Universitat de Girona

RICE SEEDS AS BIOFACTORIES FOR THE PRODUCTION OF ANTIMICROBIAL PEPTIDES

Laura Montesinos Barreda

Dipòsit legal: Gi. 1052-2014
http://hdl.handle.net/10803/135054

Universitat de Girona

PhD Thesis

Rice seeds as biofactories for the production of antimicrobial peptides

Laura Montesinos Barreda

Girona, 2014

Universitat de Girona

PhD Thesis

Rice seeds as biofactories for the production of antimicrobial peptides

Laura Montesinos Barreda

2014

Programa de Doctorat:

Tecnologia

Dirigida per:

Blanca San Segundo de los Mozos

Esther Badosa Romañó

Emilio Montesinos Seguí

Memòria presentada per optar al títol de doctora per la Universitat de Girona

Laura Montesinos Barreda Dra. Blanca San Segundo de los Mozos

Dra. Esther Badosa Romañó Dr. Emilio Montesinos Seguí

Universitat de Girona

Blanca San Segundo de los Mozos, Professora d'Investigació del CSIC al Centre de Recerca Agrigenòmica (CRAG), **Emilio Montesinos Seguí,** Catedràtic de l'Àrea de Producció Vegetal del Departament d'Enginyeria Química, Agrària i Tecnologia Agroalimentària de la Universitat de Girona, i **Esther Badosa Romañó**, Investigadora contractada del Institut de Tecnologia Agroalimentària de la Universitat de Girona,

DECLAREM:

Que el treball titulat "Rice seeds as biofactories for the production of antimicrobial peptides", que presenta **Laura Montesinos Barreda** per a l'obtenció del títol de doctora per la Universitat de Girona, ha estat realitzat sota la nostra direcció.

I, perquè així consti i tingui els efectes oportuns, signen aquest document a Girona, el 23 de Gener de 2014.

Vist-i-plau directors de la Tesi

Dra. Blanca San Segundo de los Mozos

Dra. Esther Badosa Romañó

Dr. Emilio Montesinos Seguí

Universitat de Girona

Emilio Montesinos Seguí, Catedràtic de l'Àrea de Producció Vegetal del Departament d'Enginyeria Química, Agrària i Tecnologia Agroalimentària de la Universitat de Girona, director del Centre d'Innovació i Desenvolupament en Sanitat Vegetal (CIDSAV), i coordinador dels projectes de recerca "Desarrollo de métodos de control del fuego bacteriano mediante mejora del agente de biocontrol *Pseudomonas fluorescens* EPS62e y nuevos péptidos antimicrobianos sintéticos" (AGL2006-13564-C02-001), y "Safe and Efficient Plant Systems for Antimicrobial Peptide Production- SEPSAPE" (ERA-NET FP7 PLANT-KBBE Ref. EUI2008-03572), en els que es circumscriu la tesi doctoral titulada "Rice seeds as biofactories for the production of antimicrobial peptides" amb el número de registre 1089 del dia 1-12-2008 realitzada per **Laura Montesinos Barreda**,

DECLARA QUE:

Aquesta tesi està sotmesa a la propietat intel·lectual compartida amb els investigadors que participen en els esmentats projectes (Article 2. Apartat 2. RD 1326/2003 de 24-10-2003; Llei de la Propietat Intel·lectual, RD 1/1996 de 12-04-1996).

I perquè consti als efectes oportuns, signa aquest document a Girona, 23 de Gener de 2014.

Vist-i-plau

Dr. Emilio Montesinos Seguí

This work has been supported in part by projects “Desarrollo de métodos de control del fuego bacteriano mediante mejora del agente de biocontrol *Pseudomonas fluorescens* EPS62e y nuevos péptidos antimicrobianos sintéticos" (AGL2006-13564-C02-001), and "Safe and Efficient Plant Systems for Antimicrobial Peptide Production- SEPSAPE" (ERA-NET FP7 PLANT-KBBE Ref. EUI2008-03572).

En el fons, els científics som gent amb sort: podem jugar al que vulguem durant tota la vida

Lee Smolin

Als meus pares

Acknowledgments

Finalment ha arribat el moment d'escriure els agraïments, sens dubte ha estat la part més difícil i alhora la més gratificant. M'agradaria agrair a tots el que heu participat de manera directa o indirecta en aquest procés. Espero no deixar-me a ningú...

Aquesta tesi va partir d'un projecte molt innovador i arriscat, on hi ha hagut molts moments de desesperació i alguns d'enorme satisfacció. Ara, mirant amb perspectiva, estic orgullosa i satisfeta d'haver escollit aquest camí, que acaba amb una tesi i un enorme enriquiment tant a nivell personal com científic. Em sento feliç d'haver format part d'aquest projecte i voldria agrair als meus TRES directors la oportunitat de poder realitzar aquesta tesi doctoral sota la seva direcció.

A la Dra. Blanca San Segundo pel seu gran coneixement. Per els anys que vaig estar al CRAG, començant per el treball de final de carrera, passant pel treball final de Màster i pels primers anys de tesi, i per haver-me donat l'oportunitat de pertànyer a la gran família del "Laboratori Rosa". Moltes gràcies per tot!! Tot el que he aprés amb tu m'ha obert les portes al fascinant món de la genètica molecular vegetal. A més, et vull agrair de manera molt especial la dedicació i el gran esforç que has fet en la correcció d'aquesta tesi. A la Dra. Esther Badosa, que va ser la primera en guiar-me per un laboratori, pel seu suport, per les llargues hores al laboratori, i per ser-hi sempre que l'he necessitat! De veritat, moltes gràcies! Finalment, donar les gràcies al Dr. Emili Montesinos, per donar-me l'oportunitat de formar part d'un grup com el CIDSAV, per la seva confiança i paciència, i sobretot pel seu constant entusiasme i implicació en aquest treball. A tots tres, moltes gràcies per formar-me i donar-me llibertat en aquest projecte!!!!

A la Dra. Maria Coca per tota l'ajuda dins i fora del laboratori, el coneixement, els consells i els ànims rebuts durant aquests anys. Moltes gràcies Maria!!!!

A les Dres. Pilar Fontanet i Gisela Peñas per les primeres transformacions d'arròs, i molt especialment a la Mireia Bundó per tota la implicació en la obtenció i la cura de plantes transgèniques d'arròs en el CRAG.

Moltíssimes gràcies a la Mercè Miquel i a la Mireia Sancenón per les infinites seqüències de DNA analitzades.

A la Mina per totes les converses a primera hora del matí, per tot l'afecte i les abraçades, que són tan necessàries!!

A les *Pink Ladies*, amb qui hem compartit molt bons moments dins i fora del laboratori i que han fet que el dia a dia al laboratori sigui divertidíssim. Per la seva amistat, per tots els atacs de riure i pels *Lady's night*!!! A la Belén, la Christelle, l'Angel, la Paula, la Lídia, la Bia, la Sònia i la Sílvia. Una menció especial a les Drs. Sílvia, Sònia i Bia per tota l'ajuda, disponibilitat, paciència, experiència, protocols i consells que han estat imprescindibles per la realització d'aquesta tesi, gairebé tot el que se ho he aprés de vosaltres!

Un agraïment molt especial al Dr. Emmanuel Guiderdoni per acollir-nos a la plataforma *The Rice Functional Genomics platform (REFUGE)* a Montpellier, i concretament a la Delphine Mieulet per la seva gran hospitalitat i l'experiència tramesa en la generació de plantes transgèniques d'arròs, que posteriorment ens ha servit de tantíssima ajuda!

Als Drs. Michel Rossignol i Esther Izquierdo de la plataforma *Mass Spectrometry Proteomics Platform-MSPP, Laboratoire de Protéomique Fonctionnelle*, Montpellier, per la imprescindible i impecable feina feta en la identificació dels pèptids.

Als membres del LIPSSO d'ara i d'abans, a l'Eduard Bardají, la Marta Planas, la Lídia Feliu, la Marta Domènech, la Montse Talleda, l'Imma Güell i en Roberto, per tota l'ajuda, els bons consells i toooooots els pèptids!

A les noies de tecnologia dels aliments, especialment a la Dra. Maria Pla per tota la seva experiència, amabilitat i savis consells. A les Drs. Anna Nadal i Maria Montero per compartir el BP100 en moments de frustració i també en els bons moments, i per les seqüències sintètiques dels pèptids. A la Nuri Company, per les xerrades entre westerns, pèptids i arrossos.
A tots els membres del CIDSAV d'ara i d'abans, que d'una manera o altre heu contribuït a la realització d'aquesta tesi! A en Josep Pereda, per tot el temps dedicat a l'hivernacle, i a fer quilos i quilos de substrat d'arròs!!! A l'Esther, en Jordi, la Lídia, l'Isa, L'Anna, la Neus i la Bea, per tots els consells i ajuda i sobretot perquè després de tants moments, tantes hores i tants anys de convivència, hem format una gran família.
També agrair a tots els estudiants en pràctiques, perquè una mica d'aquesta tesi es vostra. Especialment a l'Anna Domènech, amb qui vam formar un molt bon equip! Per ser tant treballadora i competent i perquè els arrossos mai han estat tant bonics com quan els cuidaves tu!!!! Moltes gràcies!

Als bioquímics, en especial a la Laia, la Jess, en David, en Jaime i l'Anna, perquè feu de cada trobada un dia memorable i especial! Per les incomptables nits, confidències i viatges, i perquè us estimo!!!!

A la colla d' *Agullana* perquè m'encanta compartir amb vosaltres, cervesetes, dinars, sopars i *lo que surja*. A en Curu, en Víctor, l'Anna, en Gerard, la Cris, la Gemma, en David, en Met. En especial a en Curu i a la Cris perquè sou un sol!! Cris, mil gràcies per la portada de la tesi!!!!

Als gironins (o casi), l'Annabel, en Marc, en Gus, la Xantal, la Gemma, en David, l'Ivan, la Laura, Enrique i Anna, la Carlota, en Ruben, en Xevi, l' Isa, en Gol, l'Anna Q, en Met, la Bàrbara, en Pere, en Lluís i l'Adna per tots els moments viscuts i per ser una gran família!!!! En especial a la Carlota, la Gemma, en Ruben, en Xevi, l' Isa, en Gol i l'Anna Q. A la Dra. Isabel·la, per compartir pràcticament tooooot el dia, dins i fora de la feina. El gimnàs, l'anglès, el pàdel (dit així sembla que realment hi juguem), el córrer, el barri, la feina, el cotxe, el despatx, cervesetes.....Però sobretot pels infinits atacs

de riure, per les aventures i desventures, per ser-hi en els bons i mals moments, i per tota l'ajuda durant la tesi, que ha estat molta.
A la *Dottoressa* Anneta, la meva ànima bessona, la meva germana, què dir-te que no sàpigues!!! Estic molt orgullosa de tu, de tot el camí que has fet i com l'has fet! Perquè m'estimes d'una manera incondicional i perquè SEMPRE, SEMPRE, SEMPRE m'has fet costat en totes les decisions que he pres! Perquè quan recordo els moments més importants de la meva vida, sempre hi estàs present! Pels moments viscuts i els que ens queden per viure!

A en Valentí i a la Maria perquè des del primer dia m'han considerat una més de la família, perquè us desviviu per tots nosaltres, i perquè sabeu gaudir com ningú de la vida! També a la Nuri, l'Alfons, en Marc i en Pol.

A la meva família, pels que hi son i pels que sempre hi seran. Per haver entès i valorat l'esforç que ha suposat aquest treball. Perquè sempre m'he sentit molt estimada i perquè cada dia penso en la gran sort que he tingut amb tots vosaltres. En especial a les meves àvies, les dones més valentes, per qui sento una profunda tendresa i admiració!

I sobretot als meus pares, de qui més he aprés de la vida! Per la seva confiança, pels principis inculcats, per la llibertat, pel seu suport constant, per les ganes de viure, per estimar-me tant! I per ensenyar-me a celebrar les coses que valen la pena, encara que estiguis en un mal moment. No tinc paraules que puguin expressar la meva immensa gratitud. Estic segura que aquest treball us fa tanta il·lusió com a mi. Us estimo i us dedico aquesta tesi, que és tan meva com vostra.

A tu Ferran, perquè et trobo a faltar quan no hi ets, perquè em trobes a faltar quan no hi soc. Perquè ets feliç quan jo ho soc, perquè pateixes si jo pateixo, perquè m'estimes. Fa anys em vas dir que havíem passat 1/5 part de la nostra vida junts i les 4/5 parts restants esperant-nos, jo ara et dic que soc la persona més afortunada del mon per tenir per davant tota una vida al teu costat. Moltes gràcies per la teva paciència, els ànims, i per fer-me feliç! I

no puc deixar-me a en Yuyu, jejeje, per tota la companyia, i *entrebancs* durant les llargues hores d'escriptura de la tesi!

Moltes gràcies a tots

Publications

Badosa, E., Moiset, G., Montesinos, L., Talleda, M., Bardají, E., Feliu, L., Planas, M. and Montesinos, E. 2013. Derivatives of the antimicrobial peptide BP100 for expression in plant systems. Plos One. 8:e85515.

Abbreviations

2,4 D	2,4-dichlorophenoxyacetic acid
µg	Microgram
µL	Microlitre
µm	Micrometre
µM	Micromolar
µmol	Micromol
Aa	Amino acid
ABA	Abscisic acid
AGPA	Stabilization sequence (aa sequence)
AMP	Antimicrobial peptide
AMPder	Antimicrobial peptide derivative (Include BP134 and Cec A derivatives)
ANA	1-Naphthaleneacetic acid
APS	Ammonium persulfate
BAP	6-Benzylaminopurine
bp	Base pair
BP134der	BP134 peptide derivative
CaMV	*Cauliflower mosaic virus*
cDNA	Complementary DNA
CE	Collision energy
Cec A	Cecropin A peptide
CecAder	Cecropin A derivative
cfu/mL	Colony-forming units per millilitre
cm	Centimetre
Ct	Threshold cycle
Da	Dalton
ddH_2O	Double distilled water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNAse	Deoxyribonuclease
DP	Declustering potential

DTT	Dithiothreitol
EB buffer	Elution buffer
EC50	Median effective concentration
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic reticulum
EV	Empty vector
F	Forward sequence
x g	Earth's gravitational acceleration
g	Gram
GFP	Green fluorescence protein
Glb-1	26 kDa α-globulin
GluB-1	Glutelin B-1
GluB-4	Glutelin B-4
GUS	β-Glucuronidase
H	Hemolysis
h	Hour
HPLC	High Performance Liquid Chromatography
hpt II	Selection gene encoding resistance to hygromycin
HRP	Horseradish peroxidase
Ig G	Immunoglobulin G
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kb	Kilobase
kDa	Kilodalton
KDEL	ER retention signal sequence (aa sequence)
L	Litre
LB	Left border
LB broth, agar	Luria-Bertani
LC-MS	Liquid Chromatography–Mass Spectrometry
m	Metre
M	Molar
m/z	Mass divided by charge
mg	Milligram
MIC	Minimal inhibitory concentration
min	Minutes

MKC	Minimal killing concentration
mL	Millilitre
mm	Millimetre
mM	Millimolar
MS	Murashige and Skoog
ng	Nanogram
nm	Nanometre
nM	Nanomolar
Nos-T	*Nopaline synthase* terminator
OB	Oil body
OD	Optical density
Ole18	18 kDa oleosin
PB	Protein body
PB-I	Protein body type I
PB-II	Protein body type I (protein storage vacuole)
PCR	Polimerase chain reaction
PDA	Potato dextrose agar
pGlb-1	*Glb-1* promoter
pGluB-1	*GluB-1* promoter
pGluB-4	*GluB-4* promoter
pI	Isoelectric point
PLs	Phospholipids
pmol	Picomol
pOle18	*Ole18* promoter
prom	Promoter
PRS	TEV protease recognition site
PVDF	Polyvinylidene difluoride
qPCR	Quantitative PCR
R	Reverse sequence
RB	Right border
RNA	Ribonucleic acid
RNAse	Ribonuclease
rpm	Revolution per minute
RT-PCR	Reverse Transcription PCR

s	Second
S-Cec A	Serine-Cec A
SDG	Sucrose density gradient
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SP	N-terminal signal peptide
Std	Molecular standard
TAG	Triacylglycerol
TCA	Trichloroacetic acid
term	Terminator
TEV	Tobacco Etch Virus protease
TFA	Trifluoroacetic acid
Tris	Tris (hydroxymethyl) aminomethane
U	Unit
Ubi-1	*Ubiquitin* promoter
UTR	5'-untranslated region
UV	Ultraviolet
V	Volt
v/v	Volume per volume
w/v	Weight per volume
WT	Wild-type
X-Gal	5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside

Nucleotide base and amino acid abbreviations are indicated according to the IUPAC/IUBMB code (International Union of Pure Applied Chemistry/ International Union of Biochemistry and Molecular Biology).

List of figures

List of tables

Table of contents

SUMMARY

Plant diseases due to plant pathogenic fungi, viruses and bacteria are among the leading causes of crop losses, and are conventionally controlled by treating the plants with synthetic chemical pesticides. Current approaches for crop protection focused on the rational application of pesticides while reducing the number of registered compounds. Antimicrobial peptides (AMPs) emerged as promising candidates to control bacterial and fungal diseases in plants and meet the regulatory requirements about toxicity, environmental impact and biodegradability implemented in many countries. Accordingly, the objective of the present PhD thesis focused on the development of a sustainable platform for the heterologous production of synthetic antimicrobial peptides using rice seeds as biofactory.

Results previously obtained in our group on the antimicrobial activity of linear undecapeptides (cecropin A (1-7) – melittin (2-9) hybrids) obtained by combinatorial chemistry (peptide library CECMEL11), served for the design of AMP sequences and subsequent obtention of their corresponding synthetic genes for plant expression. The BP100 peptide was found to exhibit antimicrobial activity against phytopathogenic bacteria, controlling infection in host plants and, low toxicity and moderate susceptibility to proteolytic degradation. Then, the BP100 peptide (amidated form) was adapted for its biotechnological production using the non-amidated form BP134 (H-KKLFKKILKYL-OH). The heterologous expression of the BP134 peptide required several modifications to allow its expression in transgenic rice plants. Essentially, the modifications made were in the following directions: 1) to increase the length of the BP134 peptide; 2) to include stabilization, retention and/or processing sequences; and 3) to produce an oleosin_AMP fusion protein. Along with this, a BP134 derivative library comprising 52 peptides adapted for plant expression was constructed.

The modifications made on the sequence of the leader peptide had a strong impact in the antibacterial activity and toxicity of the resulting BP134 derivative peptides. Among them, the peptides BP178_KDEL, BP188_KDEL and BP192_KDEL which had optimized features such as strong antibacterial

and bactericidal effects, and low hemolytic and phytotoxic activities, were selected for their production in rice seeds. The antibacterial activity displayed by the selected peptides was comparable to that of antibiotics used for plant disease control (MIC values in the range of 0.6 to 10 µM). The Cec A and Cec A_KDEL was also included in this study as reference peptides because previous reports demonstrated potent antimicrobial activity against plant pathogens and its constitutive expression accomplished in rice plants was found to confer disease resistance.

The rice seed was chosen as a biofactory for the production of the AMP derivatives (BP134 and Cec A derivatives). The use of seeds for the production of recombinant proteins has proven to have important advantages compared to other plant tissues. Indeed, cereal seeds emerged as a convenient platform to produce and accumulate foreign proteins due to their ability to naturally store large amounts of proteins in endoplasmatic reticulum and/or Golgi-derived subcellular compartments, the so called protein bodies (PB) types I and II, respectively. Seeds also offer the possibility to target the foreign protein to the oil bodies (OB) present in the embryo (and aleurone) tissues. Among different cereal seeds, rice was chosen as a plant system because of the high grain yield of rice plants, well-established gene transfer technology, and also because rice is a self-pollinating crop.

Fourteen expression vectors harbouring a codon-optimized *AMP* gene were prepared (*BP178_KDEL*, *BP188_KDEL*, *BP192_KDEL*, *Cec A* and *Cec A_KDEL* genes). In all the cases, a seed-specific promoter was used to drive transgene expression, namely the promoters from the rice *glutelin B-1*, *glutelin B-4*, *26 kDa α-globulin* or *18 kDa oleosin* genes. The various rice promoters used in this work were isolated from rice (*Oryza sativa*) genomic DNA. Transgenic rice plants were produced by *Agrobacterium*-mediated transformation. Most transgenic plants were morphologically similar to wild-type plants, and, even though, some of them showed more vigorous root and shoot systems or higher seed productivity.

