigated. Germination rates and stress-related physiological parameters were evaluated and compared to wild type plantlets under water-deficit conditions. The results indicate an improvement in water deficit resistance of transgenic seedlings. In addition, and using immunocytochemical approaches, the subcellular localization of transgenic and wild type Rab28 proteins was examined in root cells and found to be nucleolar as in embryonic tissues.

2. Materials and methods

2.1 Isolation and plasmid construction of Rab28 gene

A full length cDNA encoding the Rab28 was synthesized from RNA of mature embryo of maize. Rab28 cDNA sequence was amplified by PCR using oligonucleotide primers (Table 1). Sequences were cloned into pCR2.1-TOPO vector, cut out and subsequently ligated into pAHC25 plasmid under the control of the maize *ubiquitin* (*ubi*) promoter (Christensen and Quail 1996). *Bar* gene was used as the plant selection marker.

2.2 Generation of Rab28 transgenic maize

Immature zygotic embryos (1.5–2.0 mm long) of maize Hi-II genotype (A188 x B73) (Armstrong et al. 1992) were cultured axis-side down on N6- based medium (Wang and Frame 2009) supplemented with 25μM silver nitrate, 100 mg l/1 casein hydrolysate, 25 mM L-proline, 2 mg l/12,4-dichlorophenoxy acetic acid, 30 g l/1 sucrose and 2.5 g l/1Gelrite (Sigma) at pH 5.8 Particle bombardment of Hi-II callus was carried out with pAHC25 carrying *Rab28* gene using a Bio Rad 1000/He biolistic system, 650-psrupture disks and 0.6μm gold particles (Wang and Frame 2009). The selection medium was N6-based medium without proline and casein hydrolysate, with 2 mg /1 bialaphos added. Calli were transferred every 2 weeks to fresh selection media. After 3–4 cycles of selection, white, rapidly growing callus clusters were picked out from non-proliferating mother calli and were transferred into fresh selection medium.

For plant regeneration, pieces of callus were transferred to a modified regeneration MS medium with 60g/l sucrose and 3mg/l bialaphos (Murashige and Skoog 1962) and incubated at 28ºC for 14 days in the dark. After this maturation period the opaque and white areas with clearly defined scutellar region were transferred to standard MS medium, and allowed to germinate at 25ºC in the light (80-100μE/m^2/s light intensity), 16h light/8h dark photoperiod. *In vitro* rooted plants were transferred to the greenhouse for acclimation and further growth till maturity. Transgenic plants were regenerated from different transformation events. Mature plants from T0 generation were self-pollinated to generate T1 seeds. T1 plants represent a mixed population containing wild-type, homozygous and

heterozygous transgenic plants. T1 transgenic plants for *rab28* were either self-pollinated or crossed with wild-type plants to generate T2 plants. After identification of *rab28* homozygous and heterozygous plants, three T2 lines (heterozygous C1, homozygous C6 and C4) were selected for physiological studies.

2.3 Molecular analysis of transgenic plants

To select for transgenic plants in T1 and T2 generations, total DNA from various independent transgenic lines was extracted according to Doyle (Doyle and Doyle 1990). PCR reactions to detect *Rab28* gene were performed with *Ubi* promoter forward and RAb28 reverse primers to amplify a 1.034 kbp fragment (Table1). Untransformed plants were used as negative control and plasmid p*AHC25*::*ubi*::*Rab28* as positive control. Expression levels of *Rab28* were detected by RT-PCR (Shou et al., 2004a) in a Roche Light Cycler 480 instrument using SYBR Green I dye and the primers listed in Table1. Primers were designed using Primer Express software (Applied Biosystems,Foster City, CA, USA). Data were normalized with *Z. mays* cyclophilin (*ZmCyp*) as internal controls for maize.

Table1: table of primer

Oligonucleotides used in this study		
Primer name	Sequence	Use (restriction site introduced)
SmaI Rab28	5-CCCGGGATGAGCCAGGAGCCAGCCGAGG-3	Forward primer for SamI
Sac I RAb28	5-AGTACTTCACTCCCAGACCGCCTCGTT-3	Reverse primer for SaCI
pUBi fe	5-CCTTCATACGCTATTTATTTGC-3	Forward primer for PCR
Rab28 rev	5-AGGTTCTCCGCCGACTGCATC-3	Reverse primer for PCR
RAb28 few	5-AGTTGCAGGTAAGGATATCACAAAGG-3	Forward primer of q-PCR
Rab28 rev	5-GCCTCGAGCGCCTCGCATCGT-3	Reverse primer of q-PCR
Cyclo few	5-GTGTGGATCTGTGAACCCCAT-3	Forward primer of q-PCR
Cyclo rev	5-CAGGTGAAACACGAATCAAGC-3	Reverse primer of q-PCR
Actin f	5-CCAAGGCCAACAGAGAGAAA-3	Forward primer of RT-PCR
Actin r	5-GATACCGGAGCCGGACTCCTCCGCCTC-3	Reverse primer of RT-PCR

Copy number determination in transgenic plants containing p*AHC25*::*ubi*::*Rab28* plasmid and in wild type control, was performed by qRT-PCR (Omar et al. 2008).

Genomic DNA was extracted and purified from 100 mg of young leaves (Nucleon Phytopure kit, Amersham). Estimations on exogenous *Rab28* gene copy numbers in

transgenic lines were performed by comparison with wild-type as endogenous reference according to Preuten et al. (Preuten et al. 2010).

2.4 Protein extraction, western blot and LEA antibodies

Protein samples were obtained from transgenic and wild type plants by cutting a piece of the third leaf (about 100 mg). Proteins were extracted in salt buffer (20mM TES-OHK, pH 8.0, 0.5M NaCl) as described (Amara et al. 2012). Salt-soluble proteins were heated at 100 °C for 10 min, cooled for 30 min and heat-soluble proteins obtained in the supernantant. Protein samples (20 μg), were analyzed by 15% SDS-PAGE gel electrophoresis and blotted onto nitrocellulose filters. Immunoblots were developed with antibodies against maize LEA proteins group 5 Rab28 (Niogret et al. 1996), group 2 Rab17 (Figueras et al. 2004) and group 3 Mlg3 (Amara et al. 2012) using the ECL detection system (Amersham Pharmacia Biotech). Heat-soluble and insoluble proteins from pollen grains and anthers of wild type plants were obtained by the same extraction procedure. Gels were visualized by silver staining (Shevchenko et al. 1996).

2.5 Stress treatments and measurement of physiological parameters

Growth and stress treatments were performed on hydroponically grown plants as previously described (Amara et al. 2012) . Osmotic stress treatments were performed on 10-days-old plants, by adding 15% PEG-8000 to MS culture medium without sucrose, during 5 days. For germination assays, 20 seeds were germinated on 15% PEG containing MS-0 medium at 25ºC in growth chamber under 16h light/8h dark period and 60% relative humidity. One week after, the percentage of germination was determined by counting radicle emergence. Total leaf area of primary leaves from maize seedlings was determined after scanning the different leaves and analyzing the data by UTHSCSA software (htpp://ddsdx.uthscsa.eedu/dig/itdesc.html). Chlorophyll content of expanded leaves was measured by extraction of total chlorophyll from two leaf discs per plant and determined by spectrophotometry (Moran 1982). Relative water content (RWC) and fresh weight (FW) values were measured on hydrated cut leaves; dry weights (DW) were obtained after oven-drying leaf samples for 72h at 70ºC. RWC was calculated by the equation of Schonfeld (Schonfeld et al. 1988). Root area was determined using WinRHIZO software (V5.0, Regent Instruments, Quebec, Canada) and Epson Perfection V700 modified flatbed scanner

according to Watt et al. (Watt et al. 2005). Lipid peroxidation was measured in leaves by determining Malondialdehyde (MDA) content (Quan 2004). All determinations were performed on transgenic lines and wild type controls, grown under optimal conditions or under 15% PEG osmotic challenge. Each experiment was replicated four times. All data were presented as mean ± standard error (SE). Comparisons between transgenic plants and WT plants were performed using Student's test. Values of $0.01<P<0.05$ were considered significant.

2.6 Nuclei isolation and fractionation

Nuclei-enriched glycerol fractions were isolated from mature maize embryos and nuclei were further purified by flotation on Percoll gradients according to Goday et al. 1994). Nuclei fractionation was performed following the procedure of De Cárcer et al. 1997.

2.7 Immunocytochemistry

Rab28 was immunolocalized in wild type maize mature embryos using thin sections of Lowicryl embedded tissue as described (Niogret et al. 1996). Basically, 2 μm sections were incubated for one hour with Rab28 antibodies (1:1000), followed by Protein A-gold (10 nm, 1:50). Gold label was enhanced by exposure to a photographic developer containing silver nitrate (Marchetti et al. 1987). Sections were stained with 1% methylene blue.

In maize root tips Rab28 protein was localized by whole-mount *in situ* immunolocalization as described for Arabidopsis roots (Sauer et al. 2006) with minor modifications After root fixation with 4% paraformadehyde, cell wall digestions were performed for with 3% Driselae (Sigma D9515) in PBS for 2 hours, instead of 2% for 1 hour. Time of permeabilization with 3% IGEPAL CA-630 (Sigma I3021) and 10% DMSO (Dimethylsulphate) was increased to 3 hours. Blockcage with 3% albumin fraction V (Sigma) was followed by incubation with Rab28 antibody or preimnume serum (1:80) as primary antibodies and FITC-conjugated anti-rabbit antibodies anti rabitt, whole molecule (Sigma F9887) (1:1000) as secondary probe. Fluorescence was observed by confocal microscopy at excitation 488 nm/emission 500–600 nm (Olympus FV 1000). Cell nuclei were visualized by Dapi staining (25 μg/ml) (Sigma D9564) .

3. Results

3.1 Molecular analysis of transgenic maize plants

Plantlets regenerated from independent events (T0 generation) were grown to maturity in the greenhouse and were analyzed by PCR and western blotting using a serum directed against Rab28 protein (Figure 1A, 1B). T1 progeny was obtained by self-pollination of T0 generation; homozygous T2 progeny was obtained by self-pollination of each transgenic T1. We selected one heterozygous line (C1) and two homozygous lines (C6 and C4) for further physiological and molecular studies. Determination of the relative copy number of *Rab28* transgene in each transgenic line compared to wild-type indicated an estimation of a single copy in heterozygous C1 line and, for homozygous C6 and C4 plants, 2 and 4 copies respectively (Figure 1C).

The expression of Rab28 mRNA in the three independent transgenic lines and wild-type controls under well-watered conditions was analyzed in leaves by semi-quantitative PCR. Higher levels of Rab28 mRNA were found in homozygous C6 and C4 lines compared to the heterozygous C1, with absence of Rab28 expression in the wild-type controls (Figure 1D). The accumulation of the corresponding protein rab28 follows identical pattern, in complete accordance to mRNA levels (Figure 1E).

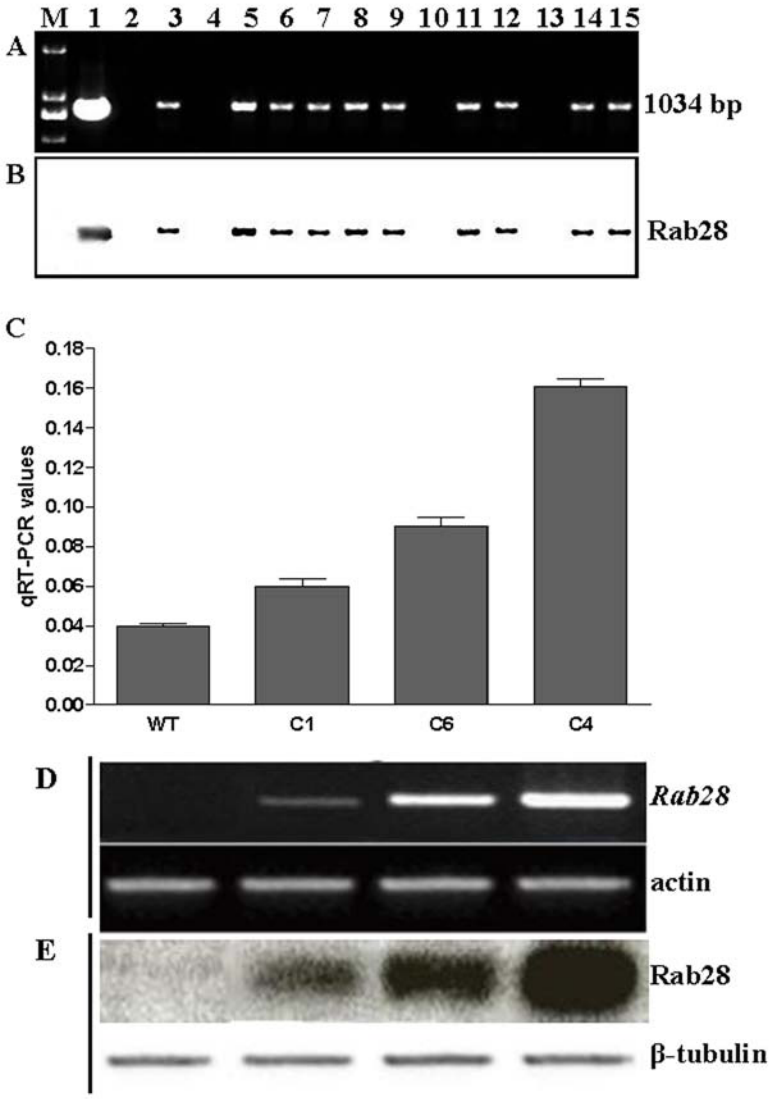


Figure 1: Analysis of Rab28 gene expression and protein accumulation in maize transgenic lines.

A. PCR analysis performed to test for the integration of 1.034kbp *Rab28* in the genomic DNA of transformants from T0 generation. pACH25 Rab28 plasmid positive control (lane 1), PCR products from DNA of wild-type leaves as the negative control (lane 2) and from DNA of leaves of Rab28 transgenic lines (lanes 3-15). M corresponds to the bp marker.**B.** Immunoblot blot performed on transgenic maize leaves to test for the presence of Rab28 protein. Protein extracts from mature maize embryos (lane 1), protein extracts from non-transformed wild type leaves (lane 2) and from transgenic leaves corresponding to the above plants (lanes 3-15). Position of Rab28 protein is indicated**.C.** Selection of three transgenic maize lines: heterozygous C1 and homozygous C6 and C4. Estimation of the relative transgene copy number in the T2 generation by qRT-PCR analysis using as reference the endogenous *Rab28* DNA levels in wild-type (WT) leaves .**D.** *Rab28* mRNA levels in transgenic maize leaves under optimal growth conditions measured by RT-PCR. Below, actin is shown as internal control. **E**. Rab28 protein in transgenic maize leaves. Immunoblot of protein extracts (15μg) developed with Rab28 polyclonal antibodies. Below, β-tubulin is shown as loading reference.

3.2 Phenotypic traits of Rab28 Transgenic Plants

Stable transformation of maize plants with *Rab28* gene did not affect the overall plant morphology. Under optimal greenhouse growth conditions traits such as plant morphology, plant height, leaf number, leaf color, length of primary leaf and roots, kernel appearance and germination rates showed no significant differences between the independent transgenic lines C1, C6 and C4 and wild-type plants. However, a significant reduction in the number of seeds produced by transgenic lines was evidenced; the average grain yield per plant in T2 generation was higher in wild-type plants compared to transgenic lines and homozygous C6 and C4 plants showed lower grain production than the heterozygous C1 plant (Table 2). This yield penalty, however, did not affect the average weight of seeds produced since no significant differences in grain weight between transgenic and wild-type seeds were observed (Table 2). In a few plants a delayed silking in relation to pollen maturation was also observed.

Table2. Seed yield and seed weight in wild-type and Rab28 transgenic lines under greenhouse growth conditions

Lines	Number of seeds/plant*	Average seed weight (g)**
WT	235	2.25±0,04
C1	123	2.33±0,02
C6	62	2.29-±0,05
C4	52	2.22± 0,03

* Values represent the number of seeds obtained in T2 plants
**Values are mean ±SE (n=10 seeds) seed weight in T2 generation

Because LEA type proteins accumulate in desiccation tolerant structures, the presence of Rab28 protein was then examined in mature dry pollen and anthers from wild type plants. Heat-soluble and heat-insoluble protein extracts from both tissues were analyzed by immunobloting. Rab28 protein which was found in thermostable fraction of mature maize embryos was, however, absent from anthers and mature pollen grains. This result indicated that Rab28 LEA protein is not a natural constituent of maize pollen proteome (Figure 2).

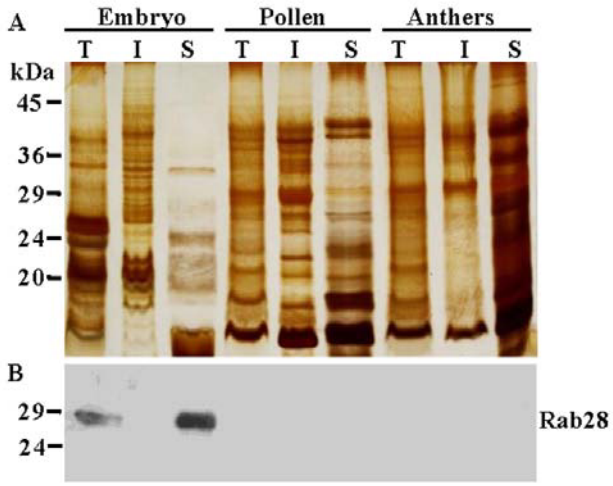


Figure 2. Absence of Rab28 LEA protein in wild-type maize pollen grains.

A. SDS gel electrophoresis of protein extracts from mature embryo, mature pollen grains and anthers from wild type maize plants. Proteins are visualized with silver staining. Protein fractions correspond to total protein extracts (T), heat-insoluble (I) and heat-soluble (S) protein extracts. The position of protein size markers is shown on the left. **B**. Western blotting of the above samples using antibodies against Rab28 protein. The position of Rab28 protein is indicated.

3.3 Overexpression of Rab28 enhances resistance to osmotic stress in transgenic maize plants

Since LEA *Rab28* transcripts are readily induced in wild-type maize leaves by drought stress (Niogret et al 1996), the effects of osmotic treatment were investigated in Rab28 transgenic maize plants. At whole plant level and under optimal growth conditions no significant differences were found between leaf and root areas of transgenic and wild type plants (Figure 3). However, under water-deficit conditions imposed by15% PEG a significant reduction in the growth of wild-type plants compared to transgenic plants was evidenced (Figure 3A). In addition to reduced size, wild-type plants exhibited also increased leaf rolling. These symptoms were delayed and attenuated in *Rab28* transgenic plants.