Subcellular fractionation in combination with immunodetection and LC-MS/MS analyses confirmed accumulation of the BP178_KDEL, BP192_KDEL,

Cec A and Cec A_KDEL peptides in the expected subcellular compartment, either the protein bodies of the endosperm or the oil bodies of the embryo.
Generally, the expression of the *BP134* and *Cec A* derivative transgenes in rice seeds resulted in a high level of protection to the infection by the bacterial phytopathogen *Dickeya* sp. 1552.10.1. The production of the BP178_KDEL, BP188_KDEL and the Cec A derivatives resulted in a better germination capability of the transgenic seeds in the presence of the fungal pathogen *Fusarium verticillioides*, compared to control plants (wild-type and empty vector controls). These results indicated that the *in planta* produced AMP derivatives were biologically active, and that the accumulation levels achieved were sufficient to confer enhanced protection to infection by the tested plant pathogens. After molecular and phenotypic characterization, homozygous lines were obtained from those lines showing the optimal characteristics.
The antimicrobial peptides were recovered from seeds of transgenic lines. Specifically, the BP178_KDEL peptide was recovered from the protein body-enriched fraction by a single purification step, and the Cec A peptide was partially purified using a Weak Cation Exchange-Reverse Phase column. Based on its electrophoretic mobility, the AMP derivatives accumulating in protein bodies exhibited an apparent molecular weight higher than expected. Most probably, this is caused by multimerization (Cec A peptide) and/or interaction with seed storage proteins present in the protein body (i.e. glutelins).
Regarding to the oleosin_PRS_Cec A fusion, it accumulated in the oil bodies of rice seeds. After proteolytic digestion of the fusion protein with the tobacco etch virus (TEV) protease, the released S-Cec A was purified by Weak Cation Exchange-Reverse Phase chromatography. The resulting S-Cec A displayed antibacterial and bactericidal activity against *Dickeya* sp. 1552.10.1.

The accumulation levels observed in seeds of the rice lines expressing the *Ole18::Ole18_PRS_Cec A* (R9), *GluB-4::Cec A* (R10), *GluB-1::Cec A_KDEL* (MC3) and *Glb-1::BP178_KDEL* (R11) were 10.8, 1.3, 1.4 and 1.8 pmols of the AMP derivative/mg of seed, respectively. Among the different strategies

assayed in the present work for the production of the Cec A peptide in seeds, the highest accumulation level was accomplished when the peptide was produced as an oleosin fusion protein (10 times higher). However, the processing of the oleosin_PRS_Cec A fusion protein was found to be complex and expensive.

Regarding to the production of AMPs accumulating in the endosperm, the BP178_KDEL, Cec A_KDEL and Cec A peptides, accumulated at similar yields for the different endosperm strategies. Interestingly, the attachment of the KDEL retention signal to the amino acid sequence of the Cec A peptide appears not to influence its accumulation level. The production of the BP178_KDEL in rice seed endosperm allowed obtaining a highly purified AMP with minimum purification and processing steps.

Summarizing, the generation of transgenic rice plants expressing *AMPder* genes under the control of a seed-specific promoter offers a promising perspective for their production. Specifically, the endosperm tissue of rice seed appears to be a suitable platform for the production of small and highly cationic AMP derivatives, because they were stable for long periods of time and do not require complex downstream processing steps.

It is expected that, the technology developed in the present work for the production of AMPs using rice seed as biofactories, may provide antimicrobial products in a sustainable platform, suitable applications as plant pesticides, as well as to produce new antimicrobials for pharmaceutical, cosmetics and food industry.

RESUM

Les malalties de les plantes causades per fongs, virus i bacteris són unes de les principals causes de pèrdues dels conreus, i tradicionalment el seu control s'ha basat en la utilització de plaguicides químics sintètics convencionals. Actualment, les aproximacions per a la protecció de cultius contra malalties se centren de manera especial en l'ús racional de plaguicides i en la reducció del número de matèries actives autoritzades. Els pèptids antimicrobians (PAMs) sorgeixen com una de les alternatives més prometedores pel control de malalties de les plantes causades per fongs i bacteris, i compleixen amb els principals requisits reglamentaris implementats en molts països, en quant a toxicitat, impacte mediambiental i biodegradabilitat.

En aquest context, la present tesi doctoral ha tingut com a objectiu el desenvolupament d'una plataforma sostenible per la producció heteròloga de pèptids antimicrobians utilitzant llavors d'arròs com a biofactoria.

Els resultats obtinguts prèviament en el nostre grup pel que fa a l'activitat antimicrobiana d'undecapeptids lineals (híbrids de cecropina A (1-7) – melitina (2-9) dissenyats i sintetitzats mitjançant química combinatòria (quimioteca CECMEL11), van ser utilitzats pel disseny de seqüències de PAMs i per l' obtenció dels corresponents gens sintètics per la seva expressió en plantes. El pèptid BP100 (forma amidada) presenta activitat antimicrobiana enfront bacteris fitopatògens, és capaç d'inhibir el desenvolupament d'infeccions en diverses plantes hoste, i alhora mostra baixa toxicitat i una moderada susceptibilitat a la degradació per proteases. Per aquests motius, el pèptid BP100 va ser el candidat per a una producció biotecnològica, implicant la seva adaptació a la forma no amidada, BP134 (H-KKLFKKILKYL-OH) per a la seva expressió. A més a més, per permetre l'expressió heteròloga del pèptid BP134 en plantes transgèniques d'arròs va ser necessari introduir modificacions a la seqüència. Essencialment, les modificacions realitzades van seguir la següent pauta: 1) augmentar la mida del pèptid; 2) incloure seqüències d'estabilització, retenció i/o seqüències de

processament; i 3) obtenir una proteïna fusió oleosina_PAM. Tenint en compte aquestes modificacions es va preparar una quimioteca de derivats del pèptid BP134 formada per 52 compostos adaptats per la seva expressió en plantes.

Les modificacions introduïdes en la seqüència del pèptid BP134 van mostrar una gran influència en l'activitat antibacteriana i toxicitat dels pèptids obtinguts. Entre els pèptids derivats de BP134, els pèptids BP178_KDEL, BP188_KDEL i BP192_KDEL van presentar característiques optimitzades com ara un potent efecte antibacterià i bactericida, baixa activitat hemolítica i fitotòxica, i per tant van ser seleccionats per ser expressats i produïts en arròs. L'activitat antibacteriana dels pèptids seleccionats era comparable a la mostrada pels antibiòtics utilitzats pel control de malalties de plantes (valors de CMI entre 0,6 i 10 μM). Els pèptids Cec A i Cec A_KDEL també es van incloure en aquest treball com a pèptids de referència donat que resultats previs havien demostrat una potent activitat antimicrobiana enfront patògens de plantes, i a més a més, la seva expressió constitutiva en plantes d'arròs conferia resistència a malalties.

Per a la producció dels PAMs (derivats de BP134 i Cec A) es va escollir la llavor d'arròs, perquè el seu ús per la producció de proteïnes recombinants presenta avantatges importants comparat amb altres teixits de la planta. Actualment, les llavors de cereals emergeixen com una plataforma molt útil per la producció i acumulació de proteïnes recombinants degut a la seva habilitat per emmagatzemar de manera natural gran quantitat de proteïnes en compartiments subcel·lulars derivats del reticle endoplasmàtic i/o Golgi, anomenats cossos proteics (CP) tipus I i II, respectivament. Les llavors també ofereixen la possibilitat de dirigir la proteïna recombinant als cossos oleics (CO) presents en els teixits de l'embrió i aleurona. L'elecció de la llavor d'arròs, respecte a altres llavors de cereals, va ser deguda al seu alt rendiment de gra, la tecnologia ben establerta de transferència de gens, i també per ser un cultiu autopolinitzant.

Un cop obtingudes les seqüències gèniques sintètiques adaptades a l'ús de codons per a la seva expressió en arròs, es varen construir catorze vectors

d'expressió amb els gens dels PAMs seleccionats (*BP178_KDEL*, *BP188_KDEL*, *BP192_KDEL*, *Cec A* i *Cec A_KDEL*) sota el control dels promotors específics de llavor, *glutelina B-1*, *glutelina B-4*, *α-26 kDa globulina* i *18 kDa oleosina*. Els promotors d'arròs utilitzats en aquest treball es van obtenir a partir d'ADN genòmic d'arròs (*Oryza sativa*).

Les plantes transgèniques d'arròs es van obtenir a partir de la transformació mitjançant *Agrobacterium tumefaciens*. La majoria de les plantes transgèniques obtingudes eren morfològicament similars a les plantes control (isogènica convencional Senia o transformada amb el vector buit), tot i que algunes mostraven un sistema radicular i foliar més vigorós, o una major producció de llavors.

El fraccionament subcel·lular juntament amb la immunodetecció i l'anàlisi mitjançant LC-MS/MS van confirmar l'acumulació dels pèptids BP178_KDEL, BP192_KDEL, Cec A i Cec A_KDEL en els compartiments subcel·lulars on s'havien direccionat, és a dir, en els cossos proteics de l'endosperm o en els cossos oleics de l'embrió de la llavor.

De manera general, les llavors transgèniques que expressaven els transgens derivats de *BP134* i *Cec A* mostraven un increment molt significatiu de la resistència a infeccions causades pel bacteri fitopatogen *Dickeya* sp. 1552.10.1. De la mateixa manera, les llavors que produïen els pèptids BP178_KDEL, BP188_KDEL i derivats de Cec A mostraven una major capacitat de germinació en presencia del fong fitopatogen *Fusarium verticillioides*, comparada amb les llavors control (Sènia i transformades amb el vector buit). Aquests resultats indicaven que els pèptids antimicrobians sintetitzats *in planta* eren biològicament actius, i que el nivell d'acumulació obtingut era suficient per conferir una major protecció a infeccions causades pels patògens assajats. Posteriorment, es van obtenir línies homozigotes d'aquelles plantes que van mostrar unes característiques òptimes en la caracterització molecular i fenotípica.

Els pèptids antimicrobians recombinants es van recuperar de les llavors d'arròs transgèniques. Concretament, el pèptid BP178_KDEL es va obtenir a partir de la fracció de cossos proteics mitjançant una única etapa de

purificació, i el pèptid Cec A va ser parcialment purificat mitjançant cromatografia d'intercanvi catiònic i fase reversa. A partir de la seva mobilitat electroforètica, es va poder observar que els derivats dels PAMs acumulats en els cossos proteics mostraven un pes molecular aparent major que l'esperat. Aquest fet es podria explicar per la multimerització (pèptid Cec A) i/o la interacció del pèptid amb proteïnes de reserva de la llavor presents en els cossos proteics (p.e. glutelines).

La proteïna fusió oleosina_PRS_Cec A s'acumulava en els cossos oleics de les llavors d'arròs. Després de la digestió de la proteïna fusió amb la proteasa del virus del gravat del tabac (TEV), el pèptid alliberat S-Cec A va ser purificat mitjançant cromatografia d'intercanvi catiònic i fase reversa. El pèptid obtingut va mostrar activitat antibacteriana i bactericida enfront *Dickeya* sp. 1552.10.1.

Els nivells d'acumulació observats en les llavors d'arròs de les línies que expressaven *Ole18::Ole18_PRS_Cec A* (R9), *GluB-4::Cec A* (R10), *GluB-1::Cec A_KDEL* (MC3) i *Glb-1::BP178_KDEL* (R11) van ser 10.8, 1.3, 1.4 i 1.8 pmols de PAM/mg de llavor, respectivament. Entre les diferents estratègies assajades, la major producció de Cec A es va obtenir mitjançant la seva producció com a proteïna fusió amb la oleosina (10 vegades superior). No obstant, el processament de la proteïna de fusió oleosina_PRS_Cec A va ser complex i costós.

En quant a la producció dels PAMs acumulats en l'endosperm, els pèptids BP178_KDEL i els derivats de Cec A, es van acumular amb rendiments similars en les diferents estratègies assajades. L'addició de la senyal de retenció KDEL a la seqüència aminoacídica del pèptid Cec A no va influir en els seus nivells d'acumulació. La producció del pèptid BP178_KDEL en l'endosperm de les llavors d'arròs va permetre l'obtenció d'un PAM altament purificat a partir d'una extracció i processament mínims.

En resum, l'obtenció de plantes d'arròs que expressen gens de *PAMs* sota el control de promotors específics de llavor, mostra una prometedora plataforma per la seva producció. Concretament, l'endosperm de la llavor d'arròs proporciona un sistema adequat per la producció de pèptids petits i

altament catiònics, degut a que es mantenen estables durant llargs períodes de temps i no requereix etapes de processament complexes.

La tecnologia desenvolupada en el present treball per la producció de PAMs utilitzant llavors d'arròs com a biofactoria, pot proporcionar en un futur proper compostos antimicrobians d'una manera sostenible, adequada per les aplicacions com a plaguicides per la protecció de les plantes, així com de nous antimicrobians pel camp farmacèutic, o de la cosmètica i la indústria alimentària.

RESUMEN

Las enfermedades de las plantas causadas por hongos, virus y bacterias se encuentran entre las principales causas de las pérdidas de cosecha, y son controladas de forma convencional mediante el uso de plaguicidas químicos sintéticos. El enfoque actual de la protección de cultivos contra enfermedades se centra especialmente en el uso racional de los plaguicidas y en la reducción del número de materias activas autorizadas. Los péptidos antimicrobianos (PAMs) ofrecen actualmente un gran potencial para el control de enfermedades de las plantas causadas por hongos y bacterias, y cumplen los principales requisitos reglamentarios de toxicidad, impacto medioambiental y biodegradabilidad implementados en muchos países.

La presente tesis doctoral tuvo como objetivo el desarrollo de una plataforma sostenible para la producción heteróloga de péptidos antimicrobianos utilizando semillas de arroz como biofactoría.

Los resultados obtenidos previamente en nuestro grupo en cuanto a la actividad antimicrobiana de undecapéptidos lineales (híbridos de cecropina A (1-7) – melitina (2-9) obtenidos mediante química combinatoria (quimioteca CECMEL11) se utilizaron para el diseño de secuencias de PAMs y para la obtención de los correspondientes genes sintéticos para su expresión en plantas. El péptido BP100 (forma amidada) presenta actividad antimicrobiana contra bacterias fitopatógenas y controla infecciones en diversas plantas huésped. Asimismo, muestra una baja toxicidad y moderada susceptibilidad a la degradación por proteasas. De este modo, el péptido BP100 fue adaptado para su producción biotecnológica utilizando la forma no amidada, BP134 (H-KKLFKKILKYL-OH). La expresión heteróloga del péptido BP134 requirió varias modificaciones para permitir su expresión en plantas transgénicas de arroz. Esencialmente, las modificaciones hechas siguieron la siguiente pauta: 1) aumento del tamaño del péptido; 2) inclusión de secuencias de estabilización, retención y/o secuencias de procesamiento; y 3) generación de una proteína de fusión oleosina_PAM. De este modo, se

construyó una quimioteca de derivados del péptido BP134 formada por 52 compuestos adaptados para su expresión en plantas.

Las modificaciones hechas en la secuencia del péptido BP134 tuvieron un fuerte impacto en la actividad antibacteriana y toxicidad de los péptidos derivados obtenidos. Los péptidos BP178_KDEL, BP188_KDEL y BP192_KDEL que presentaron características optimizadas tales como un potente efecto antibacteriano y bactericida, y baja actividad hemolítica y fitotóxica, fueron seleccionados para su expresión en semillas de arroz. La actividad antibacteriana que mostraron dichos péptidos era comparable con la mostrada por antibióticos utilizados para el control de enfermedades de plantas (valores de CMI en el rango de 0,6 a 10 μM). Los péptidos Cec A y Cec A_KDEL fueron también incluidos en el presente trabajo como péptidos de referencia debido a que resultados previos habían demostrado una potente actividad antimicrobiana frente a patógenos de plantas y, además, su expresión constitutiva en plantas de arroz confería resistencia a enfermedades.

Para la producción de los PAMs (derivados de BP134 y Cec A) se escogió la semilla de arroz, porque su uso para la producción de proteínas recombinantes presenta importantes ventajas comparado con otros tejidos de la planta. En efecto, las semillas de cereales surgen como una plataforma conveniente para la producción y acumulación de proteínas recombinantes debido a su habilidad para almacenar de forma natural grandes cantidades de proteínas en compartimentos subcelulares derivados del retículo endoplasmático y/o Golgi, los cuerpos proteicos (CP) tipo I y II, respectivamente. Las semillas también ofrecen la posibilidad de dirigir la proteína recombinante hacia los cuerpos oleicos (CO) presentes en el tejido del embrión (y aleurona). Entre las diferentes semillas de cereales, se escogió el arroz debido a su alto rendimiento de grano, la tecnología bien establecida de transferencia de genes, y también por ser un cultivo autopolinizante.

Se prepararon vectores de expresión con los genes de varios PAM (*BP178_KDEL*, *BP188_KDEL*, *BP192_KDEL*, *Cec A* y *Cec A_KDEL*) bajo el

control de los promotores específicos de semilla, *glutelina B-1*, *glutelina B-4*, *α-26 kDa globulina* y *18 kDa oleosina*. La secuencia de ADN de dichos genes sintéticos se adaptó al uso de codones de arroz. Los promotores de arroz utilizados en este trabajo fueron obtenidos a partir de ADN genómico de arroz (*Oryza sativa*). Se generaron las plantas transgénicas a partir de la transformación mediante *Agrobacterium tumefaciens*. La mayoría de las plantas transgénicas obtenidas eran morfológicamente similares a las plantas control (isogénica convencional Senia o transformada con vector vacío), incluso, algunas mostraron un sistema radical y foliar más vigoroso, o una mayor producción de semillas.

El fraccionamiento subcelular combinado con la immunodetección, y el análisis mediante LC-MS/MS confirmaron la acumulación de los péptidos BP178_KDEL, BP192_KDEL, Cec A y Cec A_KDEL en los compartimentos subcelulares esperados (cuerpos proteicos del endospermo o cuerpos oleicos del embrión).

Generalmente, las semillas transgénicas que expresaban transgenes derivados de *BP134* y *Cec A* mostraron un incremento muy significativo de la resistencia a infecciones causadas por la bacteria fitopatógena *Dickeya* sp. 1552.10.1. Asimismo, las semillas que producían los péptidos BP178_KDEL, BP188_KDEL y derivados de Cec A mostraron una mejor capacidad de germinación en presencia del hongo fitopatógeno *Fusarium verticillioides*, comparada con las semillas control (Senia y transformadas con el vector vacío). Estos resultados indicaron que los péptidos antimicrobianos sintetizados *in planta* eran biológicamente activos, y que el nivel de acumulación alcanzado era suficiente para conferir una mayor protección a infecciones causadas por los patógenos de plantas ensayados. Posteriormente, se obtuvieron líneas homocigotas de aquellas plantas que mostraron las características óptimas tras su caracterización molecular y fenotípica.

Los péptidos antimicrobianos fueron recuperados de las semillas de arroz transgénico. Específicamente, el péptido BP178_KDEL fue obtenido a partir de la fracción de cuerpos proteicos mediante una única etapa de

purificación, y el péptido Cec A fue parcialmente purificado mediante cromatografía de intercambio catiónico y fase reversa. A partir de su movilidad electroforética, se pudo observar que los derivados de PAMs acumulados en los cuerpos proteicos mostraban un peso molecular aparente mayor que el esperado. Este hecho se podría explicar por la multimerización (péptido Cec A) y/o la interacción del mismo con proteínas de reserva de la semilla presentes en el cuerpo proteico (p. e. glutelinas).

La proteína fusión oleosina_PRS_Cec A se acumuló en los cuerpos oleicos de las semillas de arroz. Posterior a la digestión de la proteína fusión con la proteasa del virus del grabado del tabaco (TEV), el péptido liberado S-Cec A fue purificado mediante cromatografía de intercambio catiónico y fase reversa. Dicho péptido mostró actividad antibacteriana y bactericida frente a *Dickeya* sp. 1552.10.1.

Los niveles de acumulación observados en las semillas de arroz de las líneas que expresaban *Ole18::Ole18_PRS_Cec A* (R9), *GluB-4::Cec A* (R10), *GluB-1::Cec A_KDEL* (MC3) y *Glb-1::BP178_KDEL* (R11) fueron 10.8, 1.3, 1.4 and 1.8 pmoles de PAM/mg de semilla, respectivamente. Entre las diferentes estrategias ensayadas, la mayor acumulación del péptido Cec A se obtuvo mediante su producción como proteína fusión con la oleosina (10 veces superior). Sin embargo, el procesamiento de la proteína fusión oleosina_PRS_Cec A fue complejo y costoso.

En cuanto a la producción de PAMs acumulados en el endospermo, los péptidos BP178_KDEL y los derivados Cec A, se acumularon con rendimientos similares en las diferentes estrategias ensayadas. La adición de la señal de retención KDEL a la secuencia aminoacídica del péptido Cec A, no influyó en sus niveles de acumulación. La producción del péptido BP178_KDEL en el endospermo de semillas de arroz permitió obtener un PAM altamente purificado a partir de una extracción y procesamiento mínimos.

En resumen, la generación de plantas de arroz que expresan genes de *PAMs* bajo el control de promotores específicos de semilla, ofrece una prometedora perspectiva para su producción. Específicamente, el

endospermo de la semilla de arroz proporciona una plataforma adecuada para la producción de péptidos pequeños y altamente catiónicos, debido a que se mantienen estables durante largos periodos de tiempo y no requiere etapas de procesamiento complejas.

Se espera que en un futuro la tecnología desarrollada en el presente trabajo para la producción de PAMs utilizando semillas de arroz como biofactoría, pueda proporcionar compuestos antimicrobianos de una forma sostenible, adecuada para aplicaciones como plaguicidas para la protección de las plantas, así como de nuevos antimicrobianos para la industria farmacéutica, cosmética e alimentaria.