Under PEG-stress conditions, the levels of *Rab28* transcripts were examined in leaves in wild-type and transgenic plants. In well irrigated wild-type plants, *Rab28* expression was absent in leaves but induced by water deficit. However, the levels of Rab28 expression were always lower than the levels of Rab28 expression observed in transgenic lines (Figure 3B). At protein level, Rab28 was accumulated by all plant leaves, but slightly higher levels were found in homozygous C6 and C4 transgenic lines. The effectiveness of the PEG stress

treatment was assessed checking for the presence of 2 (gi126632258) and Mlg3 LEA protein from group 3 (gi7387829) as stress markers since both proteins are readily accumulated in maize vegetative tissues under abiotic stress (Amara et al 2012). Using specific antibodies against Rab17 and Mlg3 LEA proteins both proteins were found to be induced by PEG stress treatment in wild-type and transgenic lines (Figure 3C); interestingly, wild-type leaves accumulated slightly higher amounts compared to transgenic leaves. Overall, these results indicate that sustained Rab28 levels in the transgenic lines contribute to maintained plant growth under water-deficit conditions.

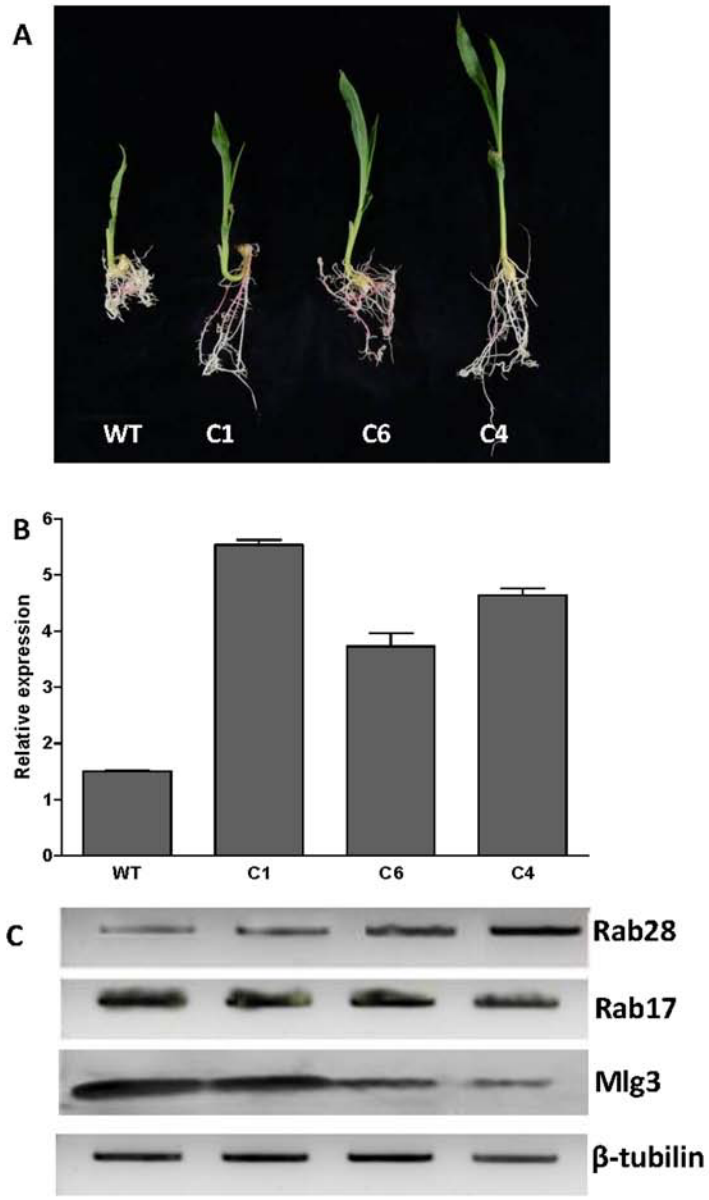


Figure 3. Effect of osmotic stress on transgenic Rab28 plants.

A. Growth inhibition in wild-type (WT), heterozygous (C1) and homozygous (C4, C6) plants caused by continuous irrigation with 15% PEG. Photograph showing the differences in plant size was depicted after five days treatment. **B**. Quantification of *Rab28* over-expression in transgenic maize lines by RT-PCR. The x-axis represents the expression level of *Rab28* transcripts in transgenic plants. Values are mean ±SE of triplicate measurements. **C**. Presence of Rab28 protein in stress-treated plants. Immunoblot of leaf protein extracts (15µg) developed with RAB28 antibodies. Below, immunoblots developed with Rab17 and Mlg3 LEA antibodies used as stress markers. β-tubulin was used as loading reference.

3.4 Physiological parameters in Rab28 transgenic plants under water stress

Because Rab28 transgenic maize plants showed enhanced resistance to osmotic challenge a series of physiological parameters relevant to water stress damage were examined in transgenic and wild-type plants under optimal or PEG growth conditions. Germination assays indicated that under standard growth conditions, germination rates of both Rab28 transgenic lines and wild type plants were similar and above 98% (Figure 4A). In the presence of 15% PEG, germination rates decreased sharply to 25% in wild-type seeds, whereas the germination capability of the transgenic lines, although reduced by the osmotic treatment, was maintained above 64% (fig 4A). Thus, Rab28 transgenic seeds show increased germination rates under conditions of water restrictions. Measurements of total leaf area (TLA) indicated that whereas under standard growth conditions all plants showed similar TLA values, PEG treatment decreased TLA values by 70% in wild-type plants, and 24%, 33% and 36% in C1, C6 and C4 transgenic lines, respectively (Figure 4B). The relative water content (RWC) in leaves of transgenic and wild-type plants ranged between 89-92% under optimal growth conditions; with PEG treatment, significantly higher water losses were detected in wild-type plants which RWC values dropped to 30%, whereas the minimum recorded RWC for transgenic lines was 65% (Figure 4C). Total leaf chlorophyll levels were similar in transgenic and wild-type plants under normal irrigation conditions; under water deficit, chlorophyll content in wild-type maize leaves was significantly decreased to 41% whereas in transgenic plants chlorophyll levels were above 68% (Figure 4D). In relation to root growth, measurements of root area also showed increased root growth of transgenic lines compared to wild type plants (Figure 4E). Collectively, all these measurements are indicators of the increased capacity of Rab28 transgenic plants for maintaining growth under water deprivation conditions imposed by PEG.

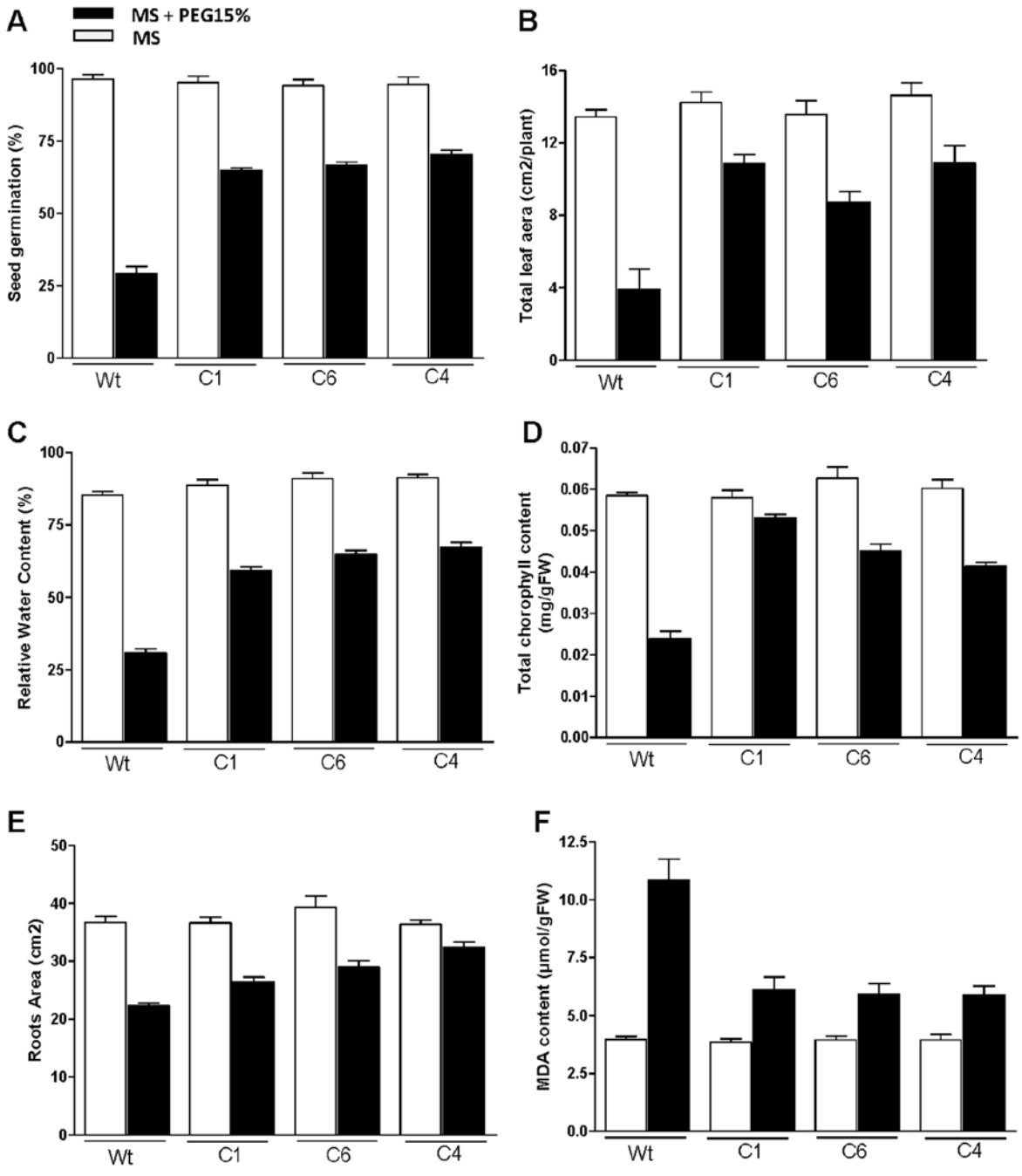


Figure 4. Germination and physiological parameters in transgenic maize plants under osmotic stress.

Germination rates of Rab28 transgenic lines and wild type controls plants under optimal conditions and in the presence of 15% PEG (A). Comparison of total leaf area (B), relative water content RWC (C), chlorophyll content (D), total root area (E) and MDA levels (F) in Rab28 transgenic plants and in wild-type controls measured in untreated controls and in 7-days old transgenic plantlets treated with 15% PEG during five days

Oxidative stress is one major cause of damage following abiotic stress. Here, the changes in lipid hydroperoxide production induced by oxidative stress were measured in maize leaf disks by determining the endogenous malondialdehyde (MDA) content. MDA levels in transgenic and control lines were similar under optimal growth conditions (Figure 4F). In general, MDA contents increase with drought and they are indicative of oxidative lipid injury. After 15% PEG treatment, MDA content in the three transgenic lines was increased by 1.5 fold, in contrast to the 2.7 fold increase scored in wild-type plants (Figure 4F). These results show that lipid peroxidation is reduced in Rab28 transgenic lines compared to control lines during water stress.

3.5 Nucleolar localization of Rab28 protein in dry embryo cells and in transgenic maize roots

In a previous work, Rab28 protein was found to be localized in the nucleolus of scutellar cells in wild-type mature maize embryos (Niogret et al 1996). Here, we confirm and extend this observation to other embryo tissues. By optical microscopy on semithin sections the nucleolar localization of Rab28 was evidenced not only in scutellar cells but also in the remaining embryo cell types, including axial, vascular and aleurone cells (Figure 5 A,B,C). In order to further characterize the subcellular location of Rab28, we performed a stepwise fractionation procedure using isolated nuclear suspensions from mature maize embryos and analyzed for the presence of Rab28 in the different nuclear fractions obtained. The fractionation procedure is based on protein solubility in non-ionic detergents and in buffers of increasing ionic strength (De Cárcer et al 1997). Rab28 is found mostly concentrated in the S2 fraction, which contains nuclear proteins soluble in low-ionic strength (Figure 5D).

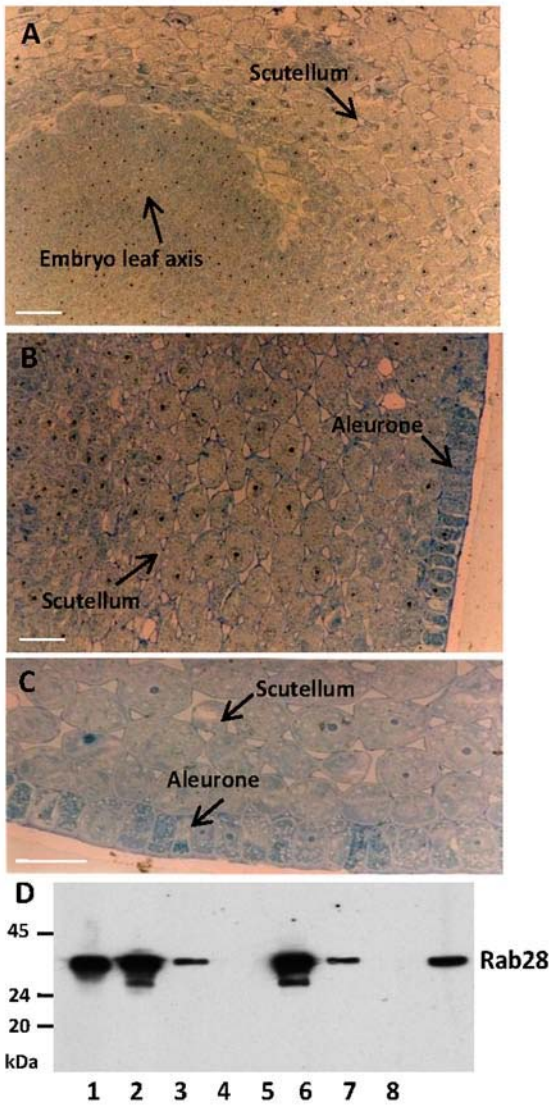


Figure 5. Subcellular distribution of Rab28 protein in wild-type mature maize embryos.

(A,B) Immunolocalization of Rab28 protein in dry maize embryo cells. Light field micrographs of transversal embryo sections. General nucleolar staining in sections of Lowicryl-embedded tissue probed with antiRab28 antibodies. Stained nucleoli are distributed throughout embryo axis, scutellum and aleurone cells. Scale bar corresponds to 20 μm. (C) Bright field micrograph of non-immune sera. Sections are counterstained with 1% methylene blue. (D) Western blotting of dry embryo nuclear extracts fractions from mature maize embryos probed with Rab28 antibodies. Fractions correspond to total initial homogenate (lane 1), nuclear-enriched pellet (lane 2), nuclear-layer obtained by flotation on a Percoll gradient (3), nuclear membranes soluble in 1% NP-40 and 0.5% DOC (lane 4), S2 fraction soluble in low-ionic strength buffer, 10 mMTris and 1mM EDTA (lane 5), chromatin fraction extracted with 0.1 mg/ml DNAse and 25M ammonium sulphate (lane 6), high-salt fraction, soluble in 2M NaCl (lane 7) and insoluble nuclear matrix (lane 8). The strongest reaction is obtained in the S2 extracts

In Rab28 maize transgenic plants, the subcellular location of Rab28 protein was investigated in roots by whole-mount *in situ* immunocytochemistry (Figure 6). In root tips, nuclei of meristematic cells occupy a high relative cellular volume with visible nucleolus which is enlarged in size in the dividing cells (González-Camacho and Medina 2006). The distribution pattern of Rab28 protein in root cells of maize transgenic lines was analyzed and compared with that of wild-type plants grown under optimal growth conditions and PEG stress. Fluorescent images indicated that in non- stress conditions, Rab28 is localized to the nucleolus in most transgenic root cells (Figure 6). Wild-type plants do not synthesize Rab28 protein under well-watered growth conditions. In the cytoplasm, the scattered fluorescence observed in most cells is due to background staining and it was also evidenced in samples incubated with pre-immune serum (not shown). Under PEG-stress no visible changes in the subcellular distribution of fluorescence could be observed and Rab28 nucleolar staining was maintained in the transgenic cells. Identical fluorescent patterns were observed in homozygous and heterozygous transgenic root cells. In wild-type roots, nucleolar staining was only visible after osmotic stress challenge (Figure 6) in accordance to Rab28 transcripts being absent from well-watered tissues and readily induced under drought (Niogret et al 1996). Collectively, these images indicate that Rab28 LEA protein in maize is delivered into the cell nucleolus not only in desiccation tolerant embryo cells but also in vegetative tissues under either optimal or under water deficit conditions.

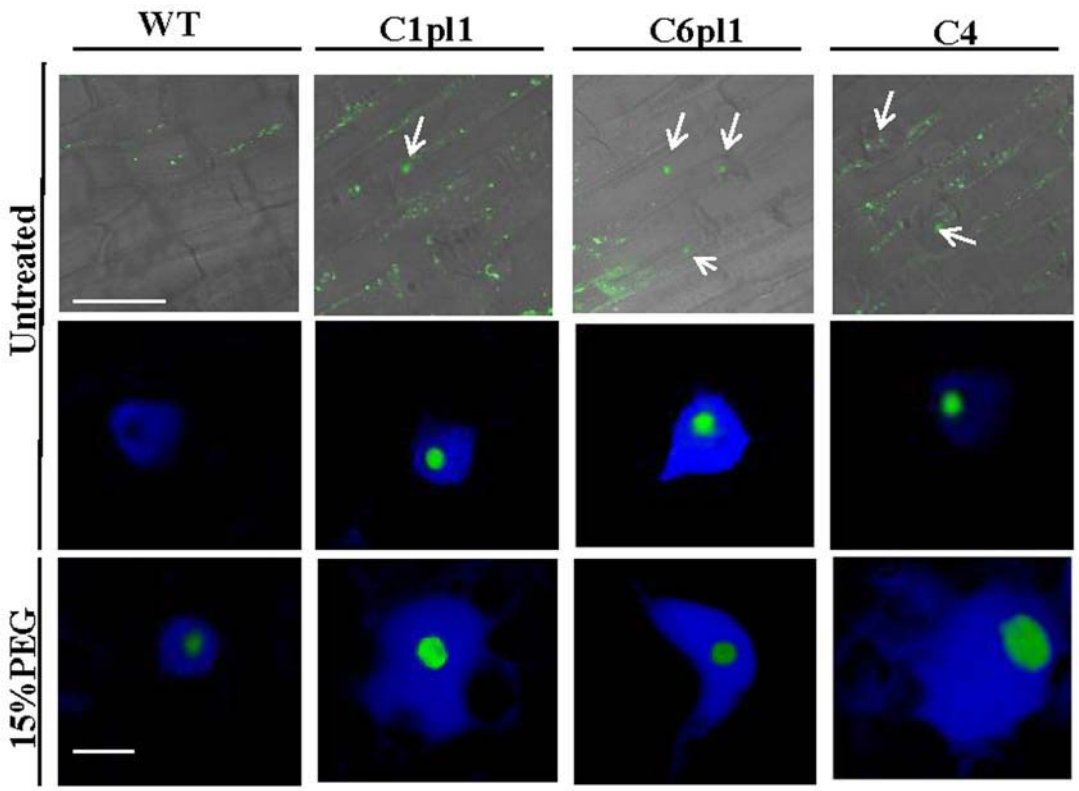


Figure 6: Immunolocalization of Rab 28 protein in transgenic maize roots.

Confocal images showing the fluorescent nucleoli in transgenic cell roots. General view under untreated conditions with arrows indicating the fluorescent stained nucleoli (top panel). Scale bars correspond to 5 µm. Details of nucleus and nucleolus in non-treated conditions (middle panel) and under 15% PEG treatment (low panel). Nucleolar fluorescence (green) corresponds to Rab28 protein; nucleus are visualized by Dapi staining (blue). Scale bars correspond 1 µm

4. Discussion

In the present work we have generated transgenic maize plants over-expressing group 5 *Rab28* LEA gene under the *ubiquitin* constitutive promoter. Previous studies have tested the effect of over-expression of LEA genes from different species of origin in a range of host species (Leprince and Buitink, 2010) but this is the first published report on a LEA gene been transgenically expressed in maize. Rab28 transgenic maize plants exhibited normal development during vegetative growth and showed enhanced resistance under water deficit conditions. The better performance of Rab28 transgenic plants was evidenced in the three independent T2 lines analyzed. All transgenic plants grew faster under PEG treatment than wild-type plants. They showed higher RWC, decreased chlorophyll degradation and higher leaf and root areas compared to wild-type plants. In addition, MDA production which is a marker for oxidative lipid damage remained reduced in the transgenic lines indicating a lower demand for the release of antioxidant species while MDA increased significantly in wild-type plants. Transgenic seeds were also able to maintain higher germination rates than wild-type seeds in the presence of PEG.

The accumulation of other stress-related LEA proteins was used as stress marker in the PEG-treated plants. The presence of group 2 Rab17 and group 3 Mgl3 LEA proteins in all treated plants was an indicator of the degree of stress challenge produced by the PEG treatment. It was interesting to observe that the levels of Rab17 and Mlg3 proteins attained were slightly increased in wild-type compared to homozygous transgenic lines.

Overall, these physiological measurements indicate that transgenic maize plants constitutively expressing Rab28 gene are able to sustain increased plant growth rates under water deficit conditions. The protective effects of Rab28 seem to lower the stress response of the transgenic plant; these results suggest that directly or indirectly Rab28 over-expression in vegetative tissues reduces the stress intensity at the cellular level and thus induces a lower response in the plant which maintains higher growth rates. Similar benefits have been reported in the over-expression of other representatives of LEA groups; in general, these studies demonstrate enhanced tolerance phenotypes, mostly to water deprivation and salinity, and indicate the potential of LEA genes to increase crop stress resistance (Leprince and Buitink 2010; Xiao et al. 2007; Yang et al. 2010).

In a previous work over-expression of maize Rab28 ortholeg, *AtRab28* led to increased germination rate in salt and osmotic conditions and to increased survival of Arabidopsis transgenic plants grown under ionic stress (Borrell et al. 2002). However, in the present study we did not observe increased tolerance of maize Rab28 transgenic plants to LiCl when compared to wild-type controls (not shown). Both genes, maize Rab28 and *AtRab28* are largely expressed in seeds but *AtRab28*, in contrast to Rab28, was never induced by water deficit or ABA in Arabidopsis vegetative tissues (Arenas-Mena et al. 1999). These differences between ortholeg genes may possibly reflect species divergence in promoters and/or protein functions among the two LEA genes.

One unwanted side-effect in Rab28 transgenic plants was the reduction in grain production per plant compared to wild-type plants under normal greenhouse growth conditions; this trait was observed in T2 generation in all transgenic events. Yield penalties have been reported by other authors in Arabidopsis plants expressing drought-related transcription factors (Shinozaki and Yamaguchi-Shinozaki 2007) and in rice over-expressing a LEA protein from group 3 (Xiao et al 2007); the use of appropriate promoters to drive the expression of genes and the selection of plants with single copies of the transgene may prevent these undesired traits (Xiao et al 2007).

Group 5 LEA proteins are considered as a minor group among the LEA family because of their lower hydrophylicity and partially folded nature. It has been predicted that they are not as heat stable as the rest of the LEA proteins (Shih et al 2008, Battaglia 2008). However, here we detected Rab28 in the heat-soluble protein extracts of mature maize embryos. In fact, we had previously identified Rab28 as component of the heat soluble maize seed proteome; however, a distinct feature of group 5 Rab28, in contrast to other more hydrophyllic LEA proteins, was its lower solubility in acid conditions (Amara et al. 2012).

Pollen grains are plant structures naturally resistant to desiccation; however, group 5 Rab28 LEA protein was absent from anthers and mature pollen grains, and this observation was extensive to group 2 Rab17 and group 3 Mlg3 LEA proteins (not shown). In other works, proteomic approaches in Arabidopsis and rice failed to identify LEA type proteins in pollen (Dai et al. 2006; Holmes-Davis et al. 2005) whereas in pollen from *Typha latifolia* and *Brassica napus,* group 3 and LEA-like proteins were reported (Wolkers et al 2001, Sheoran

et al 2009). Thus, and in contrast to seeds, LEA type proteins seem not to be general/constitutive components of pollen grains in all plant species.

It is generally assumed that the LEA family participates in stress tolerance responses although the specific mechanisms are still unclear. Their contribution in stress adaptive responses has been proposed to involve the protection and stabilization of macromolecules and/or cellular structures, through binding or replacement of water, ion sequestration, maintenance of membranes, stabilization and renaturation of unfolded proteins (revised in Tunnacliffe et al 2010, Battaglia et al 2008, Shih et al 2008, Bartels and Salamini 2001). LEA group 5 MtPM25 was found to be associated with desiccation tolerance in radicles in *Medicago truncatula*. MtPM25 protein prevented stress-induced protein aggregation *in vitro*, suggesting a role as protective molecule during drying, and additionally the protein was able to dissolve aggregates in a non-specific manner, consistent with an additional role as repair mechanism (Boudet 2006).

Protein subcellular localization may be an important clue towards understanding the functionality of LEA proteins in their participation in desiccation tolerance. Several group 5 LEA proteins, such as maize Rab28, Arabidopsis AtRab28 and Atecp31, Medicago trunculata PM25, cotton LEAD34 and carrot Ecp31, hold the QPRRP nuclear localization signal (Borrell et al 2002, Boucher et al 2010); however, except for Rab28 and AtRab28, no data is available relative to their subcellular localization. AtRab28 protein was found associated preferentially with embryo provascular tissues (Arenas-Mena et al 1999) and by immunoelectrron microscopy nuclear staining was also detected (Borrell et al 2002). Rab28 was found to be a nucleolar protein in scutellar cells from maize embryo (Niogret et al.1996). In the present study we have confirmed that in the desiccation tolerant cells of mature maize embryos Rab28 is found preferentially associated with nucleolar structures in all embryo tissues, including scutellar, axial, vascular and aleurone cells. In vegetative tissues from transgenic maize plants, we found that Rab28 protein is also accumulated in nucleolus of root cells under optimal and osmotic stress conditions. In wild-type roots, endogenous Rab28 protein synthesized after PEG treatment shows identical nucleolar localization. In the confocal images, the absence of visible changes in fluorescence distribution under optimal and stress conditions, indicates that the nucleolus seems to be the natural subcellular fate for maize Rab28 protein and suggests that Rab28 protective

function may occur in this cellular compartment. Rab28 protein was found preferentially in the nuclear fraction S2 corresponding to the ribonucleoprotein (RNP) fraction which contains the subproteome soluble in low ionic strength buffer, essentially nuclear proteins associated with RNA. In animal and plants cells, S2 fractions contain the 70-80% of RNA and nucleolar proteins involved in the synthesis and processing of ribosomal precursors including RNA polymerase and nucleolar protein kinases (De Cárcer et al. 1997).. We have tried to identify a physiological protein partner by immunoprecipitation approaches using Rab28 antibodies in maize extracts, but failed to identify any clear candidate protein/s in the maize embryo nuclear proteome (unpublished results). Affinity with hypothetical partner/s could be too weak to be identified by immunoprecipitation approaches. Alternatively, absence of specific protein partner identification could be explained by Rab28 providing general protein unspecific protection, similarly to the *in vitro* unspecific protection offered by MtPM25 protein towards the entire soluble proteome, and not to a particular client protein (Boucher et al 2010). In fact, the images by confocal microscopy show a homogeneous distribution of Rab28 protein in the entire nucleolus, not reveling visible affinities for specific structures. In this line, the specificity of action could be conferred by the subcellular localization, the cell nucleolus, and not only by the protective activity itself. One other aspect that remains to be determined is whether the nature of the biological target may not be a protein but RNA. Considering the importance of the synthesis and processing of ribosomal precursors for cell events, protection and stabilization of nucleolar proteins and RNA molecules may certainly be a relevant issue under drought conditions.

5. Conclusion

The data presented here show that over-expression of Rab28 LEA protein generates maize plants with increased osmotic stress resistance and points to *Rab28* LEA gene as a potential candidate for biotechnological purposes. Moreover, the subcellular localization Rab28 protein in the nucleolus suggests that the physiological function of Rab28 could be linked to the protection and stabilization of nucleolar components/activities under osmotic stress.

General Discussion

VI. General Discussion

LEA proteome in Arabidopsis seeds and in maize embryos

In order to analyze the LEA subproteme in Arabidopsis seeds we first established a protein extraction method suitable as a LEA protein enriching system. The method chosen consisted in the extraction of proteins from Arabidopsis seeds by heat treatment followed by solubilisation in 3% trichloroacetic acid (TCA). With this procedure most LEA-type proteins are recovered in the heat and acid soluble fraction, whereas the storage proteins, still moderately soluble after boiling become insoluble by the acid treatment. Contamination with 12S globulins and 2S albumins is a recurrent issue since storage proteins represent one third of the Arabidopsis seed dry weight. LEAs as natively unfolded proteins are thermostable. Unstructured proteins from mammalian and bacterial cells have been isolated for proteomic identification using either their solubility during acid treatment (Cortese et al. 2005) or their thermostability upon heating (Csizmók et al. 2006; Galea et al. 2006). By combining sequentially both treatments, heat and acid, it is possible to obtain a highly enriched LEA proteome from Arabidopsis seeds.

The composition of the proteins that precipitate and those that remain in solution in 3% TCA was analyzed by two different approaches: 1D SDS-PAGE coupled to LC-ESI-MSMS analysis and a gel-free protocol associated with LC-MALDI-MSMS. A total of 31 different proteins were detected by the gel-free system coupled to LC-MALDI as opposed to 25 identified by the gel-based system coupled to LC-ESI-MSMS, indicating that the gel-free approach performed better. Some very hydrophilic proteins such as AtRab18, the Arabidopsis counterpart of maize Rab17, were only detected by LC-MALDI possibly because the extremely hydrophilic tryptic peptides generated might be lost in the chromatographic separation by LC-ESI-MSMS. Nevertheless, if the gel system analysis had been done along the entire length of the SDS gel, possibly other minor proteins below the sensitivity level of the Coomassie staining might have been identified.

Eighteen LEA proteins were identified using the two analytical systems, thirteen by LC-ESI-MSMS analysis and fourteen in the gel free approach LC-MALDI-MSMS. Of these

LEA proteins, sixteen are predicted to be totally disordered and two show a partially disordered nature. By contrast the remaining identified proteins, storage and other categories, have a globular structure prediction, irrespectively of their acid solubility properties. Disordered proteins typically have a biased amino composition having higher levels of Pro, Glu, Lys, Ser, Gly and Gln amino acids and lower levels of Trp, Tyr, Phe, Cys, Ile, Leu and Val amino acids (Romero et al. 2001). All LEA proteins identified here are examples of biased amino acid composition with a lower relation of order/disorder amino acids.

Arabidopsis thaliana genome contains fifty *lea* genes (Bies-Ethève et al. 2008). Here we identify eighteen LEA proteins in the Arabidopsis seed; thus, more than one third of Arabidopsis LEA proteins are synthesized and accumulated by seeds. The composition of LEA groups indicates that group 3 is dominant. Group 4 is also highly represented but in some classifications LEA proteins from group 4 are partially included in group 3. In the group 2, three members were identified, including AtRab18. The two members of group 1 in Arabidopsis, Atm1 and Atm6, which are embryo-specific LEA proteins were both identified. The remaining proteins were mostly from the storage group and a few unrelated proteins; most of them were recovered in the acid insoluble fraction.

Overall, enriched LEA proteomes can be obtained from Arabidopsis seeds by heat and acid solubility treatments and the resulting protein mixture is susceptible to LC-ESI-MSMS and LC-MALDI-MSMS analysis. The number of LEA proteins identified is high and includes the major LEA groups. This system may be used to analyze the seed proteome in other plant species.

When applying this purification system to excised maize embryos we recovered as components of the unfolded proteome twenty-three polypeptides in the acid-soluble fraction. The proteins included the LEA family, DNA-binding proteins and stress-related proteins. The maize storage family was absent indicating that the storage proteins in embryo maize may not be as heat-soluble as in Arabidopsis; in addition, most of the storage proteins are found in the endosperm which was carefully removed to obtain excised embryos.

The LEA family was the largest group of unfolded proteins in the maize embryo and included proteins from groups 1,2,3,4 and 5. Of the thirteen LEA proteins, group 3 was over-represented, similarly to Arabidopsis. This suggests that members of group 3 may have a significant contribution in seed/embryo desiccation tolerance. Interestingly, Rab17 was the only member from group 2 found in maize embryos, although in other maize inbreds the presence of another embryo dehydrin, DHN2, of 40kDa has also been reported (Asghar et al. 1994). Searches in maize databases (http://www.maizesequence.org/blast) using the specific group LEA Pfam motifs have yielded an estimation of at least nine group 2 LEA proteins in maize embryos. In comparison to Arabidopsis seeds group 2 in maize embryos is under-represented. Since in Arabidopsis some dehydrins are not expressed in seeds (Hundermark and Hincha 2008), this could suggest a specific contribution of some group 2 LEA proteins in vegetative tissue protection during stress. In relation to group 1, Emb564 and Emb5 the two members estimated by database (http://www.maizesequence.org/blast) were both identified.

LEA post-translational modifications

Post-translational modifications of proteins are important features that might be essential for their functionality. LEA protein modifications have been focused on phosphorylation basically. In fact, phosphorylation is not uncommon in Arabidopsis seed LEA subproteome (Irar et al. 2006) and several LEA proteins from different species have been shown to be *in vivo* phosphorylated (Jiang and Wang 2004; Plana et al. 1991;Alsheikh et al. 2003; Heyen et al. 2002; Rohrig et al. 2006). Except for deamination and oxidation in LEAM, a mitochondrial pea LEA protein, so far other LEA protein modifications have not been reported (Tolleter et al. 2007).

Our approach was to analyze the 2D spots correspondent to three LEA proteins, Emb564, Rab17 and Mlg3, from groups 1, 2 and 3 respectively and identify their post-translational modifications. We combined 2D analysis, the use of specific antibodies against the LEA proteins and spot analysis by LC-MSMS.

We found that group 3 Mlg3LEA is not post-translationally modified, group 2 Rab17 is phosphorylated as previously known (see introduction maize Rab17 summary) and group 1 Emb564 displays a complex 2D pattern which combines phosphorylation, acetylation,

methylation and deamination. This is the first report for such PTMs in a group1 LEA protein. In LEAM from group 3, the occurrence of deaminations and oxidations has been proposed to contribute to the functional conformation of the protein (Tolleter et al. 2007). In maize Emb564, several of the observed PTMs are located inside the 20 amino acid conserved motif characteristic of LEA group 1. Despite that in wheat Em protein, removal of group1 conserved repeat did not affect the ability of recombinant Em protein to protect *in vitro* LDH activity from drying, other functional aspects were not evaluated (Gilles et al. 2007).

Another important issue in relation to PTMs is that recombinant proteins used in *in vitro* assays may not hold the PTMs present in native proteins and thus they may not display the complete functional potential of native proteins. PTMs in Emb564 might be critical factors for protein conformation, localization, turnover and/or interactions with subcellular partners in embryo cells.

Antiaggregation properties of Emb564, Rab17 and Mlg3 LEA proteins

The antiaggregation activity of recombinant Emb564, Rab17 and Mlg3 LEA proteins was tested both *in vitro* and *in vivo* using the *E.coli* expression system. In non-stress conditions, expression of maize Emb564, Rab17 or Mlg3 proteins, *per se*, is not deleterious for bacterial cell growth as found for other LEA proteins (Campos et al. 2006). *In vitro* antiaggregation properties were evaluated on salt-soluble *E.coli* proteomes plus recombinant LEA proteins; the protein mixtures were induced to aggregate by desiccation, heat-shock or freezing. Antiaggregation protection was measured by a light scattering assay. The *in vivo* protective effects were investigated by measuring the growth of *E.coli* cells overexpressing Emb564, Rab17 and Mlg3 LEA proteins under desiccation, heat-shock or freezing stress. The correlation of *in vitro* and *in vivo* activities suggests that prevention of aggregation might be the underlying protective mechanism.

The comparison of the three LEA proteins showed that under stress trials the ability to avoid aggregation of the soluble *E.coli* proteome is different. Group 1 Emb564 showed a mild anti-aggregation activity, consistent with other reports of group 1 MtEM from *Medicago trunculata* during freezing or drying (Boucher et al. 2010).

Group 2 Rab17 was always highly efficient both in preventing denaturation of *E.coli* soluble proteome and improving bacterial cell growth during desiccation, heat-shock or low temperature. These effects were not dependent on phosphorylation of the S-segment because recombinant Rab17 is not phosphorylated by *E.coli* cells. The general protective effects of Rab17 are consistent with the wide chaperone activity reported for group 2 ERD10 and ERD14 LEA proteins from Arabidopsis (Kovacs et al. 2008).