1. INTRODUCTION

1.1 Antimicrobial peptides in living organisms

Potential crop losses (without crop protection), among the major six crops in the world, are estimated in 50-80% of the attainable yield, whereas actual losses (despite crop protection), represent 28-40% of global agricultural productivity. Plant diseases caused by fungi, viruses and bacteria reaches about 8-21% of actual losses worldwide (according to the Food and Agriculture Organization of the United Nations, FAO; Agrios, 2005; Oerke, 2006). Conventional disease control has been routinely achieved by applying high amounts of synthetic chemical pesticides. Current, approaches are, however, focused on the rational use of pesticides and in a reduction of the number of registered compounds for their use as plant protection products (Directive 2009/128/CE). Furthermore, plant pathogens often develop resistance to fungicides or bactericides that may compromise disease control in several crops. Thus, there is a strong need to develop novel disease control compounds which fulfil the strict regulations about toxicity, environmental impact and biodegradability compiled in national regulations mainly in USA and Europe (Montesinos E. et al., 2012).

Antimicrobial peptides (AMPs) emerged as an effective alternative or complement to conventional antibiotics and fungicides to control bacterial and fungal diseases in order to overcome both the increasing incidence of antibiotic and antifungal resistant pathogens and the antibiotic usage restrictions.

AMPs form an ancient type of innate immunity found in animals and plants, providing a major first-line of defense against the invading pathogens. They have been identified in living organisms ranging from prokaryotes to humans, including bacteria, fungi, plants, insects, fish, amphibians, birds and mammals **(Table 1.1)** (Boman and Steiner, 1981; Steiner et al., 1981; Zasloff, 1987; Andreu and Rivas, 1998; García-Olmedo et al., 1998; Hancock, 2001; Wang and Wang, 2004, 2009; De Smet and Contreras, 2005; Nguyen et al., 2005; Jenssen et al., 2006; Ma et al., 2008; Thomas et al., 2010).

Table 1.1 Relevant representative AMPs produced by living organisms

					Activity against								
Kingdom	**Phylla/Class**	**Peptide**	**Structure**	**Source**	**G^+**	**G^-**	**Fun**	**Vir**	**Cac**	**Par**	**Spe**	**Mac**	**References**
Bacteria	$Gram^-$	Microcin J25	β-sheet (B)	*Escherichia coli AY25*	-	+	-	-	-	-	-	-	Salomón and Farías, 1992
		Serracin-P	Unknown (B)	*Serratia plymuthica*	+	+	-	-	-	-	-	-	Jabrane et al., 2002
		Sclerosin	Unknown (LP, B)	*Pseudomonas sp. strain DF41*	+	-	+	-	-	-	-	-	Berry et al., 2012
	$Gram^+$	Polymyxin B	Non-alphabeta/looped (CLP)	*Bacillus polymyxa*	-	+	-	-	-	-	-	-	Stansly et al., 1947
		Nisin A	Non-alphabeta (L,B)	*Lactococcus lactis*	+	-	-	-	-	-	+	-	Rogers, 1928
		Plantaricin EF, JK	α-helical (B)	*Lactobacillus plantarum C1*	+	+	-	-	-	-	-	-	Anderssen et al., 1998
		Gramicidin S	β-sheet (cyclic)looped	*Bacillus brevis*	+	+	+	-	-	-	-	+	Synge, 1945
		Subtilomycin	Unknown (L)	*Bacillus subtilis MMA7*	+	+	+	-	-	-	-	-	Phelan et al., 2013
Fungi	Ascomycota	AFP	β-sheet (bridge)	*Aspergillus giganteus*	-	-	+	-	-	-	-	-	Wnendt, et al., 1990
		AnAFP	β-sheet (bridge)	*Aspergillus niger*	-	-	+	-	-	-	-	-	Gun Lee et al., 1999
		PAF	β-sheet (bridge)	*Penicillium chrysogenum*	-	-	+	-	-	-	-	-	Kaiserer et al., 2003
		Alamethicin	α-helical (PB)	*Trichoderma viride*	+	-	+	-	-	+	-	+	Meyer and Reusser, 1967
Plantae		MiAMP1	β-sheet	*Macadamia integrifolia*	+	-	+	-	-	-	-	-	Marcus et al., 1997
		MBP-1	α-helical	*Zea mays*	+	+	+	-	-	-	-	-	Duvick et al., 1992
		Ib-AMP1	β-sheet	*Impatiens balsamina*	+	+	+	-	-	-	-	-	Tailor et al., 1997
		Kalata B1	β-sheet (cyclotides)	*Oldenlandia affinis*	+	-	-	+	-	-	-	+	Tam et al., 1999
		β-purothionin	α-helical/β-sheet	*Triticum aestivum*	+	+	+	-	-	-	-	+	Fernandez de Cayela et al., 1972
		LTP110	Unknown	*Oryza sativa*	-	+	+	-	-	-	-	-	Ge et al., 1999
		Ac-AMP2	β-sheet	*Amaranthus caudatus*	+	-	+	-	-	-	-	-	Broekaert et al., 1992
		Ah-AMP1	α-helical/β-sheet (D)	*Aesculus hippocastanum*	+	-	+	-	-	-	-	-	Osborn et al., 1995
		PsD1	α-helical/β-sheet (D)	*Pisum sativum*	-	-	+	-	-	-	-	-	Almeida et al., 2000

Table 1.1 (continued)

					Activity against								
Kingdom	**Phylla/Class**	**Peptide**	**Structure**	**Source**	**G^+**	**G^-**	**Fun**	**Vir**	**Cac**	**Par**	**Spe**	**Mac**	**References**
Animalia	Porifera	Hydramacin-1	α-helical/β-sheet	*Hydra magnipapillata*	+	+	-	-	-	-	-	-	Jung et al., 2009
	Mollusca	Mytilin B	α-helical/β-sheet	*Mytilus edulis*	+	+	-	+	-	-	-	-	Charlet et al., 1996
		MGD-1	α-helical/β-sheet (D)	*Mytilus galloprovincialis*	+	-	-	-	-	-	-	-	Hubert et al., 1996
	Annelida	Hedistin	α-helical	*Nereis diversicolor*	+	+	-	-	-	-	-	-	Tasiemski et al., 2007
	Arthropoda	Cecropin A	α-helical	*Hyalophora cecropia*	+	+	+	+	+	+	-	-	Steiner et al., 1981
		Melittin	α-helical	*Apis mellifera*	+	+	+	+	+	+	-	+	Kreil, 1973
		Heliomicin	α-helical/β-sheet (D)	*Heliothis virescens*	-	-	+	-	-	-	-	-	Casteels et al., 1989
		Drosomycin	α-helical/β-sheet (D)	*Drosophila melanogaster*	-	-	+	-	-	+	-	-	Fehlbaum et al., 1994
		Androctonin	β-sheet	*Androctonus australis*	+	+	+	-	-	-	-	-	Mandard et al., 1999
		Pen4-1	α-helical	*Litopenaeus setiferus*	+	+/-	+	-	-	-	-	-	Cuthberston et al., 2002
		Tachyplesin I	β-sheet	*Tachypleus tridentatus*	+	+	-	+	+	-	-	-	Nakamura et al., 1988
		Gomesin	β-sheet	*Acanthoscurria gomesiana*	+	+	+	-	+	+	-	+	Silva et al., 2000
	Fish	Parasin I	α-helical (HD)	*Parasilurus asotus*	+	+	+	-	-	-	-	-	Park et al., 1998
		Hipposin	Unknown	*Hippoglossus hipoglusus L*	+	+	-	-	-	-	-	-	Birkemo et al., 2003
	Amphibia	Magainin 2	α-helical	*Xenopus laevis*	+	+	+	+	+	+	+	-	Zasloff, 1987
		Buforin II	α-helical (HD)	*Bufo bufo*	+	+	+	-	+	-	-	-	Park et al., 1996
		Esculentin 1	α-helical	*Rana esculenta*	+	+	-	-	-	-	-	+	Simmaco et al., 1993
		Dermaseptin-B2	α-helical	*Phyllomedusa bicolor*	+	+	+	-	+	-	-	-	Mor and Nicolas, 1994
		Temporin A	α-helical	*Rana temporaria*	+	+	-	+	-	+	-	-	Simmaco et al., 1996

Table 1.1 (continued)

Kingdom	Phylla/Class	Peptide	Structure	Source	Activity against								References
					G^+	G^-	Fun	Vir	Cac	Par	Spe	Mac	
	Reptilia	Cathelicidin-BF	α-helical(C)	*Bungarus fasciatus*	+	+	+	-	-	-	-	-	Wang et al., 2008
	Birds	Duck AvBD9	Bridge (D)	*Anas platyrhynchos*	+	+	-	-	-	-	-	-	Liao et al., 2009
		Chicken AvBD1	Bridge (D)	*Gallus gallus*	+	+	-	-	-	-	-	-	Harwig et al., 1994
		Ostrich AvBD1	β-sheet (D)	*Struthio camelus*	+	+	+	-	-	-	-	-	Sugirato et al., 2006
	Mammalia	PR-39	Extended (P-R rich)	*Sus*	+	+	-	-	+	-	-	-	Agerberth et al., 1991
		Lactoferricin B	β-sheet	*Bos Taurus*	+	+	+	+	+	-	-	-	Bellamy et al., 1992
		hBD-2	α-helical/β-sheet (D)	*Homo sapiens*	+	+	+	+	-	-	-	-	Harder et al., 1997
		LEAP-1	β-sheet	*Homo sapiens*	+	+	+	-	-	-	-	-	Krause et al., 2000; Park et al., 2001
		Indolicidin	Extended (W-P-rich) (C)	*Bos Taurus*	+	+	+	+	-	-	-	+	Selsted et al., 1992
		mCRAMP	α-helical (C)	*Mus musculus*	+	+	-	-	-	-	-	-	Gallo et al., 1997
		SMAP-29	α-helical (C)	*Ovis aries*	+	+	+	-	-	-	-	+	Bagella et al., 1995

Antimicrobial peptide classification is based on the APD2: the updated antimicrobial peptide database and its application in peptide design (Wang et al., 2009), Montesinos et al., 2012 and Marcos et al., 2008). G^+, Gram positive bacteria; G^-, Gram negative bacteria; Fun, fungi; Vir, viruses; Cac, Cancer cells; Par, parasites, Spe, sperm cells; Mac, mammal cells. B, bacteriocin; LP, lipopeptide; CLP, cyclic lipopeptide; L, lantibiotic; PB, peptaibol; D, defensins; HD, histone derived; C, cathelicidin.

AMPs are generally defined as peptides of less than 50 amino acid residues which share certain physicochemical characteristics such as net positive charge (multiple arginine and lysine residues) and the ability to adopt an amphipathic structure, which segregates the hydrophilic (≥25%) and the hydrophobic residues (≥50%) to opposite faces of the molecule (Oren and Shai, 1998; Andreu and Rivas, 1998; Hancock, 1999, 2001; Bechinger, 2004; Shai, 2002). Although most peptides are cationic, neutral and anionic peptides have also been reported (Pridmore et al., 1996; Tam et al., 1999; Gun Lee et al., 1999; Berry et al., 2012). AMPs are produced either through ribosomal synthesis (i.e. defensins and small bacteriocins) or non-ribosomal synthesis (peptaibols, cyclic peptides and pseudopeptides) (Montesinos E., 2007) and exhibit either narrow or broad-spectrum of activity against bacteria, fungi, enveloped viruses, parasites and even tumor cells, and generally show low eukaryotic cytotoxicity (Matsuzaki et al., 1995; Hancock and Lehrer, 1998; Yeaman and Yount, 2003; Glukhov et al, 2005). In fact, the ability to target the structural differences between the microbial and plant or animal cell membranes emerged as the most important and exclusive property of AMPs, which in turn results in a low probability of emergence of pathogen resistance compared to conventional antibiotics and antifungals (Yevtushenko and Misra, 2012).

Taking together the aforementioned aspects, AMP properties include: (1) generally high antimicrobial activity (minimal inhibitory concentration (MIC) at micromolar level at the same order than conventional antibiotics), (2) quick response in the target pathogen, (3) high-temperature stability, (4) generally salt tolerance and, (5) biodegradability compared to most pesticides (Destoumieux et al., 2000; Cuthbertson et al., 2006).

More than 1000 AMPs have been reported and can be classified into four major groups based on their structure. The major classes consisted of β-sheet peptides either stabilized by disulphide bonds or by backbone cyclization (cyclic peptides), and unstructured peptides, which adopt amphipathic α-helices upon binding with the membrane **(Fig. 1.1 A-C)**. Less common peptides are those adopting an extended (proline/tryptophan-rich) or loop structure which includes a disulphide bridge **(Fig. 1.1 D, E)**. Peptides consisting of α-helix/β-sheet mixed structures were also reported **(Fig. 1.1 F)**

(Hancock and Lehrer, 1998; Hancock and Diamond, 2000; Monroc et al., 2006a; Montesinos E., 2007). Some cationic peptides are enriched with amino acids residues such as proline, cysteine, histidine and tryptophan. Cysteine-rich peptides can form disulphide bonds that promote its aggregation and therefore have a reduced susceptibility to protease degradation and are the most abundant peptides found in plants (defensins, thionins or lipid transfer proteins) (Marcos et al., 2008). In addition to the natural peptides, thousands of synthetic analogues have been produced and also enclosed into the aforementioned structural groups.

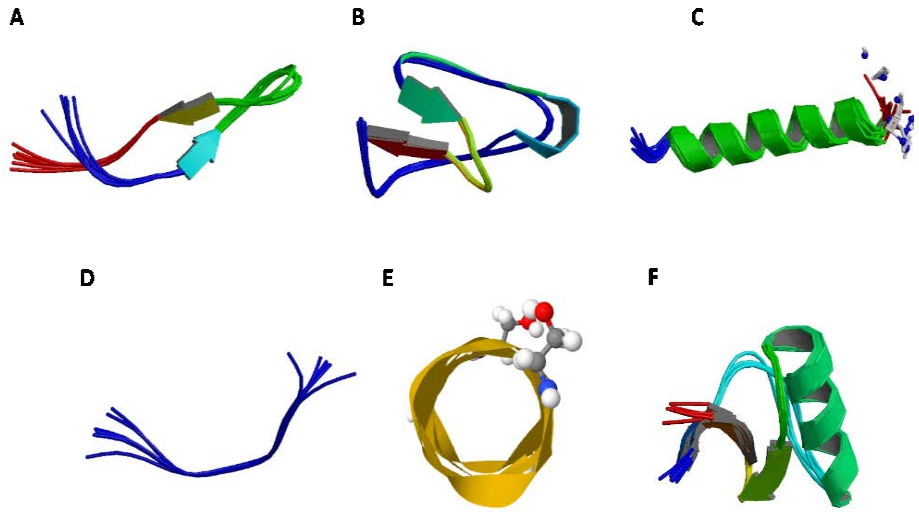


Fig. 1.1 Tertiary structures of representative examples of AMPs from living organisms. (A) β-sheet (Bovine lactoferricin B, Protein Data Bank (PDB) entry *1LFC*), (B) Cyclotides, β-sheet (Plant Kalata B1, PDB entry (*1JJZ*), (C) α-helical (Frog magainin 2, PDB entry *2MAG*), (D) Extended, W-P-rich, (Bovin indolicidin, PDB entry *1G89*), (E) looped, β-sheet (Gramicidin, PDB entry *1MAG*), and (F) α-helical/β-sheet (Plant defensin, PDB entry *1JKZ*).

A widely accepted mode of action of AMPs involves a direct electrostatic interaction between the positively charged residues with the anionic surface of the host/target cell membranes, where peptides adopt an amphipathic conformation (required for the insertion of their hydrophobic face into the membrane) (Yount and Yeaman, 2006; Epand and Epand, 2011). After reaching the threshold concentration, peptides usually destabilize the

membrane, a mechanism for which different models have been proposed: (1) toroidal pore, (2) carpet, (3) barrel stove or aggregate channel model (Ehrenstein and Lecar, 1977; Reddy et al., 2004; Huang, 2006; Giuliani et al., 2007; Melo et al., 2009; Pasupuleti et al., 2012). However, increasing studies support that some peptides kill the host cell without membrane damage by blocking essential intracellular processes after the peptide translocation across the membrane (Zhang et al., 2001; Brogden, 2005; Hancock and Sahl, 2006). In fact, since membrane permeabilization not necessarily results in cell death, increasing evidences support that additional or complementary mechanisms may act together with membrane permeabilization. However, peptide-membrane interaction appears to be the most common killing mechanism (Wu et al., 1999; Toke, 2005).

As AMPs display multiple modes of action, the emergence of resistance would require significant alteration of membrane composition. Along with this, the emergence of resistance mechanisms to AMPs in target pathogens is expected to occur at significantly lower frequencies than for traditional antibiotics.

At present, the exploitation of natural AMPs has shown some limitations as they are either produced in low amounts, or might show toxicity against non-targeted organisms. Alternatively, they might have low activity or instability and generally both extraction and purification from host organisms requires complex and costly procedures (Montesinos E. et al., 2012). Thus, increasing efforts have been made, in terms of rational AMPs design, to create synthetic peptides with improved antimicrobial activity, and low toxicity and sensitivity to protease degradation (Ferre et al., 2006, Badosa et al., 2007; Marcos et al., 2008).

1.2 Synthetic antimicrobial peptides active against plant pathogens

Rational design of AMPs represents a promising strategy to overcome the above mentioned limitations of naturally occurring AMPs and provides the opportunity to develop novel peptides exhibiting more potent antimicrobial activities with less toxicity. Advantages of using antimicrobial peptides for

plant protection include: (1) easy manipulation of AMPs, (2) potent antimicrobial activity, (3) reduced cytotoxicity, and (4) the possibility of altering their specificities. Thus, increasing areas of research are focused on the development of synthetic AMPs such as AMP analogs from natural compounds, chimeric (hybrid) AMPs, mimetic peptides, *de novo* designed AMPs and peptides derived from proteins, to obtain new compounds with improved properties and reduced size **(Table 1.2)**.

Based on the most commonly accepted mode of action, the AMPs activity results from a combination of positive net charge, hydrophobicity and amphipathicity in relation to membrane fluidity and composition in the target cell (Yang et al., 2000; Zalsoff, 2002; Hancock and Sahl, 2006; Wang, 2010; Pasupuleti et al., 2012). Therefore, to control peptide selectivity, peptide design methods focused on the net charge, helicity, hydrophobicity required for the electrostatic interactions along the target membrane, the hydrophobic moment and the influence of the angle subtended by the polar helix face (Montesinos E. et al., 2012). Generally, α-helical peptides are the basis of the design and synthesis of novel antimicrobial peptides since they are widely distributed and exhibit the broadest spectrum of activity.

The generation of synthetic peptides can take advantage of our knowledge on naturally-existing peptides or *de novo* synthesis of peptides. Our group is actively working in the design of new compounds exhibiting antimicrobial activity against the economically important plant pathogens *Erwinia amylovora*, *Pseudomonas syringae* and *Xanthomonas vesicatoria*. Several reports described new peptides with improved biological activity while decreasing hemolytic and phytotoxic activity (Ferre et al., 2006; Monroc et al., 2006a, 2006b; Badosa et al., 2007, 2009; Güell et al., 2011, 2012; Montesinos E. et al., 2012; Badosa et al., 2013).

Solid phase synthesis and combinatorial chemistry methods have provided thousands of new peptides that have been synthesized based on natural-existing peptides. In terms of novel design, additions, deletions and amino acid modifications to natural peptides appears as the first strategy. A second strategy is based on chimeric fusions of natural AMPs which include small truncated sequence containing the minimal domain required for their antimicrobial activity. The well known cecropin, magainin and melittin

peptides have been extensively used as natural models **(Table 1.2)** (Hancock et al., 1992; Jaynes et al., 1993; Chen et al., 2000; Cavallarin et al., 1998; Alan and Earle, 2002; Kamysz et al., 2005; Badosa et al., 2006).

Table 1.2 Representative synthetic antimicrobial peptides active against plant pathogens

Compound	# Aa	Source	Activity	References
Shiva-1	38	Cecropin B	B	Jaynes et al., 1993
SB-37	38	Cecropin B	B	Jaynes et al., 1993
PEP3	11	Cecropin-melittin hybrid	F	Cavallarin et al., 1998
PEP6	6	Synthetic	F	Reed et al., 1997; Ali and Reddy, 2000
BP76	11	Cecropin-melittin hybrid	F+B	Ferré et al., 2006; Badosa et al., 2009
P18	18	Cecropin-magainin hybrid	F+B	Lee et al., 2004
MB-39	39	Cecropin B	F+B	Owens and Heutte., 1997
PEP11	11	Cecropin-melittin hybrid	F+B	Cavallarin et al., 1998; Ali and Reddy, 2000
CEMA	29	Cecropin-melittin hybrid	F+B	Hancock et al., 1992
MsrA1	34	Cecropin-melittin hybrid	F+B	Osusky et al., 2000
MsrA2	31	Dermaseptin B1	F+B	Osusky et al., 2005
MsrA3	19	Temporin A1	F+B	Osusky et al., 2004
D4E1	17	Cecropin B	F+B	DeLucca and Walsh, 1999
MSI-99	23	Magainin	F+B	Alan and Earle, 2002
CAMEL	15	Cecropin-melittin hybrid	B	Kamysz et al., 2005
Rev4	13	Indolicidin	F	Li et al., 2002
10R, 11R	13, 13	Indolicidins	F+B	Bhargava et al., 2007
$LfcinB_{20-25}$	6	Lactoferricin B	F	Muñoz et al., 2006
BP100	11	Cecropin-melittin hybrid	F+B	Badosa et al., 2007; Güell et al., 2012
PAF26	6	Synthetic	F	López-García et al., 2002
PPD1	5	Synthetic	F	Reed et al., 1997
BPC194	10	Synthetic (cyclic)	B	Monroc et al., 2006b
Iseganan	17	Protegrin I	B	Chen et al., 2000
TPY	17	Tachyplesin	F	Rao, 1999
MBG01	19	Rs-AFP1	F	Schaaper et al., 2001
Myp30	24	Magainin	F+B	Li et al., 2001
D32R	46	Thionin	F+B	Vila-Perelló et al., 2003
Ovispirin	29	SMAP-29	B	Kalfa et al., 2001
Pexiganan	22	Magainin	B	Chen et al., 2000

F, antifungal; B, antibacterial; F+B, antifungal and antibacterial

To develop and identify lead peptides with improved biological properties, an initial peptide library must be obtained by solid-phase chemistry. Lead peptides are then identified by testing for their antimicrobial activity against the selected plant pathogens. Subsequently, taking into account the structure and the sequence composition of lead candidates, a second library is synthesized by combinatorial chemistry, where amino acids from selected positions are changed and optimized compounds identified (Montesinos E. et al., 2012). Likewise, peptide libraries can be obtained including peptides representing all possible combinations without sequence restriction. Further screening allows identifying a small collection of novel AMP with improved physicochemical characteristics (Blondelle et al., 1996, 2003; Reed et al., 1997; López-García et al., 2002; Ferré et al., 2006; Monroc et al., 2006a; Badosa et al., 2007).