Interestingly, group 3 Mlg3 displayed maximal protective activity under dehydration and osmotic stress but was poorly efficient under heat-shock and low temperature challenges. As has been reported, depending on how water is removed, by water-loss or freezing, gradually or rapidly, it is possible that different levels of structure from random coils in solution to mainly α-helical conformation are achieved (Reyes et al. 2005, 2008, Boudet et al. 2006; Hand et al. 2011; Tolleter et al. 2007, Li and He 2009). Pea mitochondrial LEAM, also from group 3, adopts α-helical structure upon air drying but surprisingly not upon lyophillization; the authors conclude that the rapid freezing before dehydration prevents the transition to α-helix structure and that LEAM needs to adopt an helical conformation upon drying in order to interact with membranes (Tolleter et al. 2007). Therefore, increased conformation of maize Mlg3 during drying but not under rapid heat or freezing could explain why the high protective activity of Mlg3 was found exclusively associated to dehydration.

Emb564, Rab17 and Mlg3-GFP fusions in tobacco epidermal cells

To investigate *in vivo* the subcellular localization and the potential protective effects of the different LEA proteins in plant vegetative tissues, transient expressions of GFP-fused LEA proteins were analyzed in agroinfiltrated tobacco leaves. By confocal microscopy fluorescence from Emb564, Rab17 or Mlg3 GFP fusion proteins was found to be cytosolic and nuclear in epidermal tobacco cells. The cytosolic fluorescence cytoplasm was excluded from the vacuoles and the cell wall, and decorated the entire cytosol. The fluorescence distribution in the soluble cytosolic and nuclear compartments indicates the functions of Emb564, Rab17 and Mlg3 proteins might take place in these compartments. Cytosolic localization is not uncommon to LEA proteins (reviewed in Tunnacliffe et al 2007).

Maize group 1 Emb564 protein, despite being seed specific, was stable when overexpressed in agroinfiltrated tobacco leaf cells. The expression of Emb564 only in maize embryonic tissues, is consistent with group 1 LEA proteins being preferentially accumulated in embryo development and not in vegetative tissues (Battaglia et al. 2008; Hundertmark and Hincha 2008, Manfre et al 2009); it has been shown that Atm1 plays a role in the dehydration of Arabidopsis seeds (Manfre et al. 2009). Because maize Emb564 protein holds such a complex combination of PMTs it is plausible that these modifications may be relevant for interaction with seed specific partners. Uncovering the subcellular localization of Emb564 in maize embryogenic tissues will contribute to clarify the specific role of this LEA protein in the drying embryo.

Under dehydration conditions the GFP fluorescent pattern did not change in the three LEA proteins investigated. However, epidermal cells expressing group 3 Mlg3-GFP LEA proteins were significantly bigger compared to cells expressing Emb564-GFP, Rab17-GFP or GFP after the water deficit treatment. Thus, group 3 Mlg3 LEA protein, in contrast to Emb564 or Rab17, has the ability to minimize the *in vivo* cell shrinkage effects due to water losses. During desiccation cells not only increase the molecular composition of the aqueous compartments, or crowding effect, but change their physical dimensions and shape (Mouillon et al. 2008). In cell shrinkage due to water loss, the rate at which the dehydration occurs in physiological systems is essential for cell adjustments. Interestingly, Mlg3 counteracts the shrinkage effects caused by osmotic treatment in an *in vivo* plant system, and protects bacterial growth during osmotic stress but not during rapid freezing or heat-shock. Using synthetic model peptides of group 3 LEA proteins from insects, plants and nematodes it has been shown the occurrence of reversible structural changes between random coils to α-helical coiled coils in response to water activity (Shimizu et al. 2009). In this line, desiccation-induced folding would be a requisite for Mlg3 functionality. Identification of specific intracellular targets for Mlg3 LEA protein in the process of dehydration under different experimental conditions may contribute to define how group 3 Mlg3 LEA protein minimizes the cell shrinkage effects due to water losses.

Under heat-shock tobacco epidermal cells expressing Rab17-GFP maintained a higher viability compared to cells overproducing Emb564-GFP, Mlg3-GFP or GFP. This effect was not dependent on Rab17 phosphorylation because tobacco cells agroinfiltrated with

mRab17, a mutant version of Rab17 LEA protein which cannot be phosphorylated by CK2 (Riera et al. 2004) were resistant to heat-shock as wild type Rab17-GFP fusions. Thus over-expression of Rab17-GFP fusion protein, irrespectively of its phosphorylation status, preserved epidermal cell viability from the damaging effects of thermal shock. This effect relates to the high antiaggregation activity and protection conferred by the non-phosphorylated form of group 2 Rab17 to bacterial cells during heat shock, in contrast to group 1 Emb564 or group 3 Mlg3 LEA proteins.

But more importantly, we provide evidence for the *in vivo* colocalization of Rab17 protein to leaf oil bodies under stress. Colocalization was clearly observed in heat-shocked and to a lesser extend in osmotically stressed leaves, but never in non stressed leaves. Although in higher plants, oil bodies are mostly found in seeds where they have a role as food reserves for germination (Huang et al. 2009), oil bodies are also synthesized by leaves (Lersten et al. 2006, Huang et al. 2009) and a role for oil bodies in adaptation to drought and freezing tolerance has been proposed (Aubert et al. 2010, Lersten et al. 2006; Shimada and Hara-Nishimura 2010). Binding of Rab17 to leaf oil bodies during heat-shock makes this organelle one biological target for Rab17 LEA protein under stress and supports a role for dehydrins in membrane protection. The association of dehydrins to membranes has been previously reported. Wheat WCOR414 and Arabidopsis LTi29 acidic dehydrins were immunodetected in the vicinity of the plasma membrane during cold acclimation in aerial tissues (Danyluk et al. 1998; Puhakainen et al. 2004). *In vitro* studies with Arabidopsis dehydrins ERD10 and ERD14, and maize DHN (Rab17) have show the binding to acidic phospholipid vesicles (Kovacs et al. 2008; Koag et al. 2003); binding and gain in conformation were attributed to the K-segment (Koag et al. 2009). Recently, the contribution of dehydrin K-segments from group 2 Lti30 LEA protein and the relevance of the flanking His side chains as regulators of the interaction between the K-segments and membranes in a pH-dependent manner has been reported (Eriksson et al. 2011). In the present study, we show that the interaction dehydrin-lipid vesicle is triggered by stress in a living cell.

Another aspect is that the unfolded nature of LEA proteins, may favor their multifunctionality (Tompa and Kovacs 2010; Tunnacliffe et al. 2010). Rab17 binds to oil bodies and is highly efficient in preventing protein aggregation by stress denaturation in

E.coli and in tobacco epidermal cells under heat shock; these properties may contribute to the general stabilization of the cellular proteome. However, the functionality of the phosphorylated S-segments is still an open question. In dehydrins ERD10, ERD14 and COR47, phosphorylation determines *in vitro* ion binding capacity and chaperone calcium-dependent function (Alsheikh et al. 2003; Heyen et al. 2002). Here we show that phosphorylation is not involved in Rab17 *in vivo* anti-aggregation properties, in consistency with other *in vitro* reports (Kovacs et al. 2008; Koag et al. 2009) and that Rab17 phosphorylation is not required for interacting with lipid bodies. From previous studies it is known that phosphorylation of Rab17 promotes the arrest of seed germination under salinity (Riera et al. 2004) and that Rab17 phosphorylation enhances protein *in vitro* self-association (M.Figueras, unpublished results). Phosphorylation of Rab17 could participate or modulate interactions with other intracellular partners, perhaps more specifically in seeds where the complete water loss endured by this tissue may enable protein conformations and cellular interactions which may never occur in transient dehydrations in vegetative tissues.

Transgenic maize plants over-expressing LEA *rab28* gene

This work was undertaken in order to investigate the effects of introducing a LEA gene in a crop plant such as maize which productivity is largely affected by drought (Boyer and Westgate, 2004). Thus, transgenic maize plant expressing maize *rab28* LEA gene under the constitutive maize *ubiquitin* promoter were generated. The expression pattern of group 5 *lea* gene *Rab28* resembles that of most LEA proteins: expression along the embryogenesis and in vegetative tissues induction by dehydration and ABA (Pla et al. 1991). Differently from most LEA proteins, the protein in the dry mature embryos is basically located in the nucleolus (Niogret et al.1996).

Rab28 transgenic maize plants exhibited normal development during vegetative growth and showed enhanced resistance under water deficit conditions. Transgenic plants grew faster under PEG treatment than wild-type plants. The physiological parameters analyzed showed that transgenic plantlets had higher RWC, decreased chlorophyll degradation and higher leaf and root areas compared to wild-type plants. MDA production which is a marker for oxidative lipid damage remained reduced in the transgenic lines indicating a lower demand

for the release of antioxidant species while MDA increased significantly in wild-type plants. Transgenic seeds were also able to maintain higher germination rates than wild-type seeds under water deficit. Thus, transgenic Rab28 plants are able to able to sustain increased plant growth rates under water deficit conditions; this points to the potential of this LEA gene to increase crop stress resistance. Transgenic studies overexpressing other LEA genes in diverse host species plant species have also highlighted the benefits of LEA expression in relation to tolerance against abiotic stress (Leprince and Buitink 2010, Yang et al 2010).

One unwanted side-effect in the transgenic plants was the reduction in grain production under normal greenhouse conditions; this effect was more severe in the homozygous transgenic lines. Transgenic approaches in maize often reduce yield since this trait is associated with many genes which can be disrupted. Yield penalties have been reported by other authors in Arabidopsis plants expressing drought-related transcription factors (Shinozaki and Yamaguchi-Shinozaki 2007) and in rice over-expressing a LEA protein from group 3 (Xiao et al 2007); the use of appropriate promoters to drive the expression of genes and the selection of plants with single copies of the transgene may prevent these undesired traits (Xiao et al 2007).

In transgenic maize plants, Rab28 protein was found accumulated in nucleolus of root cells under optimal and osmotic stress conditions. In embryo tissues Rab28 had been previously found localized to the nucleolus of scutellar cells (Niogret et al.1996). Here we have extended this information and have found that Rab28 protein is preferentially associated to the nuclear S2 fraction corresponding to the ribonucleoprotein (RNP) fraction. This opens the possibility of whether the nature of the biological target may not be a protein but RNA. These data suggest that the protective effects of Rab28 observed in the transgenic plants against osmotic stress might be related with the stabilization and protection of the components of the cell nucleolus.

CONCLUSIONS

VII. Conclusions

In the first chapter of present work we apply two LC-MSMS approaches suitable to identify proteins in moderately complex samples. We use a two-step fractionation procedure, heat followed by acid treatment, which has proved to be an efficient method to isolate a protein fractions highly enriched in LEA-type proteins from Arabidopsis seeds. With this procedure a highly purified LEA proteome is recovered in the acid-soluble fraction whereas most of the otherwise ever present seed storage proteins remain acid-insoluble. More interestingly, the amount of LEA-proteins in the acid-soluble fraction is high enough to be visualized by Coomassie staining in SDSPAGE gels. This opens the possibility for large-scale analysis of LEA-type proteins, the search for their *in vivo* posttranslational modifications and the comparison of total LEA proteomes in stress resistant versus sensitive seed varieties and in vegetative tissues under stress.

In the second chapter, we present evidence illustrating differences among the LEA representatives of group 1, 2 and 3 in maize. Group 1 Emb564 stands out for displaying a complex combination of different PTMs, including phosphorylation, acetylation, methylation and deamination in the native protein, which may be relevant for its seed-specific role. Group 3 Mlg3 shows high anti-aggregative properties specifically associated with dehydration conditions which, in agroinfiltrated tobacco epidermal cells, result in the reduction of cell shrinkage effects due to cell water losses. Finally, group 2 Rab17, which displays a general potent anti-aggregation activity that prevents protein denaturation under stress, also binds to leaf oil bodies under heat shock *in vivo*, supporting a role for dehydrins in membrane protection.

In the third chapter of this thesis we show that over-expression of Rab28 LEA protein generates maize plants with increased osmotic stress resistance and points to *Rab28* LEA gene as a potential candidate for biotechnological purposes. Moreover, the subcellular localization Rab28 protein in the nucleolus suggests that the physiological function of Rab28 could be linked to the protection and stabilization of nucleolar components/activities under osmotic stress.

SPANISH SUMMARY

VIII. Spanish Summary

De acuerdo con los requisitos del formato de la tesis doctoral escrita en un idioma no oficial de la Universidad de Barcelona, se presenta el resumen de esta memoria en castellano. Esta tesis doctoral, con título ¨**Abiotic stress in plants: Late Embryogenesis Abundant proteins**¨, ha sido realizada en el laboratorio del Prof. M. Pagés bajo la dirección de Dr. Adela Goday, en el Departamento de Genética Molecular del CSIC, Center for Research in Agricultural Genomics (CRAG) con la financiación del programa F.I. (Formació Personal Investigador), l'Agencia de Gesti D'Ajuts Universitaris i de Recerca (AGAUR) la Generalitat de Catalunya, 2011FI_B2 0012, España. La tesis comenzó en el año académico 2006-2007. Esta tesis doctoral ha sido inscrita en el programa de Biotecnología Vegetal de la Facultad de Farmacia de la Universidad de Barcelona.

El grupo del Prof. M. Pagés, Laboratorio de Genética Molecular de Plantas, CSIC, CRAG está dedicado a estudiar los mecanismos de tolerancia al estrés abiótico en plantas. Con este objetivo ha realizado diversos trabajos dirigidos a identificar y caracterizar genes y proteínas implicados en la resistencia al estrés hídrico. Las investigaciones abarcan aproximaciones diversas, como genes reguladores, factores de transcripción, mecanismos de transducción de señal, regulación por fosforilación y estudio de las proteínas LEA ó Late Embryogenesis Abundant.

Introducción

Las proteínas LEA, originalmente fueron descritas en las semillas de algodón; se acumulan en grandes cantidades en estructuras tolerantes a la desecación (semillas, polen) y en tejidos vegetativos sometidos a estrés abiótico, sequía, salinidad y frío. También se hallan en organismos anidrobióticos, en plantas de resurrección, algunos invertebrados y microorganismos. La presencia de proteínas LEA se correlaciona con la adquisición de tolerancia a la desecación. Desde un principio se les atribuyó un papel en las respuestas de las plantas en la adaptación al estrés (revisado en Bartels and Salamini 2001, Tunnacliffe 2007, Shih et al. 2010, Tunnacliffe 2010, Hand et al. 2011).

Las proteínas LEA se clasifican en diversos grupos en función de dominios y secuencias de aminoácidos específicos (Wise 2010, Batagglia et al 2008, Bies-Ethève et al 2008). Los grupos 1, 2 y 3 son los más relevantes ya que abarcan la mayoría de las proteínas de la familia LEA. Una característica general de estas proteínas es su elevada hidrofilicidad, alto contenido de aminoácidos cargados y su falta de estructura en estado hidratado. A pesar de encontrarse mayoritariamente en forma de "random coil", algunas adquieren un cierto grado de estructura durante la deshidratación o en la presencia de agentes promotores de α-hélices (Shih et al. 2010, Hand et al. 2011). A nivel celular se han hallado en todas las localizaciones, citosol, núcleo, nucleolo, mitocondria, cloroplasto, vacuola, retículo endoplásmico, peroxisoma y membrana plasmática, donde se supone ejercen su función protectora frente al estrés (Tunnacliffe and Wise 2007, Hundertmark and Hincha 2008). En relación a las modificaciones post-traduccionales, algunas se hallan fosforiladas (Jiang and Wang 2004; Plana et al. 1991, Heyen et al. 2002, Rohrig et al. 2006). Los efectos protectores de las varias proteínas LEA se han demostrado mediante ensayos *in vitro* y en aproximaciones transgénicas que han dado lugar a fenotipos resistentes a la sequía, sal y frío. Por lo general, se considera que estas proteínas contribuyen a la protección y a la estabilización de macromoléculas y estructuras celulares en las respuestas de adaptación al estrés en plantas; sin embargo, sus funciones específicas aún no han sido esclarecidas. A nivel molecular se ha propuesto que las funciones de las proteínas LEA pueden ser variadas: estabilización y renaturalización de proteínas, mantenimiento de membranas, en combinación, o no, con azúcares, tampones de hidratación (substitución de moléculas de agua), afinidad por iones y función antioxidante (Tunnacliffe and Wise 2007, Shih et al. 2010, Batagglia et al. 2008).

Objetivos

En esta tesis se pretende ampliar los conocimientos sobre las proteínas LEA y sus funciones relativas a la tolerancia a la sequía. Los resultados están presentados en forma de capítulos.

Capítulo 1: Aproximación proteómica para el estudio de la composición de proteínas LEA en semillas de Arabidopsis. Se establece un método para obtener poblaciones enriquecidas en proteínas LEA basado en un procedimiento de extracción que combina la estabilidad al calor y la solubilidad en ácido de las proteínas de tipo LEA, proteínas no estructuradas en estado hidratado. Las fracciones proteicas obtenidas fueron analizadas por espectrometría de masas (MS).

Capítulo 2: Se analiza el contenido de proteínas LEA en embriones maduros de maíz, utilizando el método descrito anteriormente. Del conjunto de proteínas LEA identificadas, se seleccionaron Emb564, Rab17 y Mlg3 pertenecientes a los grupos 1, 2 y 3 respectivamente, para su posterior estudio dirigido a detectar diferencias funcionales entre ellas. Con este objetivo, se analizaron las modificaciones post-traduccionales (PTM) de las proteínas, las propiedades antiagregantes de las correspondientes proteínas recombinantes *in vitro*, su capacidad de protección *in vivo* frente al estrés abiótico en células de *E.coli* y en hojas de tabaco y su localización subcelular.

Capítulo 3: Se generaron plantas transgénicas de maíz sobre-expresando el gen *rab28* del grupo LEA 5 bajo el control del promotor de la *ubiquitina* con el fin de investigar su nivel de tolerancia al estrés hídrico. Los fenotipos de las plantas transgénicas fueron caracterizados en condiciones óptimas de crecimiento y bajo estrés osmótico. Paralelamente, se realizó el estudio de la localización subcelular de la proteína Rab28.

Resultados y Discusión

Capítulo I: Identificación por LC-MSMS de proteínas de semillas de *Arabidopsis thaliana* estables al calor y ácido

Las proteínas LEA acumuladas en semillas de Arabidopsis fueron enriquecidas por tratamiento con calor seguido de tratamiento con ácido (3%TCA). Estas propiedades ya habían sido utilizadas anteriormente por separado para obtener proteomas no estructurados de células de mamífero y bacterias (Cortese et al. 2005, Csizmok et al. 2006, Galea et al. 2006). Aquí se han combinado ambos tratamientos para obtener fracciones enriquecidas en proteínas LEA a partir de semillas. Las proteínas de reserva de Arabidopsis, globulinas 12S y albúminas 2S, son mayoritarias en semilla y suelen ser una fuente importante de

contaminación; estas proteínas son parcialmente termoestables pero insolubles en ácido, por lo que se concentran sobretodo en la fracción insoluble en ácido obtenida. La identificación de proteínas se realizó mediante dos sistemas: 1) análisis por geles de acrilamida (SDS-PAGE) de las fracciones proteicas aisladas, teñido de bandas por coomassie, disección de bandas e identificación de polipéptidos por LC-ESI-MSMS y 2) análisis de la fracción estable al calor y soluble en ácido por LC-MALDI-MSMS. Se identificaron un total de 18 proteínas LEA. Su composición reveló preponderancia de aminoácidos Pro, Glu, Lys, Ser, Gly y Gln y niveles bajos de Trp, Tyr, Phe, Cys, Ile, Leu and Val, en concordancia con su naturaleza desordenada. El genoma de Arabidopsis contiene 50 genes *lea*, en consecuencia al menos un tercio se expresan en semilla. Se identificaron proteínas LEA de los grupos 1, 2, 3, 4 y 5, siendo el grupo 3 el grupo más representado. Se identificaron tres proteínas del grupo 2 y la totalidad de grupo 1, Atm1 y Atm6; las proteínas del grupo 1 en Arabidopsis son específicas de semilla y no se inducen en tejidos vegetativos bajo estrés (Manfre et al. 2009).

El enriquecimiento de proteínas solubles al calor y ácido facilita enormemente el análisis a gran escala de proteínas de tipo LEA, comparación de la composición de la familia LEA entre semillas de especies o variedades distintas y el análisis de sus modificaciones post-traduccionales. Además, el sistema puede ser adaptado para analizar y comparar la inducción de proteínas LEA en distintos tejidos vegetativos bajo distintas situaciones de estrés.

Capitulo 2: Proteínas LEA en maíz: aproximación proteómica y funcional

Proteómica

El maíz es una planta de interés agronómico. El proteoma LEA se obtuvo a partir de embriones maduros diseccionados para separarlos del endospermo, tejido muy rico en proteínas de reserva. La identificación por MS de los polipéptidos del proteoma no-estructurado reveló que en embrión de maíz la familia LEA es mayoritaria, seguida por las proteínas DNA-binding. Se identificaron 13 proteínas LEA de los grupos 1, 2, 3, 4 y 5. Al igual que en Arabidopsis, el grupo 3 fue el más numeroso, indicando que las proteínas de este grupo deben de contribuir significativamente a la tolerancia a la desecación del

embrión. Del grupo 2 ó deshidrinas, sólo se identificó Rab17, aunque otras variedades de maíz contienen además la deshidrina DH2 (Asghar et al. 1994). La búsqueda en las bases de datos (http://www.maizesequence.org/blast) indicó una estimación mínima de 9 deshidrinas en el genoma del maíz. En comparación con Arabidopsis, el grupo 2 en el embrión de maíz está poco representado. Puesto que algunas deshidrinas no se expresan en semillas de Arabidopsis (Hundermark and Hincha 2008) ello podría indicar una contribución específica de algunas proteínas LEA del grupo 2 en la respuesta al estrés en tejidos vegetativos. En relación al grupo 1, las dos proteínas estimadas, Emb564 y Emb5, ambas fueron identificadas.

Modificaciones post-traduccionales (PTM)

Las PTM de las proteínas son características que pueden ser esenciales para su funcionalidad. En las proteínas LEA, los estudios se han centrado en la fosforilación, especialmente en deshidrinas (Plana et al. 1991; Alsheikh et al. 2003; Jiang and Wang 2004, Heyen et al. 2002; Rohrig et al. 2006, Irar et al. 2006). Sin embargo, aparte de la desaminación y oxidación de LEAM, proteína LEA mitocondrial del guisante, no se han descrito otras PTM en la familia LEA. En este trabajo se analizó el patrón bidimensional de tres proteínas LEA seleccionadas: Emb564 del grupo 1, Rab17 del grupo 2 y Mlg3 del grupo 3. Se combinó el uso de anticuerpos específicos generados contra las correspondientes proteínas recombinantes, el análisis por electroforesis bidimensional (2D) y análisis de spots por LC-MSMS. Los resultados mostraron que Mlg3 no presenta PTM y aparece en forma de spot único en geles 2D; Rab17 está altamente fosforilada y aparece en forma de una ristra de spots en la zona ácida del gel de 2D tal como ya se había descrito (Vilardell et al. 1990, Plana et al. 1991); Emb564 presenta un patrón bidimensional complejo que combina fosforilación, acetilación, metilación y desaminación. Algunas de estas PTM se hallan en el motivo conservado de 20 aminoácidos característico de las proteínas LEA del grupo 1. Aunque en la proteína Em de trigo del grupo 1 la eliminación de este motivo conservado no modifica la actividad protectora de la proteína recombinante *in vitro* (Gilles et al. 2007), en la proteína LEAM del grupo 3 la desaminación y oxidación contribuyen a la conformación funcional de la proteína (Tolleter et al. 2007). Un aspecto importante relativo a las MPT es el hecho de que las proteínas recombinantes empleadas en

los ensayos *in vitro* no tengan las mismas PMT que las proteínas nativas, con lo cual podrían no ser totalmente funcionales. En el caso de Emb564, el patrón complejo de PMT que presenta puede ser relevante para su conformación, localización, recambio y/o interacción con dianas intracelulares *in vivo*.

Propiedades antiagregantes de las proteínas Emb564, Rab17 y Mlg3

Las actividades antiagregantes de Emb564, Rab17 y Mlg3 se analizaron *in vitro* e *in vivo* en un sistema de expresión en *E.coli*. En ningún caso la expresión *per se* disminuyó la tasa de crecimiento de las bacterias, en contraste con otras observaciones en que proteínas LEA recombinantes reducen el crecimiento de *E.coli* (Campos et al. 2006). La actividad antiagregante de cada una de las proteínas LEA se estudió *in vitro* evaluando la desnaturalización del proteoma soluble en sal de *E.coli*, bajo condiciones de desecación (secado al vacio, 2 horas), choque térmico (48ºC, 45 min) y congelación (-20º, 2 horas); la agregación del proteoma se cuantificó midiendo la absorbancia a 340 nm. *In vivo*, los efectos protectores de las proteínas LEA se evaluaron comparando las tasas de crecimiento de las bacterias en condiciones óptimas y bajo estrés osmótico con 10% PEG (polietilenglicol), choque térmico o congelación. Hay que mencionar que los tres tratamientos producen directa o indirectamente un déficit hídrico puesto que la temperatura incremente la evaporación y la congelación reduce la disponibilidad de agua libre. Los resultados obtenidos mostraron una clara correlación entre el comportamiento *in vitro* e *in vivo* de las proteínas.

De las tres proteínas LEA estudiadas, Emb564 del grupo 1 muestra una actividad antiagregante reducida, en concordancia con las observaciones realizadas previamente con MtEM, otra proteína del grupo 1 de *M. trunculata,* durante estrés por sequía o congelación (Boucher et al.2010).

Rab17 del grupo 2 es altamente eficaz como antiagregante y mejora substancialmente el crecimiento bacteriano bajo estrés osmótico, choque térmico o congelación. Esta elevada capacidad antiagregante es independiente de fosforilación puesto que Rab17 recombinante no se fosforila en *E.coli*. Los efectos protectores de Rab17 concuerdan con la amplia

actividad chaperona atribuida a las deshidrinas ERD10 y ERD14 de Arabidopsis (Kovacs et al. 2008).

Mlg3 del grupo 3 es un caso interesante porque es muy eficaz como antiagregante y protector de crecimiento bajo sequía, igual o más que Rab17, y poco eficiente bajo tratamiento de choque térmico o frío. Este comportamiento se puede atribuir al hecho de que dependiendo de la velocidad a la que se pierde agua, gradual o rápidamente, se consiguen distintos niveles de conformación α-hélice a partir de ausencia de conformación o random coil (Reyes et al. 2005, 2008, Boudet et al. 2006; Hand et al. 2011; Tolleter et al. 2007, Li and He 2009). De hecho, la proteína LEAM mitocondrial del guisante, del grupo 3, adopta estructura en α-hélice al secarse, pero no al liofilizarse; la congelación rápida que precede a la liofilización evita la transición a α-hélice la cual permite su interacción con membranas (Tolleter et al. 2007). Así pues, la adquisición de conformación por parte de Mlg3 durante la deshidratación y no durante el choque térmico o congelación rápida, podría justificar su actividad protectora exclusivamente durante la deshidratación.

Expresión de proteínas de fusión Emb564-GFP, Rab17-GFP y Mlg3-GFP en tabaco

Para determinar la localización subcelular y la capacidad de protección de las proteínas Emb564, Rab17 y Mlg3 frente al estrés, se agroinfiltraron hojas de tabaco con las proteínas LEA fusionadas a la proteína fluorescente GFP. Mediante microscopia confocal se observó que las tres proteínas de fusión se distribuían en el citosol y núcleo de células epidérmicas agroinfiltradas, con un patrón uniforme, soluble sin mostrar afinidad por estructuras determinadas; quedaban excluidas de vacuolas, pared celular y otros compartimentos. Esta distribución indica que muy probablemente las funciones de Emb564, Rab17 y Mlg3 ocurran en la fracción soluble del citosol y núcleo.

Aunque la proteína de fusión Emb564-GFP es estable en hojas de tabaco agroinfiltradas, Emb564 en maíz sólo se expresa en embriones y no es inducible por estrés en tejidos vegetativos tal como ya ha sido descrito para otras proteínas del grupo 1 (Battaglia et al. 2008; Hundertmark and Hincha 2008, Manfre et al. 2009). En maíz, la complejidad de las PTM observadas en la proteína nativa Emb564 podrían ser relevantes para la funcionalidad de esta proteína; en este sentido, el determinar la distribución intracelular de Emb564 en el

embrión maduro del maíz sería conveniente para conocer el papel que juega esta proteína en el embrión durante su maduración y desecación.

En condiciones de estrés osmótico (riego de las plantas de tabaco agroinfiltradas con 15% PEG durante 2 días) la distribución intracelular soluble de Emb564-GFP, Rab17-GFP y Mlg3–GFP no varía. Sin embargo, se observó que las células que sobreexpresaban Mlg3-GFP eran de mayor tamaño que las restantes. Durante la desecación, la concentración intracelular aumenta y el volumen de la célula se reduce debido a la pérdida de agua produciendo un "encogimiento" celular (Mouillon et al. 2008). Mlg3 compensa, al menos en parte, los efectos de "encogimiento" celular de las células epidérmicas ocasionados por el tratamiento con PEG. En un modelo usando péptidos sintéticos de una proteína LEA del grupo 3, se constató que la transición de random coil aα -hélice, coiled coil, dependía de la disponibilidad de agua (Shimizu et al. 2009). Siguiendo este modelo, la adquisición de cierta estructura por parte de Mlg3, podría ser un requisito previo a su funcionalidad; en este contexto, la identificación de posibles dianas intracelulares bajo condiciones de deshidratación progresiva podría esclarecer cómo Mlg3 reduce los efectos de "encogimiento" celular producidos durante el déficit hídrico.

Con el choque térmico (plantas de tabaco agroinfiltradas sometidas a 48ªc durante 45 min) las células epidérmicas expresando Rab17-GFP mostraron una viabilidad superior a las restantes. Mediante geles 2D se comprobó que la proteína de fusión Rab17-GFP se hallaba efectivamente fosforilada en hojas de tabaco. Algunas deshidrinas, como Rab17, contienen en su secuencia el motivo S compuesto de varias Ser consecutivas, que *in vivo* se halla altamente fosforilado (Vilardell et al. 1990, Plana et al. 1991); de hecho, Rab17 es la proteína fosforilada más abundante del embrión maduro de maíz (Goday et al. 1994). Empleando una versión de Rab17 mutada que no puede fosforilarse por llevar una alteración en el consenso EDD para la CK2 (Riera et al. 2004), se comprobó que el incremento de viabilidad celular observada con Rab17 también ocurría en las células incapaces de fosforilar Rab17. Así pues, la protección ejercida por Rab17 frente al choque térmico no depende de su estado de fosforilación. El efecto protector probablemente sea consecuencia directa de la gran capacidad antiagregante de la proteína Rab17 no fosforilada observada en células de *E.coli*.

Sin embargo, el resultado importante relativo a Rab17 fue observar *in vivo* su colocalización en cuerpos oleicos (oil bodies) de hoja bajo estrés. Esta colocalización ocurría después del choque térmico, y en menor grado en las plantas tratadas con PEG, pero nunca en plantas sin tratar. En plantas, la mayoría de cuerpos oleicos se encuentran en la semilla donde tienen un función de reserva para la germinación (Huang et al. 2009); sin embargo, también se han descrito en hoja (Lersten et al. 2006, Huang et al. 2009) donde se les atribuye una función relacionada con la tolerancia a la sequía y bajas temperaturas (Aubert et al. 2010, Lersten et al. 2006; Shimada and Hara-Nishimura 2010). La interacción de Rab17 con cuerpos oleicos *in vivo* después de un choque térmico indica que este orgánulo podría ser una de las dianas intracelulares de Rab17 bajo estrés, y apoya el papel de las deshidrinas en la protección de membranas celulares.

La asociación de deshidrinas a membranas ya había sido observada previamente. WCOR414 de trigo y Lti29 de Arabidopsis se localizaran por imunocitoquímica cerca de la membrana plasmática de tejidos aéreos durante la aclimatación a bajas temperaturas (Danyluk et al. 1998; Puhakainen et al. 2004). In vitro, las deshidrinas ERD10 and ERD14 de Arabidopsis y DHN (Rab17) del maíz se unen a vesículas acídicas de fosfolípidos (Kovacs et al. 2008; Koag et al. 2003); la unión y adquisición de cierta conformación fue atribuida al motivo K (Koag et al. 2009), motivo presente en todas las proteínas LEA del grupo 2, EKKGIMDKIKEKLPG (Galau and Close 1992). En un estudio reciente (Eriksson et al. 2011), se ha demostrado que la participación del motivo K de la deshidrina Lti30 de Arabidopsis en la interacción con membranas, y especialmente de las cadenas de lisina flanqueantes, es un efecto dependiente del pH. En el presente estudio se demuestra que la interacción deshidrina Rab17-cuerpo oleico está inducida por estrés en células vivas.

Otro aspecto a considerar es el hecho de que la naturaleza desordenada de las proteínas LEA puede favorecer la multifuncionalidad (Tompa and Kovacs 2010; Tunnacliffe et al. 2010). Hemos comprobado que Rab17 evita de un modo eficaz la agregación proteica consecuencia de la desnaturalización por estrés, tanto en *E.coli* como en células epidérmicas de tabaco; estas propiedades pueden contribuir a la estabilización general del proteoma celular bajo un estrés. Sin embargo, la funcionalidad de los motivos S presentes en algunas proteínas LEA del grupo 2, los cuales *in vivo* se hallan altamente fosforilados,

sigue siendo una cuestión abierta. En las deshidrinas de Arabidopsis ERD10, ERD14 y COR47, la fosforilación determina *in vitro* la capacidad de unión a iones y la actividad chaperona dependiente de calcio (Alsheikh et al. 2003; Heyen et al. 2002). En el presente estudio se demuestra que la fosforilación no está implicada ni en la actividad antiagregante de Rab17 *in vivo*, en concordancia con otros trabajos (Kovacs et al. 2008; Koag et al. 2009), ni en la interacción Rab17 con cuerpos oleicos. Sin embargo, la fosforilación de Rab17 interviene en el paro germinativo que sufren las semillas de Arabidopsis en condiciones de salinidad (Riera et al. 2004). Así pues, la fosforilación de Rab17 podría participar o modular activamente interacciones con otras dianas celulares, quizás de un modo más específico en las semillas, donde los niveles de pérdida de agua alcanzados son muy superiores a los sufridos por otros tejidos; de este modo, en un embrión de maíz maduro y deshidratado se podrían favorecer estructuras o interacciones que nunca ocurran en las deshidrataciones parciales y transitorias que se dan en los tejidos vegetativos.

En conclusión, en el capítulo 2 de la presente tesis se analiza el proteoma LEA de embriones de maíz y se ponen de manifiesto las diferencias existentes entre las proteínas LEA Emb564 del grupo 1, Rab17 del grupo 2 y Mlg3 del grupo 3. Emb564 destaca por las numerosas PMT, fosforilación, acetilación, metilación y desaminación de la proteína nativa que serían relevantes para su función específica en semilla. Mlg3 es muy eficaz como antiagregante en condiciones de deshidratación en *E. coli* y en células epidérmicas de tabaco agroinfiltradas, la presencia de Mlg3 reduce considerablemente los efectos de "encogimiento" celular provocados por la pérdida de agua. Finalmente, Rab17 muestra propiedades antiagregantes generales y muy potentes que evitan la desnaturalización de proteínas en condiciones de sequía, choque térmico ó bajas temperaturas; bajo choque térmico Rab17 interacciona *in vivo* con cuerpos oleicos, lo que apoya el papel de protección de membranas de las deshidrinas durante el estrés.

Capitulo 3: Plantas transgénicas de maíz que sobreexpresan el gen LEA *rab28*.

En este capítulo se analizan los efectos de la sobreexpresión de un gen LEA en el maíz, planta de interés agronómico y cuya productividad se ve muy afectada por la sequía (Boyer and Westgate, 2004). Se generaron plantas transgénicas de maíz expresando el gen LEA *rab28* del grupo 5 bajo el promotor constitutivo de la *ubiquitina*. El patrón de expresión de

rab28 en maíz sigue el patrón típico de un gen LEA: aumenta a lo largo de la embriogénesis, desaparece durante la germinación y es inducible en tejidos vegetativos por deshidratación y ABA (Pla et al. 1991). A diferencia de otras proteínas LEA, Rab28 se localiza en los nucléolos de las células del embrión maduro (Niogret et al. 1996).