Several approaches have been done using cecropin::melittin and cecropin::magainin hybrids, in order to maintain or increase their antimicrobial activity while reducing the cytotoxicity effects and their protease susceptibility. For instance, Pep 3 resulted from the juxtaposition of the cecropin A (2-8 residues) and melittin (6-9) and showed increased antimicrobial spectra than the parent cecropin A and less hemolytic activity than melittin (Ali and Reddy, 2000; Cavallarin et al., 1998). Moreover, using Pep 3 as a model, several analogues with improved biological properties were reported. In fact, Ferré and co-workers (2006), obtained Pep 3 derivatives effective against economically important pathogenic bacteria (*Erwinia amylovora*, *Pseudomonas syringae* and *Xanthomonas vesicatoria*) with improved antimicrobial activity and minimized cytotoxicity and protease susceptibility compared to the previous Pep 3. Among them, the peptide BP76 (KKLFKKILKFL-NH_2) exhibits the best biological properties and was considered as potential lead compound in order to be used in plant protection and/or to be expressed in transgenic plants. Moreover, further derivatives of BP76 showed even improved biological properties than the model peptide. Specifically, the peptide BP100 (KKLFKKILKYL-NH_2) showed more potent antimicrobial activity than the previous BP76. Libraries of linear undecapetides were synthesized based on the ideal α-helix wheel diagram, both the Pep 3 and BP76-based libraries. Badosa and co-workers (2006)

reported that slight modifications in charge, helix-forming propensity and an adequate hydrophobicity, induce a slight increase in the overall activity of Pep 3 and BP76. Moreover, amino acid replacement resulted in analogues showing poor/increased antibacterial activity and decreased/increased protease susceptibility. When replaced the two amino acids located in the interface, great changes in activity were observed. Interestingly, the basic N terminus and hydrophobic C terminus seems to influence the antimicrobial activity. The activity of the resulted best analogs is comparable to that of antibiotics used for bacterial diseases control. In other studies, Oh and co-workers (2000) reported that an increase in the peptide hydrophobicity results in an increased cytotoxicity.

Apart from cecropin, melittin and magainin-derived peptides, several analogs of naturally existing-peptides have been shown to display improved biological properties relative to those of their parent peptides (Rao, 1999; Chen et al., 2000; Kalfa et al., 2001; Li et al., 2002; Vila-Perelló et al., 2003; Osusky et al., 2004, 2005; Muñoz et al., 2006).

In an attempt to examine whether cyclation of linear peptides improves their biological properties, Monroc and co-workers (2006b) synthesized a series of *de novo* designed head-to-tail amphipathic cationic cyclic peptides (4 to 10 residues) which clearly improved the biological properties of the parent linear peptides. The hydrophobicity of the non-polar face also modulates the antimicrobial and the hemolytic activity as well as the peptide selectivity of these peptides. Thereby, in some compounds, cyclation involves better stability, resistance to proteolytic degradation and target selectivity.

In the development of AMPs, the identification of best candidates initially requires a screening platform to *in vitro* test their antimicrobial activity. In most cases, antimicrobial activity of synthetic peptides is tested by *in vitro* growth inhibition assays, where designed peptides are tested against selected pathogens. In addition, contact tests, based on the principle of cell killing, are used to determine the pathogen survival at given exposure time. By using these methods, both the Minimal Inhibitory Concentration (MIC) and the Median Effective Concentration (EC50) could be determined and, subsequently, the efficacy of different peptides can be compared. When

hundreds or thousands of new compounds have to be tested, cell viability analyses using vital dyes are commonly used. Particularly, cell membrane permeation produced by peptide interaction is estimated by the dye ability to penetrate through the damaged membranes of the target bacterial or fungal cells. This method is particularly useful since it can be automated and performed by using multi-well plates (micro-plate reader assay). SYTOX green (Badosa et al., 2009; Rathinakumar et al., 2009; Kaplan et al., 2011), berberine (Reyes et al., 2004; De la Fuente-salcido et al., 2007) or 2-nitrophenyl β-D-pyranoside (ONPG) (Friedrich et al., 2000; Hadjicharalambous et al., 2008), are routinely used for membrane permeability experiments. Also, several colorimetric and fluorescent dyes can be used to detect changes in pH medium such as bromophenol blue or fluorescein. It is also noteworthy that these screening methods are highly dependent on peptide solubility, the type of growth of the target pathogen and the medium characteristics (Marcos et al., 2008; Montesinos and Bardají, 2008).

As some AMPs are non-selective between target and eukaryotic non-target cells, cytotoxicity assays are performed to investigate potential adverse effects against plant or animal cells. Cytotoxicity activity is mostly tested by assaying the hemolysis produced by designed peptides towards human red blood cells. Hemolysis is evaluated by comparing newly designed peptides with the highly hemolytic peptide melittin and can be automatized in a microplate reader assay. Phytotoxic activity is mostly examined evaluating the phytotoxicity produced by the synthetic peptides on plant cells (i.e. suspension cells), or by direct application on plant tissues. Regarding assays on tobacco plants, synthetic peptides are applied to the tobacco leaves and the lesion diameter is compared to that of the highly hemolytic peptide melittin (Nadal et al., 2012). In addition, peptides must be tested for susceptibility to degradation by proteases that might be present in living tissues of the target pathogen. Traditionally, commercial proteases such as proteinase K, extracts from plant tissues, or supernatants of cultures of pathogens have been used to evaluate the peptide stability and resistance to proteolytic degradation (Badosa et al., 2007, Marcos et al., 2008). Peptide

hydrolysis can be monitored by HPLC-mass spectrometry or Tris-Tricine-SDS PAGE gel (Shibata et al., 2003; Wei and Bobek, 2005; Makovitzi et al., 2007).

Although synthetic peptides appear as effective compounds to control economically important plant pathogenic diseases, their commercialization is hampered by the high production cost associated with the chemical synthesis of large peptides. By contrast, heterologous synthesis in living systems appears to be a promising tool for the production of functional peptides or proteins (Parachin et al., 2012). However, this approach is only achievable with linear AMPs composed of proteinogenous amino acids.

1.3 Living organisms as heterologous expression systems for recombinant peptides and proteins

Several prokaryotic and eukaryotic expression systems for the production of heterologous proteins or peptides have been described. Unfortunately, no universal expression system can guarantee high yields of the product of interest. Each expression system presents its own advantages and/or disadvantages. Microbial systems have been used as a platform for protein production since are well characterized, often provide high yield production and are easy to handle. However, microbial systems have serious disadvantages such as the inclusion body formation (when recombinant proteins are produced in large amounts, they usually precipitate into insoluble aggregates in inclusion bodies) and the inability to perform the post-translational modifications required for many eukaryotic proteins. The microbial host organisms include bacteria and yeasts, being *Escherichia coli*, *Saccharomyces cerevisiae* and *Pichia pastoris,* as the most commonly used expression systems (Ingham et al., 2007; Mattanovich et al., 2012; Cereghino et al., 2002). In fact, *E. coli* is the most widely used microbial system when no posttranslational protein modifications are required. Genetics (including genome sequence), easy cultivation, fast growth, a well-known physiology and the availability of a large number of cloning vectors make *E.coli* as a suitable factory for commercial recombinant protein production (Li et al., 2010). The first recombinant protein produced in a heterologous system was

insulin which was produced in *E. coli* (Riggs, 1981). On the other hand, yeast is suitable when posttranslational modifications are needed. Compared to other eukaryotic cells, yeast shows a relatively fast growth which together with the complete genome sequence and the ability to secret the recombinant proteins made yeast a good eukaryotic expression system. However, yeast cells are not always convenient since the recombinant product often becomes hyperglycosylated (Harashima, 1994; Malissard et al. 1996). Insect and mammalian cells require complex and expensive procedures and may be easily contaminated by prions, viruses or toxins. The use of transgenic animals presents additional disadvantages such as the slow-scaling up and the ethical concerns (Lico et al., 2005).

1.4 Plant biofactories

Compared to microbial, insect or mammalian cell culture systems, transgenic plants offer several advantages for the large-scale production of high-value recombinant proteins and AMPs (Sharma and Sharma, 2009; Lico et al., 2012). In fact, several therapeutic proteins have been produced using plant-derived platforms and some of them have been commercialized (Hood et al., 1997; Faye and Gomord, 2010). Specifically, the human growth hormone was the first recombinant protein expressed in transgenic tobacco plants for pharmaceutical uses (Barta et al., 1986). Nowadays, several plant-derived pharmaceuticals are in advanced clinical stages of development or on the market (Ramessar et al., 2008; Obembe et al., 2011).

Some advantages of using plants as biofactories are: (1) lower production costs than the other transgenic systems (i.e. mammalian cells), (2) low risk of contamination by human and animal pathogens, oncogenic DNA sequences or endotoxins (derived from Gram-negative bacteria cell envelopes), (3) synthesis of complex proteins correctly folded, (4) the ability to perform most of the posttranslational modifications, (5) the easy scaling up or down, and (6) in most cases, a relatively simple purification technology (Giddings et al., 2000; Hood et al., 2002; Twyman et al., 2003; Horn et al., 2004; Stoger et al., 2005; Lico et al., 2012; Kuo et al., 2013). When using plants as biofactories, different tissues can be chosen to specifically accumulate the

heterologous protein (leaves, seeds, tubers, fruits). Also, strong constitutive promoters are available for many plant species when overall transgene expression is desired. Furthermore, different strategies can be approached to reduce transgene product degradation and toxicity towards plant cells while increasing its stability. This specificity is accomplished by expressing the gene of interest under the control of tissue-specific promoters and/or by designing chimeric genes for targeting the transgene products to specific cell compartments can be achieved using the appropriate targeting signals (i.e. retention into the endoplasmatic reticulum). The system also offers the possibility of producing fusion proteins using specific protease recognition sequences for proteolytic cleavage (van Rooijen and Moloney, 1995; Qu et al., 2004; Takaiwa et al., 2007; Conley et al., 2011; Hofbauer and Stoger, 2013). Thereby, transgenic plants offer multiple possibilities for expression of heterologous genes. Moreover, the relatively low cost product recovery (depending on the selected strategy, less than 85% of other expression systems) and the potential high yield production make plants to be considered as suitable systems for production of AMPs. Although objections for the use of plants as biofactories have been reported, such as long times to product obtention, regulations, posttranslational processing, or low expression levels preventing their commercialization have been reported, major advances have been conducted to overcome these limitations (Hood et al., 2002; Sharma et al., 2004; Sharma and Sharma, 2009; Ahmad et al,. 2010; Obembe et al., 2011; Guan et al., 2013; Vyacheslavova et al., 2012; Viana et al., 2012; Kuo et al., 2013).

Successful heterologous expression of antimicrobial peptides in plants requires overcoming several challenges to avoid the risk of having low expression levels or losing activity due to AMP degradation by host proteases. A minimum size is also needed to be an expressible gene sequence. In addition, AMPs displaying high antimicrobial activity should not show cytotoxicity towards transgenic plant cells (Nadal et al., 2012). These requirements can be accomplished by designing the most appropriated strategy for each type of transgene and plant system by: (1) increasing the size of the peptide (elongation, n-merization), (2) including sequences of stabilization and targeting to specific subcellular compartments, (3)

synthesizing codon-optimized transgenes. The production of fusion proteins that include a specific proteolytic cleavage site might also facilitate subsequent purification of the peptide of interest (Twyman et al., 2003; Sharma and Sharma, 2009; Lico et al., 2012). Specifically, the length of AMPs could be increased above a minimum threshold while maintaining biological properties through any of the following strategies: (1) genetic fusion of the peptide sequence to the sequence of a specific protein (carrier), (2) peptide tandem repeats, and (3) chimeric enlargement. Targeting to, and retention in the endoplasmatic reticulum (ER), can be approached by adding a C-terminal extension (KDEL sequence). This strategy has been reported to decrease proteolytic degradation of transgene products due to either the low proteolytic activity within the lumen or to the higher activity of ER-resident chaperones, while diminishing the potential toxicity to plant cells (Nutall et al., 2002; Benchabane et al., 2008). In addition, peptide production in specific plant tissues such as seeds may decrease the downstream processing costs (Schillberg et al. 1999, 2003).

Taking into account that any change or addition in the sequence of candidate AMPs may affect its spectrum and antimicrobial activities or its toxicity (Badosa et al., 2007, 2009; Montesinos E. et al., 2012), the biological properties of any *de novo* designed sequence must be evaluated prior to be produced in a plant-derived platform.

Several reports have shown the difficulties to express peptides of less than 50 residues in plant cells. The 17 and 19 amino-acid length D4E1 and MsrA3 are among the shortest peptides efficiently expressed in plants (Osusky et al., 2000; Rajasekaran 2005). Okamoto and co-workers (1998) demonstrated that high expression levels of small peptides were accomplished *in planta* when its coding sequence is linked to the coding sequence of the GUS protein (carrier). Similarly, Yang and co-workers (2007) reported higher levels of accumulation of a pollen allergen (Cry j I) when it was expressed as a fusion protein with rice glutelin compared to single *Cry j I* gene expression. Likewise, when the coding regions of two small antimicrobial peptides were fused through a linker peptide, high accumulation levels were observed (François et al., 2002). Also, the 30 amino-acid long glucagon-like peptide-1 (mGLP-1) was directly expressed in rice seed under the control of an

endosperm promoter, but the peptide could not be detected in mature seed. Degradation of the transgene products within the cell host or transgene silencing might explain these results. However, mGPL-1 was successfully produced as a fusion protein with the green fluorescence protein (GFP) in rice seeds (Yasuda et al., 2005). Larger peptides such as 10-repeat novokinin (~11 kDa) or the T cell epitope from the Japanese cedar pollen allergens Cry j I and Cry j II (3 Crp, 5.9 kDa) could not be produced heterologously in rice seeds (Takagi et al., 2008, 2010; Wakasa et al., 2011a).

The aforementioned requirements, along with the proper selection of the promoter to drive tissue specific transgene expression and appropriate compartmentalization of transgene products, are determinants for successful production of AMPs in plants. Additional aspects that need to be considered include post-translational modifications and downstream processing, among others.

1.5 Production of AMPs in transgenic plants: applications in plant protection

As in animals, the genome of plants contains numerous genes encoding antimicrobial peptides (and proteins), that can be used as transgenes for the obtention of transgenic plants with enhanced resistance to plant pathogens (Montesinos E., 2007). Alternatively, antimicrobial genes from organisms other that plants can be also used as transgenes for plant transformation.

The production of AMPs can be accomplished by using different strategies, the stable nuclear transformation using *Agrobacterium tumefaciens* or by particle bombardment being the most commonly used methods. Nuclear transformation requires transferring the transgenes into plant cells, and integration into the host genome in a stable manner. Stable transformation also requires tissue culture steps where single transformed cells are induced to proliferate and differentiate into complete transgenic plants. Plastid or transient transformation are also used as plant production systems

In the literature there are many examples of transgenic plants with enhanced resistance to diseases caused by bacterial and fungal pathogens

(Coca et al., 2006; Montesinos E., 2007; Marcos et al., 2008; López-García et al., 2012). Although numerous transgenic plants expressing both natural and synthetic *AMPs* have been developed with the aim of providing disease resistance, less effort has been paid to the development of strategies for peptide farming and the subsequent use of the *in planta* produced AMPs in crop protection.

Regarding crop protection, some studies focused on the overexpression of own pre-existing *AMP* genes. The resulting transgenic plants, however, showed partial protection which is usually limited to a narrow spectrum of pathogens (López-García et al., 2012). By contrast, other studies oriented to the heterologous production of peptides in transgenic plants provided high levels of protection against a broad spectrum of plant pathogens (Montesinos E., 2007; López-García et al., 2012; Yevtushenko et al., 2012). In recent years, a vast number of natural and synthetic AMPs have been successfully produced in heterologous plant systems, and, their antimicrobial activity against plant pathogens has been shown to be maintained in the transgenic plant. Thereby, the expression level of the *AMPs* in such production-system is sufficient to confer enhanced protection against a variety of plant pathogens **(Table 1.3 and 1.4)**. However, higher heterologous expression levels are required if transgenic plants are intended to be used as a biofactory of AMPs. In fact, Joensuu et al. (2008) described that plant-made recombinant proteins in stable transgenic plants, usually represent <1% of total soluble protein. To produce a plant system economically competitive compared to other heterologous platforms, recombinant proteins should be at least 1% of total soluble protein (Rybicki, 2010).

Table 1.3 Natural AMPs produced in transgenic plants conferring enhanced resistance to pathogens

Peptide	**Source**	**Host**	**Reference**
MGD-1	*Mytilus galloprovincialis*	Tobacco	Arenas et al., 2006
Gamma1-hordothionin	*Hordeum vulgare*	Tobacco	Carmona et al., 1993
Cec A, B	*Hyalophora cecropia, Bombyx mori*	Rice	Coca et al., 2006; Sharma et al., 2000
Cec B	*Hyalophora cecropia*	Tomato	Jan et al., 2010
AFP	*Aspergillus giganteus*	Wheat, rice	Oldach et al., 2001; Coca et al., 2004; Moreno et al., 2005
Drosomycin	*Drosophila melanogaster*	Tobacco	Banzet et al., 2002
Tachyplesin I	*Tachypleus tridentus*	Potato	Allefs et al., 1996
Heliomicin	*Heliothis virescens*	Tobacco	Banzet et al., 2002
Thanatin	*Podisu smaculiventris*	Rice	Imamura et al., 2010
Pen4-1	*Litopenaeus setiferus*	Creeping bentgrass	Zhou et al., 2011
Magainin	*Xenopus laevis*	Tobacco	De Gray et al., 2001
SarcotoxinIA	*Sarcophaga peregrina*	Tobacco	Ohshima et al., 1999
Rs-AFP2	*Brassicaceae species*	Tobacco, tomato, rice	Terras et al., 1995; Jha et al., 2010
alf-AFP	*Medicago sativa*	Potato	Gao et al., 2000
Ace-AMP1	*Allium cepa*	Rice	Roy-Barman et al., 2006
DRR206	*Pisum sativum*	Canola	Wang et al., 1999
hBD-2	*Homo sapiens*	Arabidopsis	Aerts et al., 2007
Temporin A	*Rana temporaria*	Tobacco	Yevtushenko and Misra., 2007
LL-37	*Homo sapiens*	Cabbage	Jung, 2012
SMAP-29	*Ovis aries*	Tobacco	Morassuti et al., 2002
Lactoferricin B	*Bos taurus*	Tobacco	Fukuta et al., 2012
Attacin E	*Hyalophora cecropia*	Tobacco, apple, pear, anthurium	Déstefano-Beltrán, 1991; Ko et al., 2002; Reynoird et al., 1999; Chen and Kuehnle, 1996
Dm-AMP1	*Dalia merckii*	Rice	Jha et al., 2009
BrD1	*Brassica rapa*	Rice	Choi et al., 2009

Table 1.4 Synthetic AMPs produced in transgenic plants conferring partial resistance to pathogens

Peptide	Source	Host	Reference
MsrA3	Temporin A	Potato	Osusky et al., 2004
MSI-99	Magainin 2 analogue	Tobacco, banana, grapevine	Chakrabarti et al., 2003; Vidal et al., 2006
MsrA1	Cecropin-melittin hybrid	Potato	Osusky et al., 2000
MsrA2	Dermaseptin B1	Tobacco, potato, wheat	Osusky et al., 2005; Yevtushenko et al., 2005; Badea et al., 2013
BP100 derivatives	Cecropin-melittin hybrid	Rice	Nadal et al., 2012
SB-37	Cecropin B analogue	Potato, apple, tobacco	Arce et al., 1999
Shiva-1	Cecropin B analogue	Anthurium, Paulownia, tobacco	Kuehnle et al., 2004; Du et al, 2005; Jaynes et al., 1993
MB-39	Cecropin B analogue	Apple	Liu et al., 2001
D4E1	Cecropin B	Tobacco, cotton, poplar	Cary et al., 2000; Rajasekaran et al., 2005; Mentag et al., 2003
Rev4	Indolicidin analogue	Tobacco, arabidopsis	Xing et al., 2006
Esc28L	Esculentin-1	Tobacco	Ponti et al., 2003
Myp30	Magainin analogue	Tobacco	Li et al., 2001
PEP11	Cecropin-melittin hybrid	Tomato	Jones et al., 2004
10R, 11R	Indolicin variants	Tobacco, wheat	Bhargava et al., 2007, Badea et al., 2013
CEMA	Cecropin-melittin hybrid	Potato	Vutto et al., 2010

1.6 Rice-seed as biofactory for production of recombinant proteins and peptides

In recent years, several reports have been focused on the production of recombinant proteins and peptides in plant systems such as leafy crops, cereal and oilseeds, grain legumes, fruits, plant cell cultures, algae, moss or tubers/roots (Twyman et al., 2003). Each production-platform presents unique properties in terms of transformation methods, scalability, accumulation levels, downstream processing, protein stability as well as production costs and time. The seed-based systems provide unique advantages over other plant-based systems. Specifically, seeds emerged as a convenient platform to produce and accumulate proteins/peptides for several reasons: (1) seeds evolved as the natural storage organ in which

storage proteins stably accumulate in specialized storage organelles (2) recombinant products remain stable and functional for long-periods of time at room temperature (due to the dry nature of mature seeds), (3) safety (no contamination of animal virus or prions), (3) production is easily scaled up and down, (4) high expression levels, (5) large biomass, (6) generally, no need of processing, (7) low levels of interfering products such as phenols and alkaloids, (8) accumulation in seeds usually do not interfere with the vegetative growth of the plant, and (9) seeds can be transported to final destination neither requiring refrigeration nor fast processing (Huang, 2004; Takaiwa et al., 2007; Wu et al., 2007; Boothe et al., 2010; Giddings et al., 2000; Stoger et al., 2002, 2005; Nochi et al., 2007; Ramessar et al., 2008). Thus, taking into account the storage ability and the adequate biochemical environment of the endosperm tissue, seeds have been exploited for the production and accumulation of a variety of recombinant products into the specialized storage organs (for example, the protein bodies and protein storage vacuoles) In an economical, sustainable and safe-manner (Delaney, 2002; Stoger et al., 2005; Khan et al., 2012).