Las plantas transgénicas de maíz sobreexpresando el gen *rab28* no presentaron diferencias apreciables con las de tipo salvaje en cuanto a su desarrollo y crecimiento en condiciones óptimas de cultivo en el invernadero. Sin embargo, el condiciones de déficit hídrico provocado por la presencia 15% PEG en el medio de riego, las plantas transgénicas crecían mejor que las de tipo salvaje. En estas condiciones, el análisis de diversos parámetros fisiológicos indicaron que las plántulas transgénicas tenían un contenido hídrico (RWC) superior, menos degradación de clorofila y áreas foliares y radiculares mayores que las plántulas salvajes. La producción de MDA (malondialdehyde), marcador de daño lipídico por oxidación, era bajo en las transgénicas indicando que la necesidad de producir compuestos antioxidantes era poca; por el contrario, la liberación de MDA se hallaba incrementada en las plantas control. Además, en condiciones de estrés osmótico las plantas transgénicas mostraron una capacidad de germinación superior. Así pues, en conjunto las plantas transgénicas *rab28* crecen a una mayor velocidad en condiciones de estrés hídrico. Estos resultados sugieren el potencial del gen LEA *rab28* para aumentar la resistencia de cultivos frente al estrés. De hecho, varios estudios han mostrado el beneficio que supone la sobreexpresión de genes LEA en otras especies vegetales en relación a la tolerancia al estrés (Leprince and Buitink 2010, Yang et al. 2010).

En las plantas de maíz transgénico *rab28* se observó una reducción en el número de granos por mazorca. La reducción era más pronunciada en las plantas homozigotas que las heterozigotas. El peso y aspecto de las semillas producidas era, sin embargo, normal. No es infrecuente que las aproximaciones transgénicas en maíz resulten en una reducción de la producción de semillas, puesto que este carácter está asociado a numerosos genes. En plantas de Arabidopsis que sobreexpresan factores de transcripción de estrés y en plantas de arroz que sobreexpresan una proteína LEA del grupo 3 las penalizaciones en la producción ya se habían descrito (Shinozaki and Yamaguchi-Shinozaki 2007, Xiao et al. 2007). El uso

de promotores apropiados y la selección de plantas con copias únicas del transgen pueden evitar o reducir los efectos colaterales indeseados (Xiao et al. 2007).

Respecto a la localización intracelular, la proteína Rab28 se localizó mediante inmunocitoquímica "whole mount" y microscopía confocal en los nucléolos de las células de raíz de plantas transgénicas, tanto en condiciones de estrés osmótico como en condiciones óptimas de crecimiento. Con anterioridad, Rab28 ya había sido localizada en los nucléolos de las células del escutelo de embriones maduros de maíz (Niogret et al. 1996). En el presente estudio hemos ampliado esta información puesto que por fraccionamiento de núcleos aislados de embrión se ha comprobado que la proteína Rab28 está asociada a la fracción S2 o fracción ribonucleoproteica (RNP). La fracción S2 corresponde al proteoma soluble en tampón de baja fuerza iónica y contiene 70-80% de RNA y proteínas nucleolares implicadas en la síntesis y procesamiento de los precursores ribosomales (Cárcer et al. 1997). Esta observación abre la posibilidad de que la diana biológica de Rab28 no sea necesariamente, o no sólo, una proteína sino también RNA. Finalmente y a modo de resumen, los datos expuestos en el capítulo 3 de esta tesis demuestran que la sobreexpresión del gen *rab28* de la familia LEA, grupo 5, en plantas de maíz genera plantas más resistentes al estrés osmótico, lo que indica el potencial biotecnológico de este gen LEA; a nivel subcelular la proteína Rab28 tiende a acumularse en el nucleolo de la célula, tanto en embriones como en tejido vegetativo, lo que sugiere que Rab28 podrían contribuir a la estabilización y protección de los componentes de la estructura nucleolar durante el estrés osmótico.

BIBLIOGRAPHY

IX. Bibliography

Abba, S., Ghignone, S. and Bonfante, P. (2006) A dehydration-inducible gene in the truffle Tuber borchii identifies a novel group of dehydrins. *BMC Genomics* 7: 39.

Ahuja, I., de Vos, R.C., Bones, A.M. and Hall, R.D. (2010) Plant molecular stress responses face climate change. *Trends Plant Sci* 15: 664-674.

Ahuja I, d.V.R., Bones AM, Hall RD (2010) Plant molecular stress responses face climate change. . *Trends Plant Sci* 15: 664-674.

Alpert, P. (2005) The Limits and Frontiers of Desiccation-Tolerant Life. *Integrative and Comparative Biology* 45: 685-695.

Alsheikh, M., Heyen, B. and Randall, S. (2003) Ion binding properties of the dehydrin ERD14 are dependent upon phosphorylation. *J Biol Chem* 278: 40882 - 40889.

Alsheikh, M., Svensson, J. and Randall, S. (2005) Phosphorylation regulated ion-binding is a property shared by the acidic subclass dehydrins. *Plant Cell Environ* 28: 1114 - 1122.

Amara, I., Odena, A., Oliveira, E., Moreno, A., Masmoudi, K., Pages, M. and Goday, A. (2012) Insights into Maize LEA proteins: from proteomics to functional approaches. *Plant Cell Physiol* 53: 312-329.

Arenas-Mena, C., Raynal, M., Borrell, A., Varoquaux, F., Cruz Cutanda, M., Stacy, R.A.P., Pagès, M., Delseny, M. and Culiáñez-Macià, F.A. (1999) Expression and cellular localization of Atrab28 during Arabidopsis embryogenesis. *Plant Molecular Biology* 40: 355-363.

Armstrong, C.L., Romero-Severson, J. and Hodges, T.K. (1992) Improved tissue culture response of an elite maize inbred through backcross breeding, and identification of chromosomal regions important for regeneration by RFLP analysis. *TAG Theoretical and Applied Genetics* 84: 755-762.

Asghar, R., Fenton, R.D., DeMason, D.A. and Close, T.J. (1994) Nuclear and cytoplasmic localization of maize embryo and aleurone dehydrin. *Protoplasma* 177: 87-94.

Aubert, Y., Vile, D., Pervent, M., Aldon, D., Ranty, B., Simonneau, T., Vavasseur, A. and Galaud, J.-P. (2010) RD20, a Stress-Inducible Caleosin, Participates in Stomatal Control, Transpiration and Drought Tolerance in Arabidopsis thaliana. *Plant and Cell Physiology* 51: 1975-1987.

Baker J, S.C., Dure L III (1988) Sequence and characterization of 6 Lea proteins and their genes from cotton. *Plant Mol Biol* 11: 277-291.

Baker, J., Steele, C. and Dure, L.I. (1988) Sequence and characterization of 6 Lea proteins and their genes from cotton. *Plant Mol Biol* 11: 277-291.

Bartels, D. and Salamini, F. (2001) Desiccation Tolerance in the Resurrection PlantCraterostigma plantagineum. A Contribution to the Study of Drought Tolerance at the Molecular Level. *Plant Physiology* 127: 1346-1353.

Bartels, D. and Sunkar, R. (2005) Drought and salt tolerance in plants. *Crit Rev Plant Sci* 24: 23 - 58.

Battaglia, M., Olvera-Carrillo, Y., Garciarrubio, A., Campos, F. and Covarrubias, A.A. (2008) The enigmatic LEA proteins and other hydrophilins. *Plant Physiol* 148: 6-24.

Battista, J., Park, M.-J. and McLemore, A. (2001) Inactivation of two homologues of proteins presumed to be involved in the desiccation tolerance of plants sensitizes Deinococcus radiodurans R1 to desiccation. *Cryobiology* 43: 133 - 139.

Bies-Etheve, N., Gaubier-Comella, P., Debures, A., Lasserre, E., Jobet, E., Raynal, M., Cooke, R. and **Delseny, M**. (2008) Inventory, evolution and expression profiling diversity of the LEA (late embryogenesis abundant) protein gene family in Arabidopsis thaliana. *Plant Mol Biol* 67: 107-124.

Blackman, S.A., Obendorf, R.L. and Leopold, A.C. (1995) Desiccation tolerance in developing soybean seeds: The role of stress proteins. *Physiologia Plantarum* 93: 630-638.
Blum, A. (2011) The Interdrought Conference in Perspective. *Journal of Experimental Botany*.
Bodnar, W., Blackburn, R., Krise, J. and Moseley, M. (2003) Exploiting the complementary nature of LC/MALDI/MS/MS and LC/ESI/MS/MS for increased proteome coverage. *Journal of The American Society for Mass Spectrometry* 14: 971-979.
Borovskii, G., Stupnikova, I., Antipina, A., Vladimirova, S. and Voinikov, V. (2002) Accumulation of dehydrin-like proteins in the mitochondria of cereals in response to cold, freezing, drought and ABA treatment. *BMC plant biology* 2: 5.
Borovskii, G.B., Stupnikova, I.V., Antipina, A.I., Downs, C.A. and Voinikov, V.K. (2000) Accumulation of Dehydrin-like-proteins in the Mitochondria of Cold-treated Plants. *Journal of Plant Physiology* 156: 797-800.
Borrell, A., Cutanda, M.C., Lumbreras, V., Pujal, J., Goday, A., Culiáñez-Macià, F.A. and Pagès, M. (2002) Arabidopsis thaliana Atrab28: a nuclear targeted protein related to germination and toxic cation tolerance. *Plant Molecular Biology* 50: 249-259.
Boucher, V., Buitink, J., Lin, X., Boudet, J., Hoekstra, F.A., Hundertmark, M., Renard, D. and **Leprince, O**. (2010) MtPM25 is an atypical hydrophobic late embryogenesis-abundant protein that dissociates cold and desiccation-aggregated proteins. *Plant, Cell & Environment* 33: 418-430.
Boudet, J., Buitink, J., Hoekstra, F., Rogniaux, H., Larre, C., Satour, P. and Leprince, O. (2006a) Comparative analysis of the heat stable proteome of radicles of Medicago truncatula seeds during germination identifies late embryogenesis abundant proteins associated with desiccation tolerance. *Plant Physiol* 140: 1418 - 1436.
Boudet, J., Buitink, J., Hoekstra, F.A., Rogniaux, H., Larré, C., Satour, P. and Leprince, O. (2006b) Comparative Analysis of the Heat Stable Proteome of Radicles of Medicago truncatula Seeds during Germination Identifies Late Embryogenesis Abundant Proteins Associated with Desiccation Tolerance. *Plant Physiology* 140: 1418-1436.
Boyer, J.S. (1982) Plant productivity and environment. *Science* 218: 443-448.
Boyer, J.S. and Westgate, M.E. (2004) Grain yields with limited water. *Journal of Experimental Botany* 55: 2385-2394.
Bracha-Drori, K., Shichrur, K., Katz, A., Oliva, M., Angelovici, R., Yalovsky, S. and Ohad, N. (2004) Detection of protein–protein interactions in plants using bimolecular fluorescence complementation. *The Plant Journal* 40: 419-427.
Bray, E. (1993) Molecular responses to water deficit. *Plant Physiol* 103: 1035 - 1040.
Bray, E.A. (1997) Plant responses to water deficit. *Trends in Plant Science* 2: 48-54.
Browne, J.A., Dolan, K.M., Tyson, T., Goyal, K., Tunnacliffe, A. and Burnell, A.M. (2004) Dehydration-Specific Induction of Hydrophilic Protein Genes in the Anhydrobiotic Nematode Aphelenchus avenae. *Eukaryotic Cell* 3: 966-975.
Buitink, J. and Leprince, O. (2004) Glass formation in plant anhydrobiotes: survival in the dry state. *Cryobiology* 48: 215-228.
Cacela, C. and Hincha, D.K. (2006) Monosaccharide composition, chain length and linkage type influence the interactions of oligosaccharides with dry phosphatidylcholine membranes. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1758: 680-691.
Campbell, S.A. and Close, T.J. (1997) Dehydrins:genes, proteins,and associations with phenotypic traits. *New Phyto* 137: 61-74.
Campos, F., Zamudio, F. and Covarrubias, A.A. (2006) Two different late embryogenesis abundant proteins from Arabidopsis thaliana contain specific domains that inhibit Escherichia coli growth. *Biochem Biophys Res Commun* 342: 406-413.

Campos, H., Cooper, M., Habben, J.E., Edmeades, G.O. and Schussler, J.R. (2004) Improving drought tolerance in maize: a view from industry. *Field Crops Research* 90: 19-34.
Cattivelli, L. and Bartels, D. (1990) Molecular Cloning and Characterization of Cold-Regulated Genes in Barley. *Plant Physiol* 93: 1504-1510.
Close, T. (1996) Dehydrins: emergence of a biochemical role of a family of plant dehydration proteins. *Physiol Plant* 97: 795 - 803.
Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J* 16: 735-743.
Cohen, A., Plant, Á.L., Moses, M.S. and Bray, E.A. (1991) Organ-Specific and Environmentally Regulated Expression of Two Abscisic Acid-Induced Genes of Tomato. *Plant Physiol* 97: 1367-1374.
Colmenero-Flores, J.M.C., Francisco,Garciarrubio, Alejandro,Covarrubias, Alejandra A. (1997) Characterization of Phaseolus vulgaris cDNA clones responsive to water deficit: identification of a novel late embryogenesis abundant-like protein. *Plant Molecular Biology* 35: 393-405.
Conde, C., Silva, P., Agasse, A., Tavares, R.M., Delrot, S. and Gerós, H. (2007) An Hg-sensitive channel mediates the diffusional component of glucose transport in olive cells. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1768: 2801-2811.
Cortese, M.S., Baird, J.P., Uversky, V.N. and Dunker, A.K. (2005) Uncovering the Unfoldome: Enriching Cell Extracts for Unstructured Proteins by Acid Treatment. *Journal of Proteome Research* 4: 1610-1618.
Cramer, G.R., Urano, K., Delrot, S., Pezzotti, M. and Shinozaki, K. (2011) Effects of abiotic stress on plants: a systems biology perspective. *BMC plant biology* 11: 163.
Csizmók, V., Szollosi, E., Friedrich, P. and Tompa, P. (2006) A Novel Two-dimensional Electrophoresis Technique for the Identification of Intrinsically Unstructured Proteins. *Molecular & Cellular Proteomics* 5: 265-273.
Cuming, A. (1999) LEA proteins. In R Casey, PR Shewry, eds, Seed Proteins. *Kluwer Academic Publishers, Dordrecht, The Netherlands*: 753–780.
Curry, J., Morris, C.F. and Walker-Simmons, M.K. (1991) Sequence analysis of a cDNA encoding a Group 3 LEA mRNA inducible by ABA or dehydration stress in wheat. *Plant Molecular Biology* 16: 1073-1076.
Chakrabortee, S., Boschetti, C., Walton, L.J., Sarkar, S., Rubinsztein, D.C. and Tunnacliffe, A. (2007) Hydrophilic protein associated with desiccation tolerance exhibits broad protein stabilization function. *Proc Natl Acad Sci U S A* 104: 18073-18078.
Chakrabortee, S., Meersman, F., Kaminski Schierle, G.S., Bertoncini, C.W., McGee, B., Kaminski, C.F. and Tunnacliffe, A. (2010) Catalytic and chaperone-like functions in an intrinsically disordered protein associated with desiccation tolerance. *Proceedings of the National Academy of Sciences* 107: 16084-16089.
Chaves, M.M. (1991) Effects of Water Deficits on Carbon Assimilation. *Journal of Experimental Botany* 42: 1-16.
Dai, S., Li, L., Chen, T., Chong, K., Xue, Y. and Wang, T. (2006) Proteomic analyses of Oryza sativa mature pollen reveal novel proteins associated with pollen germination and tube growth. *Proteomics* 6: 2504-2529.
Dalal, M., Tayal, D., Chinnusamy, V. and Bansal, K.C. (2009) Abiotic stress and ABA-inducible Group 4 LEA from Brassica napus plays a key role in salt and drought tolerance. *J Biotechnol* 139: 137-145.
Danyluk, J., Perron, A., Houde, M., Limin, A., Fowler, B., Benhamou, N. and Sarhan, F. (1998) Accumulation of an Acidic Dehydrin in the Vicinity of the Plasma Membrane during Cold Acclimation of Wheat. *The Plant Cell Online* 10: 623-638.

De Cárcer, G., Cerdido, A. and Medina, F.J. (1997) NopA64, a novel nucleolar phosphoprotein from proliferating onion cells, sharing immunological determinants with mammalian nucleolin. *Planta* 201: 487-495.

Delseny, M., Bies-Etheve, N., Carles, C., Hull, G., Vicient, C., Raynal, M., Grellet, F. and Aspart, L. (2001) Late Embryogenesis Abundant (LEA) protein gene regulation during Arabidopsis seed maturation. *Journal of Plant Physiology* 158: 419-427.

Dosztányi, Z., Csizmok, V., Tompa, P. and Simon, I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 21: 3433-3434.

Doyle, J.J. and Doyle, J.L. (1990) Isolation of plant DNA from fresh tissue. *Biological Sciences* 12: 13-15.

Dunker, A.K., Lawson, J.D., Brown, C.J., Williams, R.M., Romero, P., et al. (2001) Intrinsically disordered protein. *Journal of Molecular Graphics and Modelling* 19: 26-59.

Dure, I.L. (2001) Occurrence of a Repeating 11-Mer Amino Acid Sequence Motif in Diverse Organisms. *Protein and Peptide Letters* 8: 115-122.

Dure, L. (1993) A repeating 11-mer amino acid motif and plant desiccation. *Plant J* 3: 363 - 369.

Dure, L., Crouch, M., Harada, J., Ho, T.-H., Mundy, J., Quatrano, R., Thomas, T. and Sung, Z. (1989) Common amino acid sequence domains among the LEA proteins of higher plants. *Plant Mol Biol* 12: 475 - 486.

Dure, L., Greenway, S. and Galau, G. (1981) Developmental biochemistry of cottonseed embryogenesis and germination: changing messenger ribonucleic acid populations as shown by in vitro and in vivo protein synthesis. *Biochemistry* 20: 4162 - 4168.

Egerton-Warburton, L.M., Balsamo, R.A. and Close, T.J. (1997) Temporal accumulation and ultrastructural localization of dehydrins in Zea mays. *Physiologia Plantarum* 101: 545-555.

Eriksson, S.K., Kutzer, M., Procek, J., Grobner, G. and Harryson, P. (2011) Tunable Membrane Binding of the Intrinsically Disordered Dehydrin Lti30, a Cold-Induced Plant Stress Protein. *Plant Cell*.

Figueras, M., Pujal, J., Saleh, A., Save, R., Pages, M. and Goday, A. (2004) Maize Rabl7 overexpression in Arabidopsis plants promotes osmotic stress tolerance. *Annals of Applied Biology* 144: 251-257.

Finkelstein, R., Reeves, W., Ariizumi, T. and Steber, C. (2008) Molecular Aspects of Seed Dormancy. *Annual Review of Plant Biology* 59: 387-415.

Fowler, S.D. and Greenspan, P. (1985) Application of Nile red, a fluorescent hydrophobic probe, for the detection of neutral lipid deposits in tissue sections: comparison with oil red O. *Journal of Histochemistry & Cytochemistry* 33: 833-836.