Seeds of rice, maize, barley, rye and wheat emerged among the most attractive seed production-systems for recombinant products where they can remain stable for years (Stoger et al., 2005; Ramessar et al., 2008; Kawakatsu et al., 2010; Kuo et al., 2013). Unlike other plant species, the storage function of cereal crops is assumed mainly by the endosperm tissue since represents the major portion of the seed, whereas, by comparison, the embryo is significantly smaller. The cereal seed endosperm contains different types of storage organelles such as protein bodies, starch granules and other minor structures (Juliano, 1985; Shewry et al., 1995). In fact, cereal seeds evolved as a natural storage organ, accumulating carbohydrates, proteins and lipids. Interestingly, cereal crops are the unique species in which two-types of protein bodies (protein bodies and protein storage vacuoles) coexists as separate units in the same cell. Moreover, since cereals are the most important food crops, their physiology and genetics have been extensively studied. Several companies are dedicated to exploit cereal seeds as production platforms of pharmaceutical proteins (Ramessar

et al., 2008). The first commercial plant-derived proteins, avidin (Hood et al., 1997) and β-glucuronidase (Witcher et al., 1998) were obtained using maize seed as a biofactory. Although transgenic maize reaches the highest levels of recombinant protein accumulation, it is a cross-pollinated crop. In fact, an important limitation to the use of a plant as biofactory lies in the risk of out-crosslinking which imposes the need of using containment conditions for plant growth (Montesinos E. and Bardají, 2008). By contrast, rice, wheat and barley are self-pollinating crops. However, wheat and barley present less seed biomass, approximately two to three times lower than rice or maize, respectively. Although the limitation of grain yields, several proteins have been accumulated at high level in barley grains (Shunmann et al., 2002; Erlendsoon et al., 2010; Ventria Biosystems Inc.; ORF Genetics (Orfeus™ system)). Wheat has been used in very few studies, most probably because transformation methods progressed slowly.

Amongst monocots, rice is the most suitable crop for genetic manipulation due to its small genome size, the well-established gene transfer technology available for this cereal specie and easy cropping conditions. Transformed cells can be either used for the obtention of whole transgenic plants or can be maintained as cell cultures which can be upscaled using bioreactors (Shin et al., 2003; Su et al., 2012; Kuo et al., 2013). Also the complete rice (*Oryza sativa*) genome is available for both *japonica* and *indica* subspecies (Goff et al., 2002; Yu et al., 2002). As mentioned above, rice is a self-pollinating crop, reducing the risk of gene flow. In addition, rice seed protein content is lower than in other seed crops, facilitating the product recovery. Notably, rice is one of the most important crops, representing one-half of the global population staple food. In recent years, great advances have been made in obtaining transgenic rice showing increased biotic and abiotic stress tolerance, enhanced nutritional properties, or improved grain yield (Kathuria et al., 2007).

Thereby, rice seed has emerged as a powerful tool platform for large-scale production of recombinant peptides or proteins (Yang et al., 2008; Wang et al., 2013). Along with this, increasing advances have focused on the searching of suitable strong rice seed promoters to produce the recombinant proteins at desired levels in a temporal and spatial restricted-

manner (Wu et al., 1998; Gibbings et al 2003; Qu and Takaiwa, 2004). Constitutive promoters such as *ubiquitin* (*Ubi-1*) are also available for rice. The biotech company, Ventria Bioscience (www.ventria.com), have reported a barley and rice-based platform system for the production of pharmaceutical recombinant proteins. Takaiwa and co-workers (2007) reported that the expression of codon-optimized transgenes allows the stable accumulation of their corresponding products in rice endosperm tissue in large quantities. The success of these strategies is due to the fact that the rice seed endosperm is a specialized storage tissue composed of four main proteins which accumulates stably and in large quantities within specialized storage organelles.

Rice seed storage proteins are mainly composed of dilute-alkali (and/or acid) soluble glutelins, alcohol-soluble prolamins and saline-soluble globulins, being glutelins the major storage protein in rice seeds (up to 80% of protein content) **(Fig. 1.2)**.

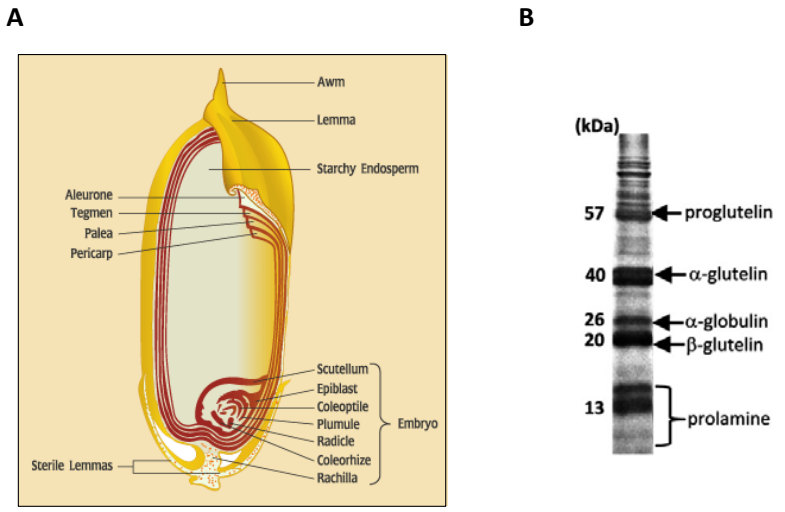


Fig. 1.2 Structure/Anatomy and storage protein composition of a mature rice seed. (A) Tissues in rice seed (http://www.teksengricemill.com/knowled/structure.htm). (B) SDS-PAGE electrophoretic profile of the rice seed storage proteins (Fukuda et al., 2013).

The rice genome contains 15 genes encoding glutelins, whereas prolamins are encoded by 34 genes in the rice genome (Takaiwa et al., 1987; 1991b; Kim et al., 1988; Kawakatsu et al., 2008; Xu and Messing, 2009).

Rice seed storage proteins accumulate in two types of protein bodies localized in the starchy endosperm (Bechtel et al., 1980). Prolamins are synthesized and co-translationally translocated into the ER lumen, and then accumulate in the protein body PB-I (Tanaka et al., 1980; Yamagata and Tanaka, 1986; Nagamine et al., 2011). By contrast, glutelins and α-globulins, which are also synthesized in the ER are targeted and accumulate at the protein storage vacuole (PB-II), via the Golgi apparatus or directly from ER.

Glutelins are synthesized as 57 kDa precursors and then exported to the PB II where they are proteolitically processed into 37-39 kDa acidic (α) and 19-20 kDa basic (β) subunits, interlinked through disulphide bounds (Wen and Luthe, 1985; Tanaka et al, 1980; Yamagata et al., 1982; Yamagata and Tanaka, 1886; Krishnan et al., 1986, Takaiwa et al., 1999). The α and β subunits are further separated into three polypeptide products, designed as α-1, α-2, α-3 for acidic subunit and β-1, β-2 and β-3 for basic subunit (Kagawa et al., 1988). Based on the homology of amino acid sequences, glutelins were classified into four families: Glu A, Glu B, Glu C and Glu D, being type-A and type-B the major subfamilies (Takaiwa et al., 1991; Qu et al., 2008; He et al., 2013).

Rice seed prolamin content is estimated to be around 18-30% of total seed protein and are classified into three groups (10, 13 and 16 kDa prolamin) based on their molecular weight according to electrophoretical mobility on SDS-PAGE gels (Ogawa et al, 1987; Kim and Okita, 1988; Masumura et al., 1989; Hibino et al., 1989; Li et al., 1993; Mitsukawa et al., 1999; Kawakatsu et al., 2008; Saito et al., 2012). It was demonstrated that at least three polypeptides were defined among the 16 kDa prolamin while four to five polypeptides can be distinguished among 13 kDa prolamin (Ogawa et al, 1989; Kumumaru et al., 1988;). Mitsukawa and co-workers (1999) classified 13 kDa prolamins into four groups, based on the amino acid sequence homology and the number of cysteine residues.

Based on their electrophoretic profile and net charge, prolamins are highly variable among different rice cultivars (Ogawa et al., 1987; Shyur et al., 1994; Muench et al., 1999).

Finally, the globulin fraction represents up to 8-12% of total seed protein content (Juliano, 1985; Ju et al., 2001) and is composed of two major

polypeptides of 26 kDa and 16 kDa (Houston et al., 1970; Cagampang et al., 1976; Komatsu and Hirano, 1992). The major globulin in rice seed, the 26 kDa α-globulin, is composed of a single polypeptide chain with an apparent molecular weight, estimated by SDS-PAGE gels, closely to 25-26 kDa (Houston and Mohammad, 1970; Perdon and Juliano, 1978; Pan and Reeck, 1988; Nakase et al., 1996; Ellepola et al., 2005). Interestingly, 26 kDa α-globulin is also known as 19 kDa globulin since the protein size predicted from its amino acid sequence is 19 kDa (Shorrosh et al., 1992; Krishnan and Pueppke, 1993). However, both nomenclatures are currently in use. The slower migration of α-globulins by SDS-PAGE gel have been reported to be associated with some conformational changes related to the SDS-containing buffer (Nakase et al., 1996).

The main physicochemical properties of rice seed storage proteins are compiled in **Table 1.5**.

Table 1.5 Physicochemical properties of major rice seed storage proteins

						Disulphide bonds		Molecular weight (Da)		Theoretical pI[c]		
Family	**Protein**	**Accession[a]**	**# Aa**	**# chains/subunit**	**Protein unit**	**Intrachain**	**Interchain**	**Precursor**	**Subunit[b]**	**Chains**	**Subunit**	**Solubility**
Glutelin	Glutelin type-A1	P07728	499	2; α (25 – 306) β (307 – 499)	Hexamer	46 ↔ 79	122 ↔ 313	56,247	α, 31892.17 β, 21760.61	α, 6.60 β, 9.86	9.10	Dilute acid Alkali
	Glutelin type-A2	P07730	499	2; α (25 – 306) β (307 – 499)	Hexamer	46 ↔ 79	122 ↔ 313	56,306	α, 31904.28 β, 21806.69	α, 6.67 β, 9.76	9.02	Dilute acid Alkali
	Glutelin type-A3	Q09151	496	2; α (25 – 305) β (306 – 496)	Hexamer	45 ↔ 78	121 ↔ 312	56,015	α, 31924.29 β, 21539.56	α, 6.19 β, 9.76	8.80	Dilute acid Alkali
	Glutelin type-B1	P14323	499	2; α (25 – 302) β (303 – 499)	Hexamer	45 ↔ 78	121 ↔ 309	56,551	α, 31756.30 β, 22160.09	α, 8.44 β, 9.69	9.30	Dilute acid Alkali
	Glutelin type-B2	Q02897	495	2; α (25 – 298) β (299 – 495)	Hexamer	45 ↔ 78	121 ↔ 305	56,047	α, 31265.68 β, 22123.97	α, 7.90 β, 9.58	9.14	Dilute acid Alkali
	Glutelin type-B4	P14614	500	2; α (25 – 303) β (304 – 500)	Hexamer	45 ↔ 78	121 ↔ 310	56,818	α, 32037.70 β, 22168.16	α, 7.13 β, 9.58	9.01	Dilute acid Alkali
	Glutelin type-B5	Q6ERU3	500	2; α (25 – 303 β (304 – 500)	Hexamer	45 ↔ 78	121 ↔ 310	56,808	α, 32037.70 β, 22158.12	α, 7.13 β, 9.58	9.01	Dilute acid Alkali
Globulin	α-globulin 19 kDa globulin	Q0DH05 P29835	186	1	Monomer, multimers	Cys residues		21,055	18915.07		6.57	Saline
Prolamin	10 kDa prolamin	Q0DN94	134	1	Monomer, multimers	Cys rich		14,682	12265.81	-	7.62	Alcohol
	13 kDa prolamin 13-I	P17048	156	1	Monomer, multimers	Cys rich		17,805	15827.99	-	8.44	Alcohol
	13 kDa prolamin 13-IIa	Q0DJ45	150	1	Monomer, multimers	No cys residues		16,869	14946.87	-	8.22	Alcohol
	13 kDa prolamin 13-IIb	Q42465	151	1	Monomer, multimers	No cys residues		16,995	15059.00	-	7.02	Alcohol
	13 kDa prolamin 13-III	P20698	156	1	Monomer, multimers	Cys rich		17,562	15591.96	-	8.20	Alcohol
	16 kDa prolamin	Q5Z9M9	149	1	Monomer, multimers	Cys rich		16,675	14787.07	-	8.01	Alcohol

a) Primary accession number; (b) 13 kDa prolamin nomenclature was based on Mitsukawa et al., (1999); (c) Contrarily to theoretical values, Ju et al., (2001) by using turbidity measurements revealed isoelectric points of pH 4.3 and pH 7.9 for globulin fraction (Houston et al., 1970; Perdon and Juliano, 1978; pH 4.5 Pan and Reeck., 1988) and pH 4.8 for glutelin fraction. Physicochemical properties were obtained from UniProt (http://www.uniprot.org/), Gramene (http://www.gramene.org/), and KOME (http://cdna01.dna.affrc.go.jp/cDNA/) databases.

As shown in **Table 1.5**, most seed storage proteins contain cysteine residues. Their cysteine content, and the formation of both intramolecular and intermolecular disulphide bounds, might affect their solubility in non-reducing conditions. In fact, Mitsukawa and co-workers (1999) demonstrated that 13 kDa prolamins were soluble in non-reducing solutions when cysteine amino acids form intramolecular disulphide bonds, while they were not soluble when these polypeptides form intermolecular bindings. In the same way, the 16.6 KDa prolamin low solubility was associated to intermolecular disulphide bounds that promote aggregate formation of the individual polypeptides. Moreover, Kawagoe and co-workers (2005) noted that α-globulin cysteine residues were required for their transport to PB II.
Taking advantage of the storage properties of the rice endosperm, AMPs can be expected to accumulate in stable manner in the rice seed storage organelles. In addition, the stable intracellular environment of PB allows to insulate the heterologous protein from cellular degradation mechanisms while avoiding the enormous stress subjected to the ER (Galili, 2004; Barbante et al., 2008). However, seed protein aggregate formation might negatively interfere in the recombinant peptide purification, since the strongly cationic peptides might interact with the protein aggregates, thus, hindering their subsequent purification.

Although rice endosperm predominantly contain protein bodies and starch as storage organelles, the presence of oil bodies (prevalent in the seeds of oil crops such as safflower, sesame, soybean, rapeseed and sunflower) both in embryo and aleurone layer may offer an additional strategy to target the foreign protein to this organelle which in turn simplifies the recombinant protein isolation (van Rooijen and Moloney, 1995; Moloney, 1997 patent US5650554; Boothe et al., 2010). Oil bodies are small, spherical and discrete intracellular organelles in which a lipid matrix made of triacylglycerols (TAGs) is enclosed by a monolayer of phospholipids and structural membrane proteins. TAGs are used as a nutrient reserve for seed germination and embryo's postgerminative growth. The most abundant rice seed oil body proteins are named oleosins, the 16 kDa (or L) and 18 kDa (or H) oleosins, which play a structural role maintaining the structural integrity and stability

of the oil bodies during seed desiccation. Oleosins cover the whole surface of the OBs (Tzen and Huang, 1992; Chuang et al., 1996). They are alkaline proteins comprising three structural domains: (1) a N-terminal amphipathic domain, (2) a hydrophobic central domain (hairpin), and (3) a C-terminal amphipathic α-helical domain (Tzen and Huang, 1992; Hsieh and Huang, 2004; Chen et al., 2012). The hydrophobic central domain embebs into the TAGs matrix, while the N and C-terminal domain are exposed in the oil body surface, in the cytosol. Oleosins together with TAGs and phospholipids are synthesized in the endoplasmatic reticulum, from which OBs are formed and released by a budding process. Seed oil body, oil body proteins profile as well as a scheme of the oleosin protein are shown in **Fig. 1.3**.

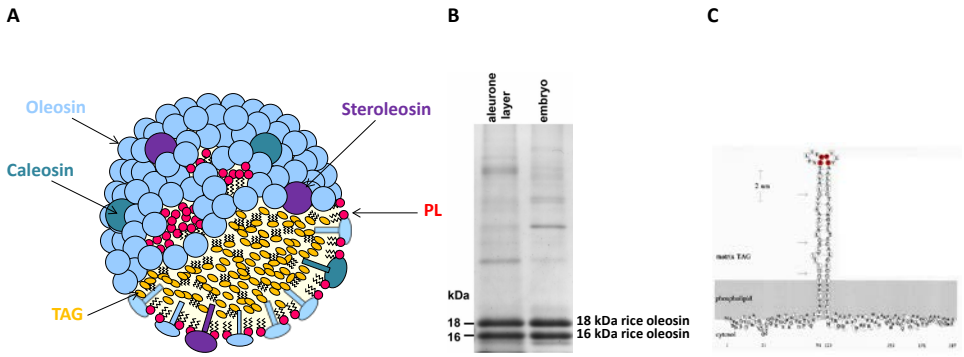


Fig. 1.3 Structure of seed oil bodies. (A) An oil body composed of oleosins and phospholipids (PLs) surrounding the triacylglycerol matrix (TAGs) (Adaptation from Hsieh and Huang, 2004). (B) SDS-PAGE of oil body proteins extracted from rice seeds (Chen et al., 2012). (C) The three structural domains of an oleosin molecule, the N-terminal amphipathic domain, the hydrophobic hairpin central domain, and the C-terminal amphipathic α-helical domain (Plant Lipid Biochemistry, Huang, 2011).

SemBioSys Genetic Inc. (http://www.sembiosys.ca/, not available from November 2013) has exploited the oleosin_fusion platform to produce and accumulate foreign proteins and high value-polypeptides such as insulin and hirudin in the oil body surface of oilseeds (namely safflower, *Carthamus tinctorius*). Recombinant products are expressed in oilseed rape or safflower as a fusion protein and may be isolated by a simple flotation step. A protease recognition site is added between the recombinant protein and

oleosin in order to deliver the recombinant protein from its carrier after protease digestion.

As shown in **Table 1.6**, rice has been successfully used to produce a great variety of valuable recombinant proteins including subunit vaccines, cytokines, antigens, antibodies, hormones or enzymes, among others (Khan et al., 2012; Wang et al., 2013). Notably, most research on AMPs is devoted to exploit the peptide properties for clinical purposes which include the ultimate goal of developing effective human therapeutics (Ramessar et al., 2008; Yevtushenko et al., 2012; Kuo et al., 2013). However, little attention has been paid to the development of seed-based platforms for the large production of AMPs designed to be used as active ingredients for plant protection. In addition, little effort has been made in quantifying the amount of the recombinant AMP produced (Parachin et al., 2012). The vast majority of rice seed-based produced peptides were obtained by expressing fusion genes (linked to the coding sequence of a protein carrier such as storage proteins) **(Table 1.6)**. Heterologous peptides have been also successfully synthesized as multimers, and several reports show the difficulty to express peptides of less than 50 residues in the rice seed endosperm, which generally results in non detection of the expected products. This fact has been routinely associated to the minimum expressible transgene size, the fast degradation of transgene products within the host cell or transgene silencing (Matoba et al., 2001; Yasuda et al., 2005; Takagi et al., 2008, 2010).

When producing a peptide as a fusion to a protein carrier or as a repeated tandem, the production yield should be estimated considering the size of the peptide (molar amounts that are relative to the total size of the expressed recombinant protein). This approach results very useful to compare the resulting peptide yield by using different expression strategies. Accordingly, maximum yields of the relevant recombinant proteins produced in rice seeds range from 0.02 to 608.8 pmols/mg of seed **(Table 1.6)**. Specifically, *TPC7_KDEL*, *Der p 1_KDEL* and *7 Crp_KDEL* engineered under the control of the *GluB-1* promoter accomplished the highest production levels (608.8, 375.0, 272.7 pmols/mg, respectively). The presence of the KDEL retention signal in the coding sequence of the 7 *Crp* transgene clearly increased the heterologous protein accumulation compared with that obtained omitting

the ER retention signal. The lowest production level was found when expressing the chimeric gene *GLUA-4x Type II-collagen* under the control of the *GluA* promoter (0.02 pmols/mg of seed). However, when 3 Crp was intended to be produced directly in rice seed endosperm, no peptide product was detected. By contrast, GLUA-3 Crp and CBT- 3 Crp fusion proteins were successfully produced and accumulated in the endosperm tissue. Regarding 3 Crp heterologous production, a reduction of the accumulation levels was observed when CBT was used as a protein carrier (51.5 pmols/mg of GLUA-3 Crp and 1.2 pmols of CBT-3 Crp/mg). In other studies, the 10 fold repeated-novokinin transgene was intended to be expressed under the control of *glutelin B-4* promoter, but the expected peptide was not detected. However, high amounts of the 18 fold repeated-novokinin accumulated in the rice seed when this larger peptide was expressed under the control of the *GluB-1* promoter. Thus, it seems that the type of promoter, as well as the size of the transgene are critical for a proper expression.

Based on the maximum yields and strategies reported in **Table 1.6**, it becomes clear that no universal expression strategy can guarantee high levels of recombinant proteins or peptides in transgenic rice seeds even if the transgene sequence (codon optimization) and the regulatory regions of the transgene are carefully chosen.

Table 1.6 Relevant recombinant proteins and peptides heterologously produced in transgenic rice seed tissue

Product	Mw (kDa)	Tissue	Localization	Promoter	Strategy	Yield (µg/mg)	pmols/mg seed	References
Recombinant proteins								
ScFv T84.66	27	E+L+C	PB-I, PB-II	*Ubi-1*	KDEL ER	0.0038	0.1	Torres et al., 1999, 2001
HSA	70	E	n.s	*Gt13α*	-	2.75	39.3	He et al., 2011
IL-10	~20	S	n.s	*GluB-1*	6x His+ KDEL ER	0.05	2.5	Fujiwara et al., 2010
TPC7	17 and 20 (glyc)	E	TPC7 body[a]	*GluB-1*	KDEL ER	10.35*	608.8, 517.5	Wang et al., 2013
CTB	12, 15 (r.c); 55-65 (u.c)	E	PB-I, PB-II	*GluB-1*	KDEL ER	1.5*	125.0, 100.0; 27.3	Nochi et al., 2007
hLF	80	S	n.s	*GluB-1*	-	4.5	56.3	Lin et al., 2010
Sasa IFN	18	L+S	n.s	*CaMV 35S*	-	n.s	-	Fukuzawa et al., 2010
rhGM-CSF	18 (a.b. 19-100)	E	PBs, A	*Gt13α*	-	0.7*	38.9	Ning et al., 2008
Recombinant peptides								
7 Crp (7x)	11	E	PB-I*	*GluB-1*	KDEL ER, no KDEL	3*, 0.75*	272.7, 68.2	Takagi et al., 2005a Takaiwa et al., 2009
3Crp-7Crp	16.9	E	PB-I	*GluB-1*	KDEL ER	0.9*	53.3	Takagi et al., 2010
3 Crp, GLUA-3Crp, CBT-3Crp	5.9, 34, 17	E	n.d, PB-II, n.s	*GluB-1*	KDEL ER, GLUA, CBT	0, 1.75*,0.02 *(F)	n.d, 51.5, 1.2	Takagi et al., 2008, 2010
Der p 1_{45-145}	12	E	PB-I	*GluB-1*	KDEL ER	4.5*	375.0	Suzuki et al., 2011
$Aβ_{1-42}$	~9	E	PB-I	*GluB-1*	KDEL ER	0.4*(TR)	44.4	Oono et al., 2010
GFP-Aβ	n.s	AL	-	n. s	GFP	0.4*(F)	-	Yoshida et al., 2011
mDer f 2	~15-20	E	Derf 2 body[a]	*GluB-1*	KDEL ER	0.75-1.5*	50.0-100.0	Yang et al., 2012
10x novokinin, 18x novokinin	11, 20	n.det, N	-	*GluB-4, GluB-1*	KDEL ER	0, 4.25*	n.d, 212.5	Wakasa et al., 2011a
A1aB1b29Lac (29x Lac)	~50-60	E	ER[a]	*ESP*	part of A1aB1b*	2 (TR)	33.3-40.0	Cabanos et al., 2013
A1aB1bCrp1 and 2	56, 35 and 21[b]	E	PB II	*GluB-1*	part of A1aB1b*	0.35* (F)	6.3, 10.0, 167	Takagi et al., 2005b

Table 1.6 (continued)

Product	Mw (kDa)	Tissue	Localization	Promoter	Strategy	Yield (µg/mg)	pmols/mg seed	References
mGLP-1, GFP-mGLP-1	3.8, 31.2	E	-	*GluB-1*	6x His+ KDEL RE, GFP +KDEL RE	n.d, n.s	n.d, -	Yasuda et al., 2005
IIAEK (6x)	~36	E	PB II	*ESP*	GLU	0.0016 (TR)	44.4	Wakasa et al., 2011b
$XCII_{250-270}$ (4x)	65, 32/33[b]	E	PBs	*GluA*	GLUA	0.00098	0.02, 0.03,0.03	Hashizume et al., 2008

E, endosperm; L, leaf; C, callus; AL, aleuron layer; N, nucleolus; glyc, glycosylated form; r.c, reducing conditions; u.r, unreducing conditions; a.b, additional bands; n.d, not detected; n.det, not determined; n.s, not specified. (*), indicate that maximum yield was reported as µg/grain. To compare yields between products and strategies, grain unit was conversed to 20 mg of seed.

When yield is not calculated for individual peptide, F, fusion protein; TR, tandem repeats, *ESP* endosperm specific promoters. (a) ER-derived body, different from native PBI and PBII. (b) Molecular masses correspond to protein precurssor, acidic and basic subunit, respectively.

ScFvT84.88, Single-chain Fv antibody; HSA, Human serum albumin; IL-10, Interleukin-10; TPC7, Tree pollen chimera 7 derived from hypoallergenic birch pollen allergen Bet v 1; CTB, Cholera toxin B subunit; hLF, human lactoferrin; Sasa IFN, Atlantic salmon interferon; rhGM-CSF, Human granulocyte macrophage-colony stimulating factor; 7 Crp, composed of human T cell epitopes of Cry j 1 and Cry j II; GLUA, Glutelin A; Crp 3, composed of three of T cell epitopes derived from the Japanese cedar pollen allergens Cry j I and Cry j II; Der p 1, epitope of mite allergen ; $A\beta_{1-42}$, 2 x β-amyloid peptide; GFP , Green fluorescence protein; GFP-Aβ, GFP-Amyloid β-peptide; mDer f 2, epitope of mite allergen; A1aB1b29Lac, lactostatin derived from β-lactoglobulin as a fusion protein of with the soybean seed storage glycinin (A1aB1b); A1aB1bCrp1 and 2, Mouse dominant T cell epitope peptides of Cry j I and Cry j II allergens of Japanese cedar pollen as a fusion protein of with the soybean seed storage glycinin; mGLP-1, codon-optimized Glucagon-like peptide 1; IIAEK, derived from bovine milk β-lactoglobulin; $4XCII_{250-270}$, Type II-collagen.

1.7 Context of the present study

Antimicrobial peptides have emerged as a promising alternative or complement to conventional pesticides to control bacterial and fungal diseases (Montesinos, E., 2007; Montesinos, E. et al., 2012). In this context, the research groups of Plant Pathology (CIDSAV) and LIPPSO of the University of Girona have been actively working together in obtaining synthetic AMPs with low cytotoxicity and resistance to proteases displaying high antibacterial activity against economically important plant pathogens, such as the plant pathogenic bacteria *Erwinia amylovora*, *Pseudomonas syringae* and *Xanthomonas vesicatoria*, and the fungal pathogens *Fusarium oxysporum* or *Penicillium expansum* (Projects AGL2006-13564-C02-00, AGL2001-2354, AGL2001-2349-C3-01, AGL2003-03354, BQU2003-07420-C05, AGL2004-07799-C03-01, CTQ2006-04410/BQU, PPT-060000-2008-2, 2005SGR00835, 2005SGR00275, AGL2009-13255-C02-01). Due to the high cost associated with the chemical synthesis of large peptides, the heterologous production in living systems appears to be a suitable tool for the obtention of functional AMPs. From results obtained in our groups, the most promising AMPs were derivatives of the CECMEL11 library composed of linear undecapeptides. Thus, collaborations were established with two research groups with experience in developing genetically modified plants at different levels: the GMO plants group of the Food Technology (FT) area of the University of Girona, and the group of Plant Responses to Biotic and Environmental Stress (Department of Molecular Genetics, Centre for Research in Agricultural Genomics (CRAG)).

Thus, CIDSAV-LIPPSO-FT and CRAG have collaborated in an European Trilateral Plant KBBE project SEPSAPE (EUI2008-03572, Safe and Efficient Plant Systems for Antimicrobial PEptide production). This PhD thesis has been developed as part of the SEPSAPE project.

The basic steps initially planned, and followed in the present work, for the development of a rice seed-based platform for the heterologous production of synthetic AMPs are presented in **Fig. 1.4**. Synthetic AMPs are designed using natural compounds as models by means of function-structure approaches, and lead peptides are identified from an initial peptide library obtained by combinatorial chemistry methods. Then, based on lead peptides structure and sequence composition, a peptide library adapted to plant expression is designed, synthesized and screened for antimicrobial activity against target plant pathogens. After selection of suitable rice seed-specific promoters, codon-optimized *AMP* genes are synthesized. Later, gene cassettes consisting of a specific endosperm promoter (the signal peptides coding region of the corresponding gene was also included), the coding region of the synthetic *AMP* gene of interest, and the terminator DNA sequence, are introduced into a suitable plant expression vector. Next, the vector for expression of the *AMP* of interest is used for rice transformation (*A. tumefaciens*-mediated transformation). After plant regeneration, the phenotypic and molecular characterization of rice plants is performed. Subsequently, independent rice lines are selected and used as the parental lines to obtain T2 homozygous progeny plants. Seed amplification of selected transgenic lines is performed for preparative production of rice seeds (starting material for AMP purification). Finally, the rice seed material is processed, and the recombinant AMP isolated and further characterized.

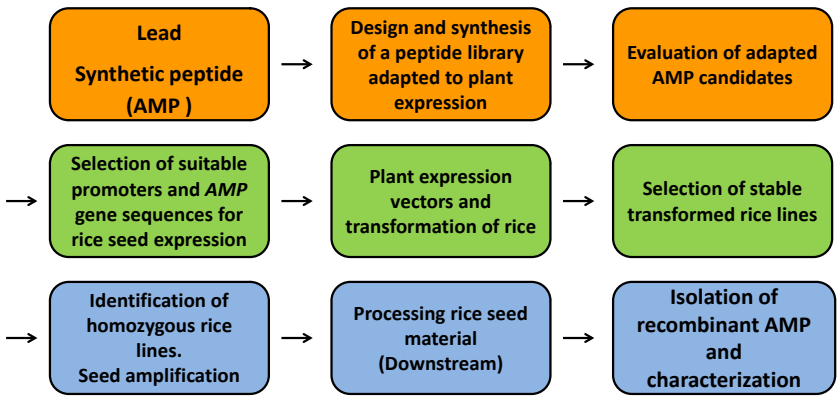


Fig. 1.4 Flow diagram used in the present work for the development of stable transgenic rice plants that produce synthetic antimicrobial peptides.

1.8 Objectives

The general objective of the present PhD thesis was to develop a sustainable platform for the production of synthetic antimicrobial peptides using rice seeds as biofactory. The specific objectives were:

1. To develop BP134 derivatives optimized for plant expression

2. To clone and build-up genetic constructs for seed-specific expression and accumulation of AMPs in the endosperm (protein body strategy) or in the embryo (oil body strategy) of rice seed.

3. To obtain stable transformed rice lines producing AMPs of interest in either the seed endosperm or the embryo tissues.

4. To purify the *in planta*-produced recombinant AMPs, to assess the yield of elite lines, and to confirm the identity and the biological activity of the transgene products.

2. MATERIALS

2.1 Microorganisms

Microorganism	Strain	Use	Origin
Fungi	*Fusarium verticillioides* strains 149, 999, 11556	Antimicrobial activity assessment	IAS-UCO, Instituto de Agricultura Sostenible-CSIC, Cordoba, Spain
	Fusarium 11427	Antimicrobial activity assessment	IAS-UCO, Instituto de Agricultura Sostenible-CSIC, Cordoba, Spain
	Fusarium semitectum 11432	Antimicrobial activity assessment	IAS-UCO, Instituto de Agricultura Sostenible-CSIC, Cordoba, Spain
	Fusarium equiseti 11439	Antimicrobial activity assessment	IAS-UCO, Instituto de Agricultura Sostenible-CSIC, Cordoba, Spain
	Fusarium avenaceum 11440	Antimicrobial activity assessment	IAS-UCO, Instituto de Agricultura Sostenible-CSIC, Cordoba, Spain
	Fusarium verticillioides CRAG 1	Seed infection assays	CRAG, Barcelona, Spain
Bacteria	*Escherichia coli* DH5αF'	Clonation process	CRAG, Barcelona, Spain
	Agrobacterium tumefaciens EAH105	Stable transformation of *Oryza sativa*. Rifampicin resistance, chromosal encoded	CRAG, Barcelona, Spain
	Pseudomonas syringae pv. syringae EPS94	Antimicrobial activity assessment	Institut de Tecnologia Agroalimentària, Universitat de Girona, Spain
	Xanthomonas axonopodis pv. vesicatoria 2133-2	Antimicrobial activity assessment	Instituto Valenciano de Investigaciones Agrarias, Valencia, Spain
	Dickeya sp.strains *1552.10.1, 1428 1a*	Antimicrobial activity assessment and seed infection assay	Instituto Valenciano de Investigaciones Agrarias, Valencia, Spain
	Pectobacterium carotovorum strains 578.3, 3902.3, 558-6-b	Seed infection assays	Instituto Valenciano de Investigaciones Agrarias, Valencia, Spain
	Erwinia amylovora PMV6076	Antimicrobial activity assessment	Institut National de la Recherche Agronomique, Angers, France

2.2 Plant material

- *Oryza sativa* ssp. *japonica* cv. Senia
- *Oryza sativa* ssp. *japonica* cv. Ariete
- *Oryza sativa* ssp. *indica* cv. IR36

Rice seeds were available at CRAG, Barcelona, Spain. Rice plants were used to perform plant transformation and to isolate gene promoters.

2.3 Kits and reaction systems

Product description	Company/Dealer
Qiaprep Spin Miniprep Kit	Qiagen Iberia, S.L., Madrid, Spain
Qiaquick Gel extraction Kit	Qiagen Iberia, S.L., Madrid, Spain
Qiagen Plasmid Midi kit	Qiagen Iberia, S.L., Madrid, Spain
DNA polimerase 500 u/µl (incl. $MgCl_2$ free buffer+ 50 mM $MgCl_2$)	Life Technologies, Invitrogen, Madrid, Spain
1 kb plus DNA ladder DNA	Life Technologies, Invitrogen, Madrid, Spain
GeneRuler™Low Range DNA Ladder	Fermentas, Thermo Scientific Molecular Biology, Barcelona, Spain.
100 mM DNTP SET	Life Technologies, Invitrogen, Madrid, Spain
pGEMT-Easy vector System I	Promega Biotech Iberica, SL, Madrid, Spain
TRIzol® plus RNA purification kit	Life Technologies, Invitrogen, Madrid, Spain
2xSYBR®Green PCR Master Mix	Life Technologies, Applied Biosystems, Madrid, Spain
Turbo DNA-free™ Kit	Life Technologies, Applied Biosystems, Madrid, Spain
GeneAmp® RNA PCR	Life Technologies, Applied Biosystems, Madrid, Spain
Western Blue®Stabilized substrate for Alkaline Phosphatase	Promega Biotech Iberica, SL, Madrid, Spain
Luminata® Forte Western HRP Substrate	Millipore iberica S.A. U., Madrid, Spain
Precision Plus Protein™ Standards Dual Color	Bio-Rad Laboratories S.A., Barcelona, Spain
Precision Plus Protein™ Standards Dual Color Xtra Standards	Bio-Rad Laboratories S.A., Barcelona, Spain
Strata™ X-CW Weak Cation Mixed Mode	Phenomenex, Madrid, Spain

2.4 Buffers and solutions

Buffers and solutions were prepared in double distilled H_2O and when necessary were autoclaved for 20 min at 121 °C. Most of them were stored at 4 °C.

For extraction and gel electrophoretic separation of nucleic acids	
TE	10 mM Tris-HCl, pH 8.0 1 mM EDTA, pH 8.0
50x TAE running buffer	2 M TriZma Base 0.1 M EDTA, pH 8.0 5.71% (v/v) glacial acetic acid
6X DNA loading buffer	30% (v/v) glycerol 0.25% (w/v) bromophenol blue 0.2% (w/v) xylene cyanol FF 0.5 M EDTA, pH 8.0
DNA extraction buffer	100 mM TriZma Base, pH 8.0 1.4 M NaCl 20 mM EDTA 2% (w/v) MATAB 1% (w/v) PEG 6000 0.5% (w/v) sodium sulphite

For Glycine SDS-PAGE electrophoresis according to Laemmly	
Protein loading buffer (TM2X)	125 mM TriZma Base 20% (v/v) glycerol 4% (w/v) SDS 0.04% (w/v) bromophenol blue 10% (v/v) β-mercaptoethanol pH 6.8
10X Running buffer	1.92 M glycine 0.25 M TriZma Base 1% (w/v) SDS pH 8.3-8.7

Upper buffer	0.5 M TriZma Base, pH 6.8 0.4% (w/v) SDS pH 6.8
Lower buffer	1.5 M TriZma Base_HCl, pH 8.0 0.4% (w/v) SDS pH 8.8

For Tris-Tricine SDS-PAGE electrophoresis according to Schägger

Cathode buffer	0.1 M Trizma base 0.1 M tricine 0.1% (w/v) SDS
Anode buffer	0.2 M TriZma Base, pH 8.0
Protein loading buffer (TTx2)	100 mM TriZma Base, pH 6.8 24% (v/v) glycerol 8% (w/v) SDS 0.2 M DTT 0.04% Coomassie G-250 pH 6.8

Polyacrylamide gel staining

Coomassie R-250 staining solution	0.25% (w/v) Coomassie R-250 40% (v/v) methanol 10% (v/v) glacial acetic acid
Coomassie R-250 distaining solution	30% (v/v) methanol 10% (v/v) glacial acetic acid
Coomassie G-250 staining solution	Solution A 2% (w/v) ortophosphoric acid 10% (w/v) ammonium sulphate Solution B 5% (w/v) Coomassie G-250 Staining solution* 98% (v/v) solution A + 2% solution B *Mix 4 parts of staining solution+ 1 part methanol

For transfer and detection of proteins and peptides	
Transblot Buffer	25 mM TriZma Base 192 mM glycine 20% (v/v) Methanol 0.1% SDS pH 6.8 Check pH before adding methanol
PBS	136 mM NaCl 2.7 mM KCl 10 mM Na_2HPO_4 2 mM KH_2PO_4 pH 7.4
PBS-T	0.1% (v/v) Tween-20 in PBS buffer
Blocking buffer	5% (w/v) BSA in PBS-T buffer

For protein and peptide extraction	
Grinding buffer	0.6 M sucrose 10 mM Na_2HPO_4 pH 7.5
Flotation buffer	0.4 M sucrose 10 mM Na_2HPO_4 pH 7.5
Floation buffer	0.25 M sucrose 10 mM Na_2HPO_4 pH 7.5
Protein gel band extraction buffer	0.125 M TriZma Base 1 mM EDTA 0.1% SDS pH 8.0

Antibiotics and other solutions (stock solutions)	
Ampicillin	50 mg/mL H_2O
Carbenicillin	100 mg/mL H_2O
Chloramphenicol	34 mg/mL ethanol
IPTG	23.8 mg/mL H_2O
Kanamycin	25 mg/mL H_2O
Rifampicin	50 mg/mL in methanol
X-Gal	20 mg/mL dimethilformamide

2.5 Media

All media were prepared in double distilled H_2O autoclaved for 20 min at 121 °C. When antibiotic was added, the media was stored at 4 °C.