Gal, T., Glazer, I. and Koltai, H. (2004) An LEA group 3 family member is involved in survival of C. elegans during exposure to stress. *FEBS Lett* 577: 21 - 26.

Galau, G.A., Hughes, D.W. and Dure, L. (1986) Abscisic acid induction of cloned cotton late embryogenesis-abundant (Lea) mRNAs. *Plant Molecular Biology* 7: 155-170.

Galea, C.A., Pagala, V.R., Obenauer, J.C., Park, C.-G., Slaughter, C.A. and Kriwacki, R.W. (2006) Proteomic Studies of the Intrinsically Unstructured Mammalian Proteome. *Journal of Proteome Research* 5: 2839-2848.

Gallardo, K., Job, C., Groot, S.P.C. and Puype, M. (2002) Proteomics of Arabidopsis seed germination. A comparative study of wild-type and geberellin-deficient seeds. . *M Plant Physiol.* 129: 823.

Garay-Arroyo, A., Colmenero-Flores, J., Garciarrubio, A. and Covarrubias, A. (2000) Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit. *J Biol Chem* 275: 5668 - 5674.

Gilles, G.J., Hines, K.M., Manfre, A.J. and Marcotte Jr, W.R. A predicted N-terminal helical domain of a Group 1 LEA protein is required for protection of enzyme activity from drying. *Plant Physiology and Biochemistry* 45: 389-399.
Giordani, T., Natali, L., D'Ercole, A., Pugliesi, C., Fambrini, M., Vernieri, P., Vitagliano, C. and Cavallini, A. (1999) Expression of a dehydrin gene during embryo development and drought stress in ABA-deficient mutants of sunflower (Helianthus annuus L.). *Plant Molecular Biology* 39: 739-748.
Goday, A., Jensen, A.B., Culianez-Macia, F.A., Mar Alba, M., Figueras, M., Serratosa, J., Torrent, M. and Pages, M. (1994) The maize abscisic acid-responsive protein Rab17 is located in the nucleus and interacts with nuclear localization signals. *Plant Cell* 6: 351-360.
Goday, A., Sánchez-Martínez, D., Gómez, J., Domènech, P.P. and Pagès, M. (1988) Gene Expression in Developing Zea mays Embryos: Regulation by Abscisic Acid of a Highly Phosphorylated 23- to 25-kD Group of Proteins. *Plant Physiol* 88: 564-569.
Goldberg, R.B., Barker, S.J. and Perez-Grau, L. (1989) Regulation of gene expression during plant embryogenesis. *Cell* 56: 149-160.
González-Camacho, F. and Medina, F. (2006) The nucleolar structure and the activity of NopA100, a nucleolin-like protein, during the cell cycle in proliferating plant cells. *Histochemistry and Cell Biology* 125: 139-153.
Goyal, K., Pinelli, C., Maslen, S.L., Rastogi, R.K., Stephens, E. and Tunnacliffe, A. (2005) Dehydration-regulated processing of late embryogenesis abundant protein in a desiccation-tolerant nematode. *FEBS Lett* 579: 4093-4098.
Goyal, K., Tisi, L., Basran, A., Browne, J., Burnell, A., Zurdo, J. and Tunnacliffe, A. (2003) Transition from natively unfolded to folded state induced by desiccation in an anhydrobiotic nematode protein. *J Biol Chem* 278: 12977 - 12984.
Grelet, J., Benamar, A., Teyssier, E., Avelange-Macherel, M.-H., Grunwald, D. and Macherel, D. (2005) Identification in Pea Seed Mitochondria of a Late-Embryogenesis Abundant Protein Able to Protect Enzymes from Drying. *Plant Physiology* 137: 157-167.
Guerche, P., Tire, C., de Sa, F.G., De Clercq, A., Van Montagu, M. and Krebbers, E. (1990) Differential Expression of the Arabidopsis 2S Albumin Genes and the Effect of Increasing Gene Family Size. *The Plant Cell Online* 2: 469-478.
Gupta, K., Agarwal, P., Reddy, M. and Jha, B. (2010) SbDREB2A, an A-2 type DREB transcription factor from extreme halophyte Salicornia brachiata confers abiotic stress tolerance in Escherichia coli. *Plant Cell Reports* 29: 1131-1137.
Hand, S., Jones, D., Menze, M. and Witt, T. (2007) Life without water: expression of plant LEA genes by an anhydrobiotic arthropod. *J Exp Zool* 307A: 62 - 66.
Hand, S.C., Menze, M.A., Toner, M., Boswell, L. and Moore, D. (2011) LEA Proteins During Water Stress: Not Just for Plants Anymore. *Annual Review of Physiology* 73: 115-134.
Hara, M., Fujinaga, M. and Kuboi, T. (2005) Metal binding by citrus dehydrin with histidine-rich domains. *J Exp Bot* 56: 2695-2703.
Hara, M., Shinoda, Y., Tanaka, Y. and Kuboi, T. (2009) DNA binding of citrus dehydrin promoted by zinc ion. *Plant Cell Environ* 32: 532-541.
Heinemeyer, J., Eubel, H., Wehmhöner, D., Jänsch, L. and Braun, H.-P. (2004) Proteomic approach to characterize the supramolecular organization of photosystems in higher plants. *Phytochemistry* 65: 1683-1692.
Heyen, B.J., Alsheikh, M.K., Smith, E.A., Torvik, C.F., Seals, D.F. and Randall, S.K. (2002) The Calcium-Binding Activity of a Vacuole-Associated, Dehydrin-Like Protein Is Regulated by Phosphorylation. *Plant Physiology* 130: 675-687.

Higashi, Y., Hirai, M.Y., Fujiwara, T., Naito, S., Noji, M. and Saito, K. (2006) Proteomic and transcriptomic analysis of Arabidopsis seeds: molecular evidence for successive processing of seed proteins and its implication in the stress response to sulfur nutrition. *The Plant Journal* 48: 557-571.

Hoekstra, F.A., Golovina, E.A. and Buitink, J. (2001) Mechanisms of plant desiccation tolerance. *Trends in Plant Science* 6: 431-438.

Holmes-Davis, R., Tanaka, C.K., Vensel, W.H., Hurkman, W.J. and McCormick, S. (2005) Proteome mapping of mature pollen of Arabidopsis thaliana. *Proteomics* 5: 4864-4884.

Honjoh, K.-I., Yoshimoto, M., Joh, T., Kajiwara, T., Miyamoto, T. and Hatano, S. (1995) Isolation and Characterization of Hardening-Induced Proteins in Chlorella vulgaris C-27: Identification of Late Embryogenesis Abundant Proteins. *Plant and Cell Physiology* 36: 1421-1430.

Houde, M., Dhindsa, R.S. and Sarhan, F. (1992) A molecular marker to select for freezing tolerance in Gramineae. *Molecular and General Genetics MGG* 234: 43-48.

Hsing, Y., Chen, Z., Shih, M., Hsieh, J. and Chow, T. (1995) Unusual sequences of group 3 LEA mRNA inducible by maturation or drying in soybean seeds. *Plant Mol Biol* 29: 863-868.

Huang, C.-Y., Chung, C.-I., Lin, Y.-C., Hsing, Y.-I.C. and Huang, A.H.C. (2009) Oil Bodies and Oleosins in Physcomitrella Possess Characteristics Representative of Early Trends in Evolution. *Plant Physiology* 150: 1192-1203.

Hundertmark, M. and Hincha, D.K. (2008) LEA (late embryogenesis abundant) proteins and their encoding genes in Arabidopsis thaliana. *BMC Genomics* 9: 118.

Imai, R., Chang, L., Ohta, A., Bray, E. and Takagi, M. (1996) A lea-class gene of tomato confers salt and freezing tolerance when expressed in Saccharomyces cerevisiae. *Gene* 170: 243 - 248.

Ingram, J. and Bartels, D. (1996) The Molecular basis od dehydratation tolerance in plants.. *Annual Review of Plant Physiology and Plant Molecular Biology* 47: 377-403.

Irar, S., Oliveira, E., Pages, M. and Goday, A. (2006) Towards the identification of late-embryogenic-abundant phosphoproteome in Arabidopsis by 2-DE and MS. *Proteomics* 6 Suppl 1: S175-185.

Ismail, A., Hall, A. and Close, T. (1999) Purification and partial characterization of a dehydrin involved in chilling tolerance during seedling emergence of cowpea. *Plant Physiol* 120: 237 - 244.

Jacyn Baker, C. and Mock, N.M. (1994) An improved method for monitoring cell death in cell suspension and leaf disc assays using evans blue. *Plant Cell, Tissue and Organ Culture* 39: 7-12.

Jensen, A.B., Goday, A., Figueras, M., Jessop, A.C. and Pagès, M. (1998) Phosphorylation mediates the nuclear targeting of the maize Rab17 protein. *The Plant Journal* 13: 691-697.

Jiang, X. and Wang, Y. (2004) β-Elimination Coupled with Tandem Mass Spectrometry for the Identification of in Vivo and in Vitro Phosphorylation Sites in Maize Dehydrin DHN1 Protein. *Biochemistry* 43: 15567-15576.

Kanias, T. and Acker, J. (2006) Mammalian Cell Desiccation: Facing The Challenges. *Cell Preservation Technology* 4: 253-277.

Kasuga, M., Liu, Q., Miura, S., Yamaguchi-Shinozaki, K. and Shinozaki, K. (1999) Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. *Nat Biotechnol.* 17: 287-291.

Kazuoka, T. and Oeda, K. (1994) Purification and Characterization of COR85-Oligomeric Complex from Cold-Acclimated Spinach. *Plant and Cell Physiology* 35: 601-611.

Kikawada, T., Nakahara, Y., Kanamori, Y., Iwata, K., Watanabe, M., McGee, B., Tunnacliffe, A. and Okuda, T. (2006) Dehydration-induced expression of LEA proteins in an anhydrobiotic chironomid. *Biochem Biophys Res Comm* 348: 56 - 61.

Kiyosue, T., Yamaguchi-Shinozaki, K. and Shinozaki, K. (1994) Characterization of Two cDNAs (ERD10 and ERD14) Corresponding to Genes That Respond Rapidly to Dehydration Stress in Arabidopsis thaliana. *Plant and Cell Physiology* 35: 225-231.
Kiyosue, T., Yamaguchi-Shinozaki, K., Shinozaki, K., Higashi, K., Satoh, S., Kamada, H. and Harada, H. (1992) Isolation and characterization of a cDNA that encodes ECP31, an embryogenic-cell protein from carrot. *Plant Molecular Biology* 19: 239-249.
Koag, M.-C., Fenton, R.D., Wilkens, S. and Close, T.J. (2003) The binding of Maize DHN1 to Lipid Vesicles. Gain of Structure and Lipid Specificity. *Plant Physiology* 131: 309-316.
Koag, M.C., Wilkens, S., Fenton, R.D., Resnik, J., Vo, E. and Close, T.J. (2009) The K-segment of maize DHN1 mediates binding to anionic phospholipid vesicles and concomitant structural changes. *Plant Physiol* 150: 1503-1514.
Kovacs, D., Agoston, B. and Tompa, P. (2008) Disordered plant LEA proteins as molecular chaperones. *Plant Signal Behav* 3: 710-713.
Lan, Y., Cai, D. and Zheng, Y. (2005) Expression in Escherichia coli of three different soybean Late embryogenesis abundant (LEA) genes to investigate enhanced stress tolerance. *J. Integr. Plant Biol* 47: 613–621.
Leprince, O. and Buitink, J. (2010) Desiccation tolerance: From genomics to the field. *Plant Science* 179: 554-564.
Lersten, N.R., Czlapinski, A.R., Curtis, J.D., Freckmann, R. and Horner, H.T. (2006) Oil bodies in leaf mesophyll cells of angiosperms: overview and a selected survey. *American Journal of Botany* 93: 1731-1739.
Li, D. and He, X. (2009) Desiccation induced structural alterations in a 66-amino acid fragment of an anhydrobiotic nematode late embryogenesis abundant (LEA) protein. *Biomacromolecules* 10: 1469-1477.
Lin, C. and Thomashow, M. (1992) DNA sequence analysis of a complementary DNA for cold-regulated Arabidopsis gene cor15 and characterization of the COR15 polypeptide. *Plant Physiol* 99: 519 - 525.
Link, A.J., Eng, J., Schieltz, D.M., Carmack, E., Mize, G.J., Morris, D.R., Garvik, B.M. and Yates, J.R. (1999) Direct analysis of protein complexes using mass spectrometry. *Nat Biotechnol.* 7: 676-682.
Liu, Y., Zheng, Y., Zhang, Y., Wang, W. and Li, R. (2010) Soybean PM2 Protein (LEA3) Confers the Tolerance of Escherichia coli and Stabilization of Enzyme Activity Under Diverse Stresses. *Current Microbiology* 60: 373-378.
Maitra, N. and Cushman, J. (1994) Isolation and characterization of a drought-induced soybean cDNA encoding a D95 family late-embryogenesis-abundant protein. *Plant Physiol* 106: 805-806.
Manfre, A., Lanni, L. and Marcotte, W. (2006) The Arabidopsis group 1 Late-embryogenesis-abundant protein ATEM6 is required for normal seed development. *Plant Physiol* 140: 140 - 149.
Manfre, A.J., LaHatte, G.A., Climer, C.R. and Marcotte, W.R. (2009) Seed Dehydration and the Establishment of Desiccation Tolerance During Seed Maturation is Altered in the Arabidopsis thaliana Mutant atem6-1. *Plant and Cell Physiology* 50: 243-253.
Mann, M., Hojrup, P. and Roepstorff, P. (1993) Use of mass spectrometric molecular weight information to identify proteins in sequence databases. *Biological Mass Spectrometry* 22: 338-345.
Marchetti, E., Querzoli, P., Moncharmont, B., Parikh, I., Bagni, A., Marzola, A., Fabris, G. and Nenci, I. (1987) Immunocytochemical Demonstration of Estrogen Receptors by Monoclonal Antibodies in Human Breast Cancer: Correlation with Estrogen Receptor Assay by Dextran-coated Charcoal Method. *Cancer Research* 47: 2508-2513.
Marttila, S., Tenhola, T. and Mikkonen, A. (1996) A barley (Hordeum vulgare L.) LEA3 protein, HVA1, is abundant in protein storage vacuoles. *Planta* 199: 602-611.

Maskin, L., Frankel, N., Gudesblat, G., Demergasso, M.J., Pietrasanta, L.I. and Iusem, N.D. (2007) Dimerization and DNA-binding of ASR1, a small hydrophilic protein abundant in plant tissues suffering from water loss. *Biochemical and Biophysical Research Communications* 352: 831-835.
McCubbin, W., Kay, C. and Lane, B. (1985) Hydrodynamic and optical properties of the wheat germ Em protein. *Can J Biochem Cell Biol* 63: 803 - 811.
Mian, A., Oomen, R.J.F.J., Isayenkov, S., Sentenac, H., Maathuis, F.J.M. and Véry, A.-A. (2011) Over-expression of an Na+and K+permeable HKT transporter in barley improves salt tolerance. *The Plant Journal* 68: 468-479.
Mills, E.N., Jenkins, J., Marigheto, N. and Belton, P.S. (2002) Storage are thermostable. *Biochem Soc* 30: 925.
Moran, R. (1982) Formulae for Determination of Chlorophyllous Pigments Extracted with N,N-Dimethylformamide. *Plant Physiol* 69: 1376–1381.
Mouillon, J. M., Eriksson, S.K. and Harryson, P. (2008) Mimicking the Plant Cell Interior under Water Stress by Macromolecular Crowding: Disordered Dehydrin Proteins Are Highly Resistant to Structural Collapse. *Plant Physiology* 148: 1925-1937.
Mouillon, J.M., Gustafsson, P. and Harryson, P. (2006) Structural Investigation of Disordered Stress Proteins. Comparison of Full-Length Dehydrins with Isolated Peptides of Their Conserved Segments. *Plant Physiology* 141: 638-650.
Murashige, T. and Skoog, F. (1962) A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum* 15: 473-497.
Nakayama, K., Okawa, K., Kakizaki, T., Honma, T., Itoh, H. and Inaba, T. (2007) Arabidopsis Cor15am Is a Chloroplast Stromal Protein That Has Cryoprotective Activity and Forms Oligomers. *Plant Physiology* 144: 513-523.
NDong, C., Danyluk, J., Wilson, K.E., Pocock, T., Huner, N.P.A. and Sarhan, F. (2002) Cold-Regulated Cereal Chloroplast Late Embryogenesis Abundant-Like Proteins. Molecular Characterization and Functional Analyses. *Plant Physiology* 129: 1368-1381.
Niogret, M.F., Culiáñez-Macià, F.A., Goday, A., Albà, M.M. and Pagès, M. (1996) Expression and cellular localization of rab28 mRNA and Rab28 protein during maize embryogenesis. *The Plant Journal* 9: 549-557.
Nylander, M., Svensson, J., Palva, E.T. and Welin, B.V. (2001) Stress-induced accumulation and tissue-specific localization of dehydrins in Arabidopsis thaliana. *Plant Molecular Biology* 45: 263-279.
Oliveira, E., Amara, I., Bellido, D., Odena, M.A., Dominguez, E., Pages, M. and Goday, A. (2007) LC-MSMS identification of Arabidopsis thaliana heat-stable seed proteins: enriching for LEA-type proteins by acid treatment. *J Mass Spectrom* 42: 1485-1495.
Oliver, M.J., Tuba, Z. and Mishler, B.D. (2000) The evolution of vegetative desiccation tolerance in land plants. *Plant Ecology* 151: 85-100.
Oliver, R.P. and Solomon, P.S. (2004) Does the oxidative stress used by plants for defence provide a source of nutrients for pathogenic fungi? *Trends in Plant Science* 9: 472-473.
Omar, A.A., Dekkers, M.G.H., Graham, J.H. and Grosser, J.W. (2008) Estimation of transgene copy number in transformed citrus plants by quantitative multiplex real-time PCR. *Biotechnology Progress* 24: 1241-1248.
Padmanabhan, V., Dias, D.M.A.L. and Newton, R.J. (1997) Expression analysis of a gene family in loblolly pine (Pinus taeda L.) induced by water deficit stress. *Plant Molecular Biology* 35: 801-807.
Pla, M., Goday, A., Vilardell, J., Gómez, J. and Pagès, M. (1989) Differential regulation of ABA-induced 23–25 kDa proteins in embryo and vegetative tissues of the viviparous mutants of maize. *Plant Molecular Biology* 13: 385-394.