Growth media	
LB Agar	10 g/L bactotryptone 5 g/L yeast extract 10 g/L NaCl *Adjust to pH 7.5* 15 g/L agar for solid media
SOB	20 g/L bactotryptone 5 g/L yeast extract 10 mM NaCl 2.5 mM KCl 10 mM $MgCl_2$ and 10 mM $MgSO_4$
YEB	5 g/L beef extract 1 g/L yeast extract 5 g/L peptone 5 g/L sucrose 0.48 g/L $MgSO_4$ *Adjust to pH 7.2* 15 g/L agar for solid media
PDA	39 g/L BD Difco™ Potato Dextrose Agar

MS	4.4 g/L MS salts including MS vitamins (Duchefa)
	10 g/L sucrose
	Adjust to pH 5.8 with KOH
	2.6 g/L phytagel

For stable transformation of *Oryza sativa*

AB Medium	3 g/L K_2HPO_4
	1 g/L NaH_2PO_4
	1 g/L NH_4Cl
	0.3 g/L $MgSO_4.7H_2O$
	0.15 g/L KCL
	10 mg/L $CaCl_2$
	2.5 mg/L $FeSO_4.7H_2O$
	5 g/L glucose
	15 g/L agar (Difcobacto)
N6 Macroelements	56.6 g/L KNO_3
	9.26 g/L $(NH_4)_2SO_4$
	8 g/L KH_2PO_4
	3.3 g/L $CaCl_{2\cdot 2}H2O$
	3.7 g/L $MgSO_4.7H_2O$
B5 Microelements	300 mg/L H_3BO_3
	758 mg/L $MnSO_4.H_2O$
	200 mg/L $ZnSO_4.7H_2O$
	75 mg/L KI
	25 mg/L $Na_2MoO_4.2H_2O$
	2.5 mg/L $CuSO_4.5H_2O$
	2.5 mg/L $CoCl_2.6H_2O$
NB Fe EDTA	2.78 g/L $FeSO_4.7H_2O$
	3.72 g/L Na_2EDTA
B5 Vitamin	100 mg/L nicotinic acid
	100 mg/L pyridoxine HCl
	1 g/L thiamine HCl

R2-I Macroelements	40 g/L KNO_3
	3.3 g/L $(NH4)_2SO_4$
	3.12 g/L $NaH_2PO_4.H_2O$
	2.46 g/L $MgSO_4.7H_2O$
R2-II Macroelements	1.46 g/L $CaCl_2.2H_2O$
R2 Microelements	1.6 g/L $MnSO_4.H_2O$
	2.2 g/L $ZnSO_4.7H_2O$
	2.83 g/L H_3BO_3
	195 mg/L $CuSO_4.5H_2O$
	125 mg/mL $Na_2MoO_4.2H_2O$
R2 Fe EDTA	1.25 g/L $FeSO_4.7H_2O$
	177 mg L Na_2EDTA
R2 Vitamin	40 mg/L thiamine HCL
NB Medium	50 mL/L N6 Macroelements
	10 mL/L NB FeEDTA
	10 mL/L B5 Microelements
	10 mL/L B5 Vitamin
	100 mg/L myo-inositol
	500 mg/L proline
	500 mg/L glutamine
	300 mg/L casein hydrolyzate
	30 g/L sucrose
	2.5 mg/L 2,4-dichlorophenoxyacetic acid (2,4-D)*
	Adjust to pH 5.8 with KOH
	2.6 g/L phytagel
NBS Medium	50 mL/L N6 Macroelements
	10 mL/L NB FeEDTA
	10 mL/L B5 Microelements
	10 mL/L B5 Vitamin
	100 mg/L myo-inositol
	500 mg/L proline
	500 mg/L glutamine

	300 mg/L casein hydrolyzate
	30 g/L sucrose
	2.5 mg/L 2,4-D*
	400 mg/L cefotaxime
	100 mg/L timentin
	50 mg/L hygromycin
	Adjust to pH 6 ,0 with KOH
	7 g/L agarose type I
PR-AG Medium	50 mL/L N6 Macroelements
	10 mL/L NB FeEDTA
	10 mL/L B5 Microelements
	10 mL/L B5 Vitamin
	100 mg/L myo-inositol
	500 mg/L proline
	500 mg/L glutamine
	300 mg/L casein hydrolyzate
	30 g/L sucrose
	5 mg/L ABA*
	2 mg/L BAP*
	1 mg/L ANA*
	100 mg/L cefotaxime
	100 mg/L timentin
	50 mg/L hygromycin
	Adjust to pH 5.8 with KOH
	7 g/L agarose type I
RN Medium	50 mL/L N6 Macroelements
	10 mL/L NB FeEDTA
	10 mL/L B5 Microelements
	10 mL/L B5 Vitamin
	100 mg/L myo-inositol
	500 mg/L proline
	500 mg/L glutamine
	300 mg/L casein hydrolyzate
	30 g/L sucrose
	3 mg/L BAP*
	0.5 mg/L ANA*

	Adjust to pH 5.8 with KOH, 3.5 g/L phytagel
R2-CL Medium	100 mL/L R2-I Macroelements
	100 mL/L R2-II Macroelements
	10 mL/L R2 FeEDTA
	1 mL/L R2 Microelements
	25 mL/L R2 Vitamin
	10 g/L Glucose
	2.5 mg/L 2,4-D*
	(100 µM) 0.1 mL/L acetosyringone*
	Adjust to pH 5.,2 with KOH
R2-CS Medium	100 mL/L R2-I Macroelements
	100 mL/L R2-II Macroelements
	10 mL/L R2 FeEDTA
	1 mL/L R2 Microelements
	25 mL/L R2 Vitamin
	10 g/L glucose
	2.5 mg/L 2,4-D*
	(100 µM) 0,1 mL/L acetosyringone*
	Adjust to pH 5.2 with KOH
	7 g/L agarose type I
R2-S Medium	100 mL/L R2-I Macroelements
	100 mL/L R2-II Macroelements
	10 mL/L R2 FeEDTA
	1 mL/L R2 Microelements
	25 mL/L R2 Vitamin
	30 g/L sucrose
	2,5 mg/L 2,4-D*
	400 mg/L cefotaxime
	100 mg/L timentin
	50 mg/L hygromycine
	Adjust to pH 6.0 with KOH
	7 g/L agarose type I
P Medium	4,4 g/L MS salts including MS vitamins (Duchefa)
	50 g/L sucrose
	Adjust to pH 5.8 with KOH, and 2,6 g/L phytagel

(*) Solvents required to solubilise phytohormones and acetosyringone are detailed above

Phytohormones	**Solvent**
1-Naphthaleneacetic acid (ANA)	Ethanol
2,4-dichlorophenoxyacetic acid (2,4 D)	Ethanol
6-Benzylaminopurine (BAP)	KOH
Abscisic acid (ABA)	KOH
Other Compounds	**Solvent**
Acetosyringone	Ethanol

2.6 Oligonucleotides

Most lyophilized oligonucleotides used in this study were synthesized either by the commercial Thermo Electron Corporation or by Roche Applied Science. 100 µM stock solutions were prepared in sterile distilled water and stored a -20 °C.

Specific oligonucleotides used to amplify the rice seed promoter and its signal peptide coding region:

Promoter	**Primer code**	**Sequence (5'-3')**	**Restriction site**
GluB-1	LM 7_F	GGGGTACCTCTAGACAGATTCTTGCTACCAA	*Kpn* I
	LM 8_R	CCGGATCC GAGCTC AAGCTTCTTGGCCATA GAACCATGGCATAAT	*BamH* I-*Sac* I-*Hind* III[a]
GluB-4	LM 1_F	GGGGTACCTACAGGGTTCCTTGCGTGAAGAA	*Kpn* I
	LM 2 _R	CCGAGCTC GGATCCATGGCATAGGAGAAGAA	*Sac* I-*BamH* I
Glb-1	LM 17_F[b]	GGGGTACCTGGAGGGAGGAGAGGGGAGA	*Kpn* I
	LM 35_R[b]	CCGGATCC GGCGCCGGAGATGGCCACCAT	*BamH* I-*Nar* I
	LM 32_F[b]	GGAGGAGAGGGGAGAGATGGTGAGAGAGGAGGAAGAAGAGGA	No RS
	LM 33_R[b]	GGAGATGGCCACCATGGCCGCCATGAGCGCCGCCGCGAAGAA	No RS
Ole18	LM 43_F	GGGAATTCGATGGTCAGCCAATACATTGATCCGTT	*EcoR* I
	LM 50_R	TGCTAAGCTAGCTAGCAAGATGAATGCAACGAAGA	No RS[c]

Restriction sites used for cloning purpose are underlined. No RS, no incorporation of restriction sites to the 5' terminal end of the oligonucleotide primer sequence; F, forward sequence; R, reverse sequence. (a) *Hind* III restriction side was added to 5' terminal of the oligonucleotide sequence when *BP134 derivatives* were cloned between the signal peptide and the terminator sequence. When *Cec* A derivatives were cloned, *BamH* I restriction side was added, using the LM2_R primer (*Sac* I-*BamH* I). (b) Two consecutive rounds of PCR had been performed to amplify the *Glb-1* promoter by using two sets of specific primers (first set, LM32_F and LM33_R). (c) Indicate native restriction site belonging to native sequence (*Bsm* I).

Specific oligonucleotides used to amplify the *nopaline synthase* terminator sequence (*Nos-T*):

Terminator	Primer code	Sequence (5'-3')	Restriction site
Nos_T pGluB-1	LM9_F	GGAAGCTTCGGGGATCGTTCAAACATTTGGCAA	*Hind* III
	LM6_R	CCGAGCTCGTTTGACAGCTTATCATCGGATCTA	*Sac* I
Nos-T pGluB-4	LM5_F	GGGGATCCCGGGGATCGTTCAAACATTTGGCAA	*BamH* I
	LM6_R	CCGATCTCGTTTGACAGCTTATCATCGGATCTA	*Sac* I
Nos-T pGlb-1	LM21_F	GGGGCGCCCGGGGATCGTTCAAACATTTGGCAA	*Nar* I
	LM6_R	CCGAGCTCGTTTGACAGCTTATCATCGGATCTA	*Sac* I
*Nos-T Kpn I**	LM40*Kpn*I_R	CCGGTACCGTTTGACAGCTTATCATCGGATCTA	*Kpn* I

Restriction sites used for cloning purpose are underlined. F, forward sequence; R, reverse sequence. (*), TNos*kpn*I was used to amplify those terminators that were part of those vectors harbouring a synthetic gene designed to be retained in the lumen of endoplasmatic reticulum. ER retention signal includes a native *Sac* I restriction site in its sequence (aag gac gag ctc).

Specific oligonucleotides used to amplify the antimicrobial peptides and the oleosin_PRS_Cec A fusion protein sequences:

Peptide/Fusion protein		Primer code	Sequence (5'-3')	Restriction site
BP188/178	*pGluB-1*	LM_F	GGAAGCTTTTCAAGAAGATCCTCA	*Hind* III
		LM46_R	CCAAGCTTATTACTAGAGCTCGTCCTT	*Hind* III
	pGluB-4	LM3_F	GGGGGATCCATGGCCAAGAAGCTCTT CAAGAAGAT	*BamH* I
		LM53_R	CCGGATCCATTACTAGAGCTCGTCCTT	*BamH* I
	pGlb-1	LM19_F	GGGGGCGCCAAGAAGCTCTTCAAGA AGAT	*Nar* I
		LM52_R	CCGGCGCCATTACTAGAGCTCGTCCTT	*Nar* I
BP192	*pGluB-4*	LM3_F	GGGGGATCCATGGCCAAGAAGCTCTT CAAGAAGAT	*BamH* I
		LM28_R	CCGGATCCATTACTAGAGCTCGTCCTT GAGGTACTTGAGGATCTTCTT	*BamH* I
	pGluB-1	LM_F	GGAAGCTTTTCAAGAAGATCCTCA	*Hind* III
		LM46_R	CCAAGCTTATTACTAGAGCTCGTCCTT	*Hind* III
	pGlb-1	LM19_F	GGGGGCGCCAAGAAGCTCTTCAAG AAGAT	*Nar* I
		LM31_R	CCGGCGCCATTACTAGAGCTCGTCCTT GAGGTACTTGAGGATCTTCTT	*Nar* I
Cec A_KDEL	*pGluB-4, pGluB-1*	LM40_F	GGGGATCCATGGCCAAGTGGAAGCTT TTCAAGAAGAT	*BamH* I
		LM41KDEL_R	CCGGATCCATTATCAGAGCTCGTCCTT CTTGGCGATTTGGTTGGCTT	*BamH* I
Cec A	*pGluB-4, pGluB-1*	LM40_F	GGGGATCCATGGCCAAGTGGAAGCTT TTCAAGAAGAT	*BamH* I
		LM42_R	CCGGATCCATTATCACTTGGCGATTTG GTTGGCTT	*BamH* I
Oleosin_ PRS_Cec A*	*pOle18*	LM56_F	GGCGTTGCATTCATCTTGCTAGCTAGC TTA	*Bsm* I
		LM58_R	CCCGAGGATGTCTTGGTG	No RS
		LM59_R	CCTAGAGGTTCTCGGTGGTGGGCGA GGATGTCTTGGTG	No RS
		LM60_R	CCAAGCTTCCACTTGCTCTGGAAGTA GAGGTTCTCGGTGGTGGG	*Hind* III[a]
Ole18	*pOle18*	LM56_F	GGCGTTGCATTCATCTTGCTAGCTAGC TTA	*Bsm* I
		LM57_R	CCAAGCTTTTACGAGGATGTCTTGGT G	*Hind* III

Restriction sites used for cloning purpose are underlined. No RS, no incorporation of restriction sites to the 5' terminal end of the oligonucleotide primer sequence; F, forward sequence; R, reverse sequence.

Asterisk indicate partial sequence of cecropin A (1-12 nucleotides from Cec A). (a) Indicate native restriction site belonging to native sequence of Cec A (*Hind* III).

Specific oligonucleotides used for DNA sequencing:

Primer code	Sequence (5'-3')
Related to gene contructs	
S1_F	CGAACACACGACGGAGCAGCT
S2_F	AGGGCACGAATACTG AGACCT
S3_F	GGATGACTCATCAATTCCACTTTACGTA
S4_F	GGCTTTCGGTGTGTAAAATAACATA
S5_F	GGTTTACATATGGGAGAGGATAGTAT
S6_F	GGCTCCTCAAATTAAAGGGCCTTTT
LM56_F	GGCGTTGCATTCATCTTGCTAGCTAGCTTA
Related to plasmid vectors	
T7_F	AATACGACTCACTATAGG
SP6_R	GGTGACACTATAGAATAC
1221_R	AGCGTGGTCGCGGCCGAGGT
1201_R	TCGAGCGGCCGCCCGGGCAGGT

Specific oligonucleotides used for quantitative polymerase chain reaction:

Promoter	Peptide/Fusion protein	Primer code	Sequence (3'-5')	Primer binding site
GluB-1	BP178	QPG1_F	TAGCATGGCGAGTTCC GTTTTC	*Promoter_SP*
		QPG1BP_R	GCGAGGTACTTGAGG ATCTTCTTG	*BP134der*
	BP188	QPG1_F	TAGCATGGCGAGTTCC GTTTTC	*Promoter_SP*
		QPG1BP_R	GCGAGGTACTTGAGG ATCTTCTTG	*BP134der*
	BP192	QPG1_F	TAGCATGGCGAGTTCC GTTTTC	*Promoter_SP*
		QPG1GLB192_R	ATTACTAGAGCTCGTC CTTGAGGT	*BP134der*
	Cec A_KDEL	QPG1_F	TAGCATGGCGAGTTCC GTTTTC	*Promoter_SP*
		QPG1/OCE_R	GCGATTTGGGTGGCTT GGC	*CecAder*
	Cec A	QPG1_F	TAGCATGGCGAGTTCC GTTTTC	*Promoter_SP*
		QPG1/OCE_R	GCGATTTGGGTGGCTT GGC	*CecAder*
GluB-4	BP178, BP188, BP192, Cec A_KDEL, Cec A	QPG4_F	ACTTTTGTGTTCTTCTC CTATGCC	*SP*
		QPG4_R	TTGCCAAATGTTTGAA CGATCCC	*Nos-T*

Promoter	Peptide/Fusion protein	Primer code	Sequence (3'-5')	Primer binding site
Glb-1	BP178	QPGLB_F	CGCCATCATCTCATCATCAGTTC	*Promoter*
		QPGLB_R	GCGAGGTACTTGAGGATCTTCTTG	*BP134der*
	BP192	QPGLB_F	CGCCATCATCTCATCATCAGTTC	*Promoter*
		QPG1GLB192_R	ATTACTAGAGCTCGTCCTTGAGGT	*BP134der*
Ole18	Ole18_PRS_Cec A	QPOLE18_F	ACCACCGAGAACCTCTACTTCC	*PRS*
		QPG1/OCE_R	GCGATTTGGGTGGCTTGGC	*CecAder*

SP, signal peptide; PRS, protease recognition site; F, forward primer; R, reverse primer; *BP134der*, BP134 peptide derivatives; *CecAder*, CecA peptide derivatives.

Gene	Primer code	Sequence (5’-3’)	Description
Actin	QPACT_F	CCTCTTCCAGCCTTCCTTCATA	Endogenous gene
	QPACT_R	GCAATGCCAGGGAACATAGTG	
hpt II	QPHIG_F	CGAAATTGCCGTCAACCAAGC	Selection gene
	QPHIG_R	CTGGAGCGAGGCGATGTTC	

F, forward primer; R, reverse primer.

2.7 Enzymes

Enzymes	Company/Dealer
Restriction endonucleases	
BamH I	Roche Diagnostics, Barcelona, Spain
Bgl I	Roche Diagnostics, Barcelona, Spain
Bsm I	Roche Diagnostics, Barcelona, Spain
EcoR I	Roche Diagnostics, Barcelona, Spain
Hind III	Roche Diagnostics, Barcelona, Spain
Kpn I	Roche Diagnostics, Barcelona, Spain
Nar I	Roche Diagnostics, Barcelona, Spain
Pst I	Roche Diagnostics, Barcelona, Spain
Sac I	Roche Diagnostics, Barcelona, Spain
Other enzymes used	
Ribonuclease T1	Calbiochem, Merck Chemicals, Madrid, Spain
Alkaline Phosphatase	Roche Diagnostics, Barcelona, Spain
T4 DNA ligase	Fermentas, Thermo Scientific Molecular biology, Barcelona, Spain
ProTEV plus protease	Promega Biotech Iberica, SL, Madrid, Spain
AcTEV™Protease	Life Technologies, Invitrogen, Madrid, Spain
TEV Protease	Sigma-Aldrich Quimica S.L., Madrid, Spain
Endoglycosidase H	New England Biolabs, IZASA, Barcelona, Spain
Protein Deglycosylation Mix	New England Biolabs, IZASA, Barcelona, Spain
Amersham ECL Glycoprotein Detection Module	GE Healthcare Life Science, Barcelona, Spain

2.8 Plasmids

Plasmid and vector system	Source	Selection gene	Description
pGEM®-T Easy	Promega Biotech Iberica, SL, Madrid, Spain	Ampicillin	Plasmid used for cloning of PCR products
pCAMBIA1300	CRAG	Kanamicin Hygromycin	Plasmid used for stable transformation of *Oryza sativa*
pCUbi::Ap-Cec A::Nos-T	CRAG	Kanamicin Hygromycin	Plasmid used for amplify cecropin A and the *Nos* terminator DNA
Riken AK243015	Rice Genome Resource Center	Ampicillin	Plasmid used for amplify the 18 kDa rice oleosin cDNA

2.9 Peptides

Most of standard peptides were synthesized by LIPPSO (University of Girona) or CASLO Laboratory ApS (Lyngby, Denmark). Cec A was purchased from Innovagen AB (Lund, Sweden). 1 mM stock solutions were prepared in sterile double distilled water, sterilized through a 0.2 µm pore filter and stored a -20°C.