Pla, M., Gómez, J., Goday, A. and Pagès, M. (1991) Regulation of the abscisic acid-responsive gene rab28 in maize viviparous mutants. *Mol Gen Genet* 230: 394-400.
Plana, M., Itarte, E., Eritja, R., Goday, A., Pagès, M. and Martínez, M.C. (1991) Phosphorylation of maize RAB-17 protein by casein kinase 2. *Journal of Biological Chemistry* 266: 22510-22514.
Preuten, T., Cincu, E., Fuchs, J., Zoschke, R., Liere, K. and Börner, T. (2010) Fewer genes than organelles: extremely low and variable gene copy numbers in mitochondria of somatic plant cells. *The Plant Journal* 64: 948-959.
Puhakainen, T., Hess, M., Makela, P., Svensson, J., Heino, P. and Palva, E. (2004) Overexpression of multiple dehydrin genes enhances tolerance to freezing stress in Arabidopsis. *Plant Mol Biol* 54: 743 - 753.
Qin, F., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2011) Achievements and Challenges in Understanding Plant Abiotic Stress Responses and Tolerance. *Plant and Cell Physiology* 52: 1569-1582.
Quan, R., M. Shang, H. Zhang, Y. Zhao and J. Zhang (2004) Engineering of enhanced glycine betaine synthesis improves drought tolerance in maize. *Plant Biotechnol. J* 2: 477-486.
Raghavendra, A.S., Gonugunta, V.K., Christmann, A. and Grill, E. (2010) ABA perception and signalling. *Trends in Plant Science* 15: 395-401.
Rahman, L.N., Smith, G.S., Bamm, V.V., Voyer-Grant, J.A., Moffatt, B.A., Dutcher, J.R. and Harauz, G. (2011) Phosphorylation of Thellungiella salsuginea dehydrins TsDHN-1 and TsDHN-2 facilitates cation-induced conformational changes and actin assembly. *Biochemistry* 50: 9587-9604.
Reyes, J., Rodrigo, M.-J., Colmenero-Flores, J., Gil, J.-V., Garay-Arroyo, A., Campos, F., Salamini, F., **Bartels, D**. and Covarrubias, A. (2005) Hydrophilins from distant organisms can protect enzymatic activities from water limitation effects in vitro. *Plant Cell Environ* 28: 709 - 718.
Reynolds, T.L.B., J.Derek. (1993) Abscisic Acid Enhances the Ability of the Desiccation-Tolerant Fern Polypodium virginianum to Withstand Drying. *Journal of Experimental Botany* 44: 1771-1779.
Ried, J.L. and Walker-Simmons, M.K. (1993) Group 3 Late-embryogenesis-abundant protein in Desiccation-Tolerant Seedlings of Wheat (Triticum aestivum L.). *Plant Physiology* 102: 125-131.
Riera, M., Figueras, M., Lopez, C., Goday, A. and Pages, M. (2004) Protein kinase CK2 modulates developmental functions of the abscisic acid responsive protein Rab17 from maize. *Proc Natl Acad Sci U S A* 101: 9879-9884.
Riera, M., Peracchia, G., De Nadal, E., Ariño, J. and Pagès, M. (2001) Maize protein kinase CK2: regulation and functionality of three β regulatory subunits. *The Plant Journal* 25: 365-374.
Rivin, C.J. and Grudt, T. (1991) Abscisic Acid and the Developmental Regulation of Embryo Storage Proteins in Maize. *Plant Physiology* 95: 358-365.
Roberts, J.K., DeSimone, N.A., Lingle, W.L. and Dure III, L. (1993) Cellular Concentrations and Uniformity of Cell-Type Accumulation of Two Lea Proteins in Cotton Embryos. *The Plant Cell Online* 5: 769-780.
Rohrig, H., Schmidt, J., Colby, T., Brautigam, A., Hufnagel, P. and Bartels, D. (2006) Desiccation of the resurrection plant Craterostigma plantagineum induces dynamic changes in protein phosphorylation. *Plant Cell Environ* 29: 1606-1617.
Romero, P., Obradovic, Z., Li, X., Garner, E.C., Brown, C.J. and Dunker, A.K. (2001) Sequence complexity of disordered protein. *Proteins: Structure, Function, and Bioinformatics* 42: 38-48.
Romo, S., Labrador, E. and Dopico, B. (2001) Water stress-regulated gene expression in Cicer arietinum seedlings and plants. *Plant Physiology and Biochemistry* 39: 1017-1026.
Rorat, T., Grygorowicz, W.J., Irzykowski, W. and Rey, P. (2004) Expression of KS-type dehydrins is primarily regulated by factors related to organ type and leaf developmental stage during vegetative growth. *Planta* 218: 878-885.

Rorat, T., Szabala, B., Grygorowicz, W., Wojtowicz, B., Yin, Z. and Rey, P. (2006) Expression of Sk3-type Dehydrin in Transporting Organs is Associated with Cold Acclimation in Solanum Species. *Planta* 224: 205-221.

Rossignol, M., Peltier, J.-B., Mock, H.-P., Matros, A., Maldonado, A.M. and Jorrín, J.V. (2006) Plant proteome analysis: A 2004–2006 update. *Proteomics* 6: 5529-5548.

Sano, M., Tokudome, S., Shimizu, N., Yoshikawa, N., Ogawa, C., et al. (2007) Intramolecular Control of Protein Stability, Subnuclear Compartmentalization, and Coactivator Function of Peroxisome Proliferator-activated Receptor γ Coactivator 1α. *Journal of Biological Chemistry* 282: 25970-25980.

Sauer, M., Paciorek, T., Benkova, E. and Friml, J. (2006) Immunocytochemical techniques for whole-mount in situ protein localization in plants. *Nat. Protocols* 1: 98-103.

Scott , A.C. and Close, T.J. (1997) Dehydrins: genes, proteins, and associations with phenotypic traits. *New Phyto* 137: 61-74.

Schonfeld, M.A., Johnson, R.C., Carver, B.F. and Mornhinweg, D.W. (1988) Water Relations in Winter Wheat as Drought Resistance Indicators. *Crop Sci.* 28: 526-531.

Seo, M., Nambara, E., Choi, G. and Yamaguchi, S. (2009) Interaction of light and hormone signals in germinating seeds. *Plant Molecular Biology* 69: 463-472.

Shen, G., Pang, Y., Wu, W., Miao, Z., Qian, H., Zhao, L., Sun, X. and Tang, K. (2005) Molecular cloning, characterization and expression of a novel jasmonate-dependent defensin gene from Ginkgo biloba. *Journal of Plant Physiology* 162: 1160-1168.

Sheoran, I.S., Pedersen, E.J., Ross, A.R. and Sawhney, V.K. (2009) Dynamics of protein expression during pollen germination in canola (Brassica napus). *Planta* 230: 779-793.

Shevchenko, A., Wilm, M., Vorm, O. and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem.* 68: 850-858.

Shih, M.-D., Hoekstra, F.-A. and Hsing, Y.-I.E (2008) Late embryogenesis abundant proteins. *Adv. Bot. Res* 48: 212–240.

Shih, M., SHsieh, JTsou, CChow, TLin, T Hsing, Y (2004) Gene cloning and characterization of a soybean (Glycine max L.) LEA protein, GmPM16. *Plant Mol Biol* 56: 689 - 703.

Shimizu, T., Kanamori, Y., Furuki, T., Kikawada, T., Okuda, T., Takahashi, T., Mihara, H. and Sakurai, M. (2009) Desiccation-Induced Structuralization and Glass Formation of Group 3 Late Embryogenesis Abundant Protein Model Peptides. *Biochemistry* 49: 1093-1104.

Shinozaki, K. and Yamaguchi-Shinozaki, K. (2007) Gene networks involved in drought stress response and tolerance. *Journal of Experimental Botany* 58: 221-227.

Siddiqui, N.U., Chung, H.-J., Thomas, T.L. and Drew, M.C. (1998) Abscisic Acid-Dependent and -Independent Expression of the Carrot Late-Embryogenesis-Abundant-Class Gene Dc3 in Transgenic Tobacco Seedlings. *Plant Physiology* 118: 1181-1190.

Silhavy, D., Hutvágner, G., Barta, E. and Bánfalvi, Z. (1995) Isolation and characterization of a water-stress-inducible cDNA clone from Solanum chacoense. *Plant Molecular Biology* 27: 587-595.

Siloto, R.M.P., Findlay, K., Lopez-Villalobos, A., Yeung, E.C., Nykiforuk, C.L. and Moloney, M.M. (2006) The Accumulation of Oleosins Determines the Size of Seed Oilbodies in Arabidopsis. *The Plant Cell Online* 18: 1961-1974.

Singh, J., Whitwill, S., Lacroix, G., Douglas, J., Dubuc, E., Allard, G., Keller, W. and Schernthaner, J.P. (2009) The use of Group 3 LEA proteins as fusion partners in facilitating recombinant expression of recalcitrant proteins in E. coli. *Protein Expression and Purification* 67: 15-22.

Skriver, K. and Mundy, J. (1990) Gene expression in response to abscisic acid and osmotic stress. *Plant Cell* 2: 503–512.

Solomon, A.S., R.Paperna, I. glazer,I. (2000) Desiccation stress of entomopathogenic nematodes induces the accumulation of a novel heat-stable protein. *Parasitology* 121: 409-416.

Soulages, J., Kim, K., Arrese, E., Walters, C. and Cushman, J. (2003) Conformation of a group 2 late embryogenesis abundant protein from soybean. Evidence of poly (L-proline)-type II structure. *Plant Physiol* 131: 963 - 975.

Soulages, J., Kim, K., Walters, C. and Cushman, J. (2002) Temperature-induced extended helix/random coil transitions in a group 1 late embryogenesis-abundant protein from soybean. *Plant Physiol* 128: 822 - 832.

Stacy, R.A.P., Nordeng, T.W., Culiáñez-Macià, Francisco A. and Aalen, R.B. (1999) The dormancy-related peroxiredoxin anti-oxidant, PER1, is localized to the nucleus of barley embryo and aleurone cells. *The Plant Journal* 19: 1-8.

Svensson, J., Palva, E.T. and Welin, B. (2000) Purification of Recombinant Arabidopsis thaliana Dehydrins by Metal Ion Affinity Chromatography. *Protein Expression and Purification* 20: 169-178.

Swire-Clark, G.A. and Marcotte, W.R. (1999) The wheat LEA protein Em functions as an osmoprotective molecule in Saccharomyces cerevisiae. *Plant Molecular Biology* 39: 117-128.

Szőllősi, E., Házy, E., Szász, C. and Tompa, P. (2007) Large systematic errors compromise quantitation of intrinsically unstructured proteins. *Analytical Biochemistry* 360: 321-323.

Tolleter, D., Jaquinod, M., Mangavel, C., Passirani, C., Saulnier, P., Manon, S., Teyssier, E., Payet, N., Avelange-Macherel, M.-H. and Macherel, D. (2007) Structure and Function of a Mitochondrial Late Embryogenesis Abundant Protein Are Revealed by Desiccation. *The Plant Cell Online* 19: 1580-1589.

Tompa, P. (2002) Intrinsically unstructured proteins. *Trends in Biochemical Sciences* 27: 527-533.

Tompa, P. (2005) The interplay between structure and function in intrinsically unstructured proteins. *FEBS Letters* 579: 3346-3354.

Tompa, P. and Kovacs, D. (2010) Intrinsically disordered chaperones in plants and animalsThis paper is one of a selection of papers published in this special issue entitled “Canadian Society of Biochemistry, Molecular & Cellular Biology 52nd Annual Meeting Protein Folding: Principles and Diseases” and has undergone the Journal's usual peer review process. *Biochemistry and Cell Biology* 88: 167-174.

Towbin, H., Staehelin, T. and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences* 76: 4350-4354.

Tunnacliffe, A., Hincha, D., Leprince, O. and Macherel, D (2010) LEA proteins: versatility of form and function.In Dormancy and Resistance in Harsh Environments. *Springer, Heidelberg*: 91–108.

Tunnacliffe, A. and Lapinski, J. (2003) Resurrecting Van Leeuwenhoek's rotifers: a reappraisal of the role of disaccharides in anhydrobiosis. *Philos Trans R Soc Lond B Biol Sci* 358: 1755-1771.

Tunnacliffe, A., Lapinski, J. and McGee, B. (2005) A putative LEA protein, but no trehalose, is present in anhydrobiotic bdelloid rotifers. *Hydrobiologia* 546: 315 - 321.

Tunnacliffe, A. and Wise, M. (2007) The continuing conundrum of the LEA proteins. *Naturwissenschaften* 94: 791-812.

Tzen, J.T., Lie, G.C. and Huang, A.H. (1992) Characterization of the charged components and their topology on the surface of plant seed oil bodies. *Journal of Biological Chemistry* 267: 15626-15634.

Tzen, J.T.C., Peng, C.-C., Cheng, D.-J., Chen, E.C.F. and Chiu, J.M.H. (1997) A New Method for Seed Oil Body Purification and Examination of Oil Body Integrity Following Germination. *Journal of Biochemistry* 121: 762-768.

Ukaji, N., Kuwabara, C., Takezawa, D., Arakawa, K. and Fujikawa, S. (2001) Cold Acclimation-Induced WAP27 Localized in Endoplasmic Reticulum in Cortical Parenchyma Cells of Mulberry Tree Was Homologous to Group 3 Late-Embryogenesis Abundant Proteins. *Plant Physiology* 126: 1588-1597.

Uversky, V., Gillespie, J. and Fink, A. (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* 41: 415 - 427.
Vilardell, J., Goday, A., Freire, M.A., Torrent, M., Martínez, M.C., Torne, J.M. and Pagès, M. (1990) Gene sequence, developmental expression, and protein phosphorylation of RAB-17 in maize. *Plant Molecular Biology* 14: 423-432.
Vinocur, B. and Altman, A. (2005) Recent advances in engineering plant tolerance to abiotic stress: achievements and limitations. *Current Opinion in Biotechnology* 16: 123-132.
Wang, K. and Frame, B. (2009) Biolistic gun-mediated maize genetic transformation. *Chapter 3 in Methods in Molecular Biology: Transgenic Maize* 526: 29-45.
Wang, X.-S., Zhu, H.-B., Jin, G.-L., Liu, H.-L., Wu, W.-R. and Zhu, J. (2007) Genome-scale identification and analysis of LEA genes in rice (Oryza sativa L.). *Plant Sci* 172: 414 - 420.
Warner, A.H., Miroshnychenko, O., Kozarova, A., Vacratsis, P.O., MacRae, T.H., Kim, J. and Clegg, J.S. (2010) Evidence for multiple group 1 late embryogenesis abundant proteins in encysted embryos of Artemia and their organelles. *Journal of Biochemistry* 148: 581-592.
Wasilewska, A., Vlad, F., Sirichandra, C., Redko, Y., Jammes, F., Valon, C., Frey, N.F.d. and Leung, J. (2008) An Update on Abscisic Acid Signaling in Plants and More. *Molecular Plant* 1: 198-217.
Watt, M., Kirkegaard, J.A. and Rebetzke, G.J. (2005) A wheat genotype developed for rapid leaf growth copes well with the physical and biological constraints of unploughed soil. *Functional Plant Biology* 32: 695-706.
Welling, A., Rinne, P., Viherä-Aarnio, A., Kontunen-Soppela, S., Heino, P. and Palva, E.T. (2004) Photoperiod and temperature differentially regulate the expression of two dehydrin genes during overwintering of birch (Betula pubescens Ehrh.). *Journal of Experimental Botany* 55: 507-516.
Wise, M. (2003) LEAping to conclusions: a computational reanalysis of late embryogenesis abundant proteins and their possible roles. *BMC Bioinformatics* 4: 52.
Wise, M.J. and Tunnacliffe, A. (2004) POPP the question: what do LEA proteins do? *Trends in Plant Science* 9: 13-17.
Wolkers, W., McCready, S., Brandt, W., Lindsey, G. and Hoekstra, F. (2001) Isolation and characterization of a D-7 LEA protein from pollen that stabilizes glasses in vitro. *Biochim Biophys Acta* 1544: 196 - 206.
Wolkers, W.F., Bochicchio, A., Selvaggi, G. and Hoekstra, F.A. (1998) Fourier Transform Infrared Microspectroscopy Detects Changes in Protein Secondary Structure Associated with Desiccation Tolerance in Developing Maize Embryos. *Plant Physiology* 116: 1169-1177.
Wolkers, W.F., Tetteroo, F.A.A., Alberda, M. and Hoekstra, F.A. (1999) Changed Properties of the Cytoplasmic Matrix Associated with Desiccation Tolerance of Dried Carrot Somatic Embryos. An in Situ Fourier Transform Infrared Spectroscopic Study. *Plant Physiology* 120: 153-164.
Xiao, B., Huang, Y., Tang, N. and Xiong, L. (2007) Over-expression of a LEA gene in rice improves drought resistance under the field conditions. *Theor Appl Genet* 115: 35-46.
Xiong, L., Schumaker, K.S. and Zhu, J.-K. (2002) Cell Signaling during Cold, Drought, and Salt Stress. *The Plant Cell Online* 14: S165-S183.
Xu, D., Duan, X., Wang, B., Hong, B., Ho, T.-H. and Wu, R. (1996) Expression of a late embryogenesis abundant protein gene, HVA1, from barley confers tolerance to water deficit and salt stress in transgenic rice. *Plant Physiol* 110: 249 - 257.
Yamaguchi-Shinozaki, K. and Shinozaki, K. (2006) Transcriptional regulatory networks in cellular rponses and toleranceto dehydratation and cold stresses. *Annual Review of Plant Biology* 57: 781-803.
Yang, S., Vanderbeld, B., Wan, J. and Huang, Y. (2010) Narrowing Down the Targets: Towards Successful Genetic Engineering of Drought-Tolerant Crops. *Molecular Plant* 3: 469-490.

Yates, J.R., Speicher, S., Griffin, P.R. and Hunkapiller, T. (1993) Peptide Mass Maps: A Highly Informative Approach to Protein Identification. *Analytical Biochemistry* 214: 397-408.
Zegzouti, H., Jones, B., Frasse, P., Marty, C., Maitre, B., Latché, A., Pech, J.-C. and Bouzayen, M. (1999) Ethylene-regulated gene expression in tomato fruit: characterization of novel ethylene-responsive and ripening-related genes isolated by differential display. *The Plant Journal* 18: 589-600.
Zienkiewicz, K., Castro, A.J., Alché, J.d.D., Zienkiewicz, A., Suárez, C. and Rodríguez-García, M.I. (2010) Identification and localization of a caleosin in olive (Olea europaea L.) pollen during in vitro germination. *Journal of Experimental Botany* 61: 1537-1546.

ANNEX