Code	Sequence
BP100	KKLFKKILKYL-NH_2
BP134	KKLFKKILKYL-OH
BP184	KLFKKILSKLFKILSKILKYL-OH
BP185	KWFKKILSKLFKILSKILKYL-OH
BP186	KLFKKILSKLFKILSKILKYL-KDEL-OH
BP187	KWFKKILSKLFKILSKILKYL-KDEL-OH
BP199	KKLFKKILKYL-AGPA-OH
BP214	KKLFKKILKYL-KDEL-OH
BP203	KKLFKKILKYL-KKLFKKILKYL-OH
BP202	KKLFKKILKYL-AGPA-KKLFKKILKYL-OH
BP193	G-KKLFKKILKYL-AGPA-KKLFKKILKYL-OH
BP195	S-KKLFKKILKYL-AGPA-KKLFKKILKYL-OH
BP200	KKLFKKILKYL-AGPA-KKLFKKILKYL-AGPA-OH
BP213	KKLFKKILKYL-AGPA-LYKLIKKFLKK-KDEL-OH
BP198	KKLFKKILKYLKKLFKKILKYL-KDEL-OH
BP197	KKLFKKILKYL-AGP-KKLFKKILKYL-KDEL-OH
BP192	KKLFKKILKYL-AGPA-KKLFKKILKYL-KDEL-OH
BP194	G -KKLFKKILKYL-AGPA-KKLFKKILKYL-KDEL-OH
BP196	S -KKLFKKILKYL-AGPA-KKLFKKILKYL-KDEL-OH
BP204	KKLFKKILKYL-KKLFKKILKYL-KKLFKKILKYL-OH
BP201	KKLFKKILKYL-AGPA-KKLFKKILKYL-AGPA-KKLFKKILKYL-OH
BP216	KKLFKKILKYL-AGPA-KKLFKKILKYL-AGPA-KKLFKKILKYL-KDEL-OH
BP169	KKLFKKIKKYL-TTGLPALISW-OH
BP205	G-KKLFKKIKKYL-TTGLPALISW-OH
BP206	S-KKLFKKIKKYL-TTGLPALISW-OH
BP170	KKLFKKILKYL-TTGLPALISW-OH
BP171	KKLFKKILKYL-AGPA-TTGLPALISW-OH
BP207	G-KKLFKKILKYL-AGPA-TTGLPALISW-OH
BP208	S-KKLFKKILKYL-AGPA- TTGLPALISW-OH
BP172	KKLFKKILKYL-TTGLPALISW-KDEL-OH
BP173	KKLFKKILKYL-AGPA-TTGLPALISW-KDEL-OH
BP189	KKLFKKILKYL-GIGAVLKVLTTGL-KDEL-OH
BP191	GIGAVLKVLTTGL-KKLFKKILKYL-KDEL-OH
BP217	KKLFKKILKYL-TTGLPALIS-AGPA-SILAPLGTT-LYKLIKKFLKK-KDEL

Code	Sequence
BP180	KKLFKKILKYL-KFLHSAK-OH
BP181	KKLFKKILKYL-AGPA-KFLHSAK-OH
BP211	G-KKLFKKILKYL-AGPA-KFLHSAK-OH
BP212	S-KKLFKKILKYL-AGPA-KFLHSAK-OH
BP182	KKLFKKILKYL-KFLHSAK-KDEL-OH
BP183	KKLFKKILKYL-AGPA-KFLHSAK-KDEL-OH
BP174	KKLFKKILLYL-GIGKFLHSAK-OH
BP175	KKLFKKILKYL-AGPA-GIGKFLHSAK-OH
BP209	G-KKLFKKILKYL-AGPA-GIGKFLHSAK-OH
BP210	S-KKLFKKILKYL - AGPA-GIGKFLHSAK-OH
BP176	KKLFKKILKYL-GIGKFLHSAK-OH
BP177	KKLFKKILKYA-GIGKFLHSAK-KDEL-OH
BP178	KKLFKKILKYL-AGPA-GIGKFLHSAK-KDEL-OH
BP179	KKLFKKILKYL-GIGKFLHSAK-KDEL-OH
BP188	KKLFKKILKYL-AVAVVGQATQIAK-KDEL-OH
BP215	KKLFKKILKYL-AGPA-VAVVGQATQIAK-KDEL-OH
BP190	AVAVVGQATQIAK-KKLFKKILKYL-KDEL-OH
BP251	KKLFKKILKYLPTTENLYFQSAVAVVGQATQIAK-OH
Cec A_KDEL	KWKLFKKIEKVGQNIRDGIIKAGPAVAVVGQATQIAK- KDEL-OH
Cec A	KWKLFKKIEKVGQNIRDGIIKAGPAVAVVGQATQIAK-OH

2.10 Antibodies

Antibodies	Company/Dealer
Rabbit IgG HRP linked whole AB	GE healthcare, Barcelona, Spain
Anti-Rabbit IgG (Fc), AP Conjugate	Promega Biotech Iberica, SL, Madrid, Spain
Antibody anti-oleosin	Produced in this work
Antibody anti-cecropin A	Avaiable at GRAG, Barcelona, Spain (Coca et al., 2006)
Antibody anti-BP192	GenScript Corp, Piscataway, USA
Antibody anti-BP178	GenScript Corp, Piscataway, USA

3. METHODS

3.1 *In vitro* activity assessment of synthetic peptides

Synthetic peptides (mostly synthesized by LIPPSO, Universitat de Girona, Spain) were evaluated for their antibacterial activity against pathogenic bacterial strains indicators. Peptides chosen to express in rice plant seeds were previously selected from a library of linear undecapeptides exhibiting antibacterial activity against phytopathogenic strains. The selection criterion was based on the balance between their antibacterial and hemolytic activity.

3.1.1 Bacterial pathogens

The plant pathogenic bacterial strains assayed in this work: *Erwinia amylovora* PMV6076 (Institut National de la Recherche Agronomique, Angers, France), *Pseudomonas syringae* pv. syringae EPS94 (Institut de Tecnologia Agroalimentària, Universitat de Girona, Spain), *Xanthomonas axonopodis* pv. vesicatoria 2133-2 and *Dickeya* sp. 1552.10.1 (Instituto Valenciano de Investigaciones Agrarias, Valencia, Spain). All bacteria strains were stored in Luria Bertani broth (LB) containing 20-30% glycerol at -80 °C. *E. amylovora*, *P. syringae* pv. syringae and *Dickeya sp.* 1552.10.1 were scrapped from LB agar after growing 24 h and *X. axonopodis* pv. vesicatoria after growing 48 h at 25 °C. Cell material was suspended in double distilled H_2O and adjusted to 10^8 cfu/mL at 600 nm using an OD v.s. Cell Viable Concentration Curve, and serially diluted 10-fold if necessary.

3.1.2 Antibacterial activity

Peptide antibacterial activity was tested following the protocol described by Badosa et al. (2007). Lyophilized peptides were solubilized in double distilled H_2O to a final concentration of 1 mM and filter sterilized through a 0.2 µm pore filter (Whatman®). For minimum inhibitory concentration (MIC) assessment, dilutions of the synthetic reference peptides were made to

obtain a final concentration of 0.6, 1.25, 2.5, 5.0, 7.5, 10, 12.5 and 20 µM. Twenty µL of each dilution were mixed in a microtiter plate with 20 µL of the corresponding suspension of bacterial indicator (*E. amylovora*, *P. syringae* pv. syringae and *X. axonopodis* pv. vesicatoria) and added to 160 µL of Trypcase soya Broth (TBS) (Biòmereux, France) to obtain a total volume of 200 µL. Three replicates for each strain, peptide and concentration were used. Also, positive controls containing water instead of peptide and negative controls containing peptide without bacterial suspension were included. Bacterial cell concentration in suspension was automatically determined by optical density measurement at 600 nm (Bioscreen C, Labsystem, Helsinki, Finland). Microplates were incubated at 25 °C with 20 s shaking before hourly absorbance measurement for 48 h. Each experiment was repeated twice. The MIC value was taken as the lowest peptide concentration with no growth at the end of the experiment.

3.1.3 Bactericidal activity

Bactericidal activity of chemically synthesized peptides was assessed by a contact time exposure test of the pathogenic bacteria to a peptide solution in water. This method is in fact a killing assay and differs from continuous exposure (bacteriostatic assay) used in screening of activity **(growth inhibition, Section 3.1.2, Methods)**. Lyophilized peptides were solubilized in double distilled H_2O to a final concentration of 1 mM, filter sterilized through a 0.2 µm pore filter (Whatman®) and subsequently diluted to obtain a final concentration of 100, 50, 25, 12.5, 6.25, 3.12 and 1.56 µM. One hundred µL of each peptide dilution were mixed in a microtiter plate with 100 µL of the corresponding suspension of bacterial indicator at 10^3 and 10^4cfu/mL (*E. amylovora* and *Dickeya* sp. 1552.10.1). Microplates were incubated at 25 °C with shaking. The minimal killing concentration (MKC) was evaluated by the viable plate count method (two times: 30 min and 2 h after contact exposure). Three replicates of each strain, peptide and concentration were used as well as positive controls containing water instead of peptide and negative controls containing peptide without bacterial suspension.

3.1.4 Hemolytic activity

The hemolytic activity of synthetic peptides was evaluated by determining hemoglobin release from erythrocyte suspensions of fresh human blood (5% v/v). Blood was aseptically collected using BD vacutainer K2E System with EDTA (Belliver Industrial State, Plymouth, UK) and stored for less than 2 h at 4 °C. Blood was centrifuged at 6000 x g for 5 min, washed three times with Tris buffer and diluted in it. Three replicates of each peptide and concentration were used. Sixty-five µL of human red blood cells were mixed with 65 µL of the peptide solution (150 µM final concentration) in a MicroAmp®96-well plates (Applied Biosystems, USA) and incubated at 37 °C for 1 h with shaking. After the plates were centrifuged at 3500 x g for 10 min, 80 µL aliquots of the supernatant were transferred to 100-well microplates (Bioscreen) and diluted with 80 µL of double distilled H_2O. Hemolysis was measured as the absorbance at 540 nm with Bioscreen plate reader (Bioscreen C, Labsystem, Helsinki, Finland). Complete hemolysis was determined in Tris buffer containing melittin (200 µM) (Sigma-Aldrich Corporation, Madrid, Spain) as a positive control. The percentage of hemolysis (H) was calculated using equation $H=100x\ [(O_p- O_b)/ (O_m- O_b)]$, where O_p was the density for a given peptide concentration, O_b for the buffer, and O_m for the melittin positive control.

3.1.5 Phytotoxic activity

Leaves from *Nicotiana tabacum* plants (20-30 days old plants) were used to test the phytotoxicity of the chemically synthesized peptides. One hundred µL of synthetic peptides at 50, 100 and 150 µM were infiltrated into the mesophylls of fully expanded leaves with a syringe without a needle (the leaves were previously wounded with a needle) as described by Nadal et al. (2012). Up to three independent inoculations were carried out per peptide and concentration, and peptides are randomly distributed in different leaves and plants. The melittin peptide (at 50, 100 and 150 µM) and double distilled water are used as a positive and negative control, respectively. The appearance of symptoms on the leaves was followed for 48-72 h after infiltration and the phytotoxicity was measured as the lesion diameter (cm).

3.2 Preparation and transformation of competent *E. coli* cells

A single colony of *E. coli* (DH5αF'strain) was picked up from a fresh LB agar plate was inoculated into 5 mL of LB broth and incubated overnight at 37 °C with a constant shaking of 250 x rpm. Subsequently, 200 mL of LB broth were inoculated with 1/100 of the overnight culture and incubated with shaking until the culture reached an OD_{600} of 0.5. Then the culture was transferred into a prechilled centrifuge flask and centrifuged at 4000 x rpm at 4 °C during 10 min (Centrifuge 5810R, Eppendorf). The pellet was resuspended in 50 mM sterile prechilled $CaCl_2$ (1/2 of initial volume) and chilled (let stand) for 30-45 min in ice. Afterwards, the bacterial culture was centrifuged at 4.000 x rpm at 4 °C for 10 min and the pellet was resuspended at 1/20 of initial volume with prechilled 50 mM $CaCl_2$ containing 15% (w/v) glycerol, and let stand 30-45 min on ice. Aliquots of 100 µL of cell suspension were frozen in liquid nitrogen and stored at -80 °C until required.

Transformation of *Escherichia coli DH5αF'*competent cells was carried out by heat shock, following the protocol described by Hanahan (1983) with some modifications. Plasmidic DNA (1 µg or the product from a DNA ligation reaction mixture) was mixed with 100 µL of competent cells. The mixture was immediately frozen in liquid nitrogen and incubated at 37 °C for 5 min. After adding 900 µL of SOB medium, the tube was incubated at 37 °C for 1 h in an orbital shaker (250 x rpm). Following centrifugation at 13000 x rpm for 1 min at room temperature, the pellet was suspended in 100 µL of SOB broth and plated on LB agar plates containing the appropriates antibiotics and incubated overnight at 37 °C. Transformed colonies appeared in 16 h.

3.2.1 Screening bacterial colonies using x-gal and IPTG selection

Bacterial cells harbouring a recombinant plasmid were identified by using the blue/white screening. For this, the bacterial cells were plated onto x-gal (40 µg/mL in dimethylformamide), IPTG (23.8 µg/mL) and ampicilin agar plates. The identification of colonies carrying recombinant plasmids is based on β-galactosidase activity. Colonies that carry wild type plasmids present active β-galactosidase; therefore, the colonies are blue. On the contrary, colonies that contain recombinant plasmids do not present β-galactosidase

activity and the colonies shown a cream-white colour (sometimes, a faint blue spot in the centre).

3.3 Small-scale plasmid DNA purification from *E. coli*

A selected single colony carrying a recombinant plasmid was transferred to a 15 mL assay tube containing 5 mL of LB broth including the appropriate antibiotic. The culture was incubated at 37 °C overnight in an orbital shaker (250 x rpm). After 16 h, the culture was ready to proceed to the plasmid DNA extraction. Plasmid DNA was isolated from an overnight bacterial culture using the QIAgen Spin Miniprep kit (QIAGEN®) according to the manufacturer's manual. This protocol allows the purification of up to 20 µg high-copy plasmid DNA. Next, 2-4 mL of bacterial culture were centrifuged at room temperature for 1 min at 13000 x rpm, and the bacterial cells were harvested. Plasmid DNA was eluted either in double distilled H_2O or EB buffer and quantified using Nanodrop N-2000 UV-Vis spectrophotometer (Thermo Scientific). The resulting DNA is suitable for analysis by electrophoresis, sequencing, transformation or DNA modification reactions. The plasmid DNA is stored at -20 °C until use.

The manipulation and analysis of nucleic acids hereafter detailed, were performed following the manual Molecular Cloning: A Laboratory Manual (Sambrook and Russell 2001), and Current Protocols in Molecular Biology (Ausubel et al., 1998) with some modifications.

3.4 Large-scale plasmid DNA purification from *E.coli*

Plasmid DNA was isolated from an overnight bacterial culture **(Section 3.3, Methods)** using the QIAgen Plasmid Midi kit (QIAGEN®) according to the manufacturer's manual. This protocol allows the purification of up to 100-200 µg high-copy plasmids DNA. For this, 50-200 mL of bacterial culture is centrifuged at room temperature for 10 min at 4000 x rpm, and the bacterial cells harvested. Plasmid DNA was eluted from the spin column either with double distilled H_2O or EB buffer and quantified using Nanodrop N-2000 UV-Vis spectrophotometer (Thermo Scientific). The resulting DNA is suitable for

analysis by electrophoresis, sequencing, transformation or DNA modification reactions. The plasmid DNA was stored at -20 °C until use.

3.5 Preparation and transformation of competent *A. tumefaciens* cells

For the preparation of competent cells, a single colony of *Agrobacterium tumefaciens* EHA105 (Hood et al., 1993) taken from a fresh YEB agar plate containing rifampicin (100 µg/mL) served for inoculation of 5 mL of YEB broth (supplemented with antibiotic) and incubated overnight at 28 °C with a constant shaking of 250 x rpm. Subsequently, 5 mL of YEB broth with antibiotic was inoculated with 2 mL of the overnight culture and incubated with shaking until the culture reached the OD_{600} of 0.5-1. Then the culture was chilled during 5 min, transferred into prechilled centrifuge flask, and centrifuged at 4.500 x rpm at 4 °C for 5 min (Centrifuge 5810R, Eppendorf). The pellet was resuspended in 1 mL of 20 mM sterile prechilled $CaCl_2$ and aliquots of 100 µL of cell suspension were frozen in liquid nitrogen and stored at -80 °C until required.

The transformation of *A. tumefaciens* EHA105 competent cells was carried out by heat shock following the protocol described by Hanahan (1983) with some modifications. Briefly, 1 µg of plasmidic DNA was mixed with 200 µL of competent cells. The mixture was immediately frozen in liquid nitrogen and incubated at 37 °C for 5 min. After adding 800 µL of YEB broth, the tube was incubated at 28 °C for 4 h in an orbital shaker (250 x rpm). Following centrifugation at 13000 x rpm for 1 min at room temperature, the pellet was suspended in 100 µL of YEB broth and plated on YEB agar plates containing the appropriate antibiotics and incubated 2 days at 28 °C. Transformed colonies appeared in 12-16 h.

3.6 Small-scale plasmid DNA purification from *A. tumefaciens*

Plasmid DNA was isolated from an overnight bacterial culture in the same way as Small-scale plasmid DNA purification from *Escherichia coli* **(Section 3.3, Methods)**.

3.7 Extraction of genomic DNA from rice (*Oryza sativa*)

Genomic DNA was isolated from leaves of rice plants as described by Murray and Thompson (1980) with some modifications. For this, leaves (1 g) were pulverized in liquid nitrogen, transferred to Eppendorf 2 mL tubes, mixed with 750 µL MATAB extraction buffer and incubated at 74 °C for 30-60 min in a bath, with occasionally shaking. When tubes were tempered, 900 µL of chloroform/isoamyalcohol were added. After inverting the tube several times, the phases were separated by centrifugation at 13000 x rpm for 5 min. Then, the supernatant was mixed with 2 µL of RNAse T1 (10 mg/mL, Sigma) and incubated at 37 °C for 30 min, 900 µL of chloroform/isoamyalcohol was added and the phases were separated by centrifugation at 13000 x rpm for 5 min. The aqueous phase was transferred into a new tube and the DNA was precipitated by adding 750 µL of isopropanol. The DNA was pelleted after centrifugation at 13000 x rpm for 15 min and washed twice in 70% (v/v) ethanol. After air-drying, the pellet was resuspended in TE buffer and the DNA was quantified using Nanodrop N-2000 UV-Vis spectrophotometer (Thermo Scientific). Genomic DNA was stored at 4 °C until use.

3.8 Amplification of DNA fragments by PCR (polymerase chain reaction)

In vitro amplification of DNA by the polymerase chain reaction was used to obtain DNA fragments from genomic DNA, recombinant plasmids and screening of bacterial recombinants. Also it was used in studies of gene expression by RT-PCR. The composition of the PCR reactions used in this study is listed in **Tables 3.1** and **3.2**.

Table 3.1 Reagents composition to amplify genomic or plasmidic DNA fragments by PCR

Stock reagents		**X1 (µL)**
10x PCR buffer		5.0
50 mM $MgCl_2$		1.5
2.5 mM dNTPs		4.0
10 µM Reverse primer		1.0
10 µM Forward primer		1.0
5 U/µL Taq DNA polymerase		0.2
50-100 ng DNA		1
Double distilled H_2O	adjust to	50.0

Table 3.2 Reagents composition for screening of bacterial recombinants by PCR

Stock reagents		**X1 (µL)**
10x PCR buffer		2.0
50 mM $MgCl_2$		1.5
2.5 mM dNTPs		4.0
10 µM Reverse primer		1.0
10 µM Forward primer		1.0
5 U/µL Taq DNA polymerase		0.2
50-100 ng DNA		-[a]
Double distilled H_2O	adjust to	20.0

(a) A single colony of transformed or non-transformed *E.coli* was pinch with sterile tip. The tip was soaked into PCR tubes as a DNA template.

Exceptionally, to increase the efficiency of amplification, a PCR enhancer (Invitrogen) was used according to the manufacturer's manual.

PCR was performed in a T3000 thermocycler (Biometra®) using the conditions detailed in **Table 3.3**.

Table 3.3 Termocycler conditions for DNA amplification by PCR

Step	**Temperature (°C)**	**Duration**	**PCR cycles**
Initial denaturation	94	3 min	1
Denaturation	94	30 s	
Annealing	58	30 s	29[b]
Extension	72	-[a]	
Final extension	72	7 min	1
Hold	4	∞	1

(a) Extension is carried out for 1 min for every 1000 pb of product to amplify.
(b) Including denaturation, annealing and extension.

3.8.1 Design of oligonucleotide primers for PCR

The criteria for primer design were as follows: (1) the size of primers (18-25 nucleotides long), (2) the base composition (G+C content should be between 40-60%), (3) no inverted repeat or self complementary sequences, (4) Tm values of a primer pair (should not differ by more >5-10 °C (Tm = 4 (G + C) + 2 (A + T) °C)), (5) the 3' base for each primer should be G or C. Also, when restriction sites added at 5' end of the primer, the primer should be extended 2-3 additional nucleotides from the recognition sequence of the restriction enzyme. The OligoAnalyzer 3.1 software (http://eu.idtdna.com/analyzer/applications/oligoanalyzer/) was used to exclude the possibility of choosing oligonucleotides which may form dimer primer.

3.8.2 Purification of PCR products

To purify double-strand DNA fragments from PCR reactions, the QIAquick PCR Purification kit (QIAGEN®) was used according to manufacturer's manual. The DNA was eluted either in doubled distilled H_2O or EB buffer and quantified using Nanodrop N-2000 UV-Vis spectrophotometer (Thermo Scientific). Purified DNA was stored at -20 °C until required.

3.9 DNA gel electrophoresis

Electrophoresis agarose gels were used to separate, identify and purify DNA fragments. Samples were mixed with 6x loading buffer before loading. To visualize the DNA fragments, gels were stained with ethidium bromide (1:10000 in TAE 1X), incubated at room temperature for 15-30 min and detected using ChemiDoc™XRS+System (BioRad, USA). In this work, fragments from 100 to 4000 bp in length were separated using a horizontal electrophoresis system (Mini-Sub cell GT and wide Mini-Sub cell GT, Bio-Rad, USA.

Agarose gel percentage for size separation used in this work is detailed as follows:

Table 3.4 Agarose gel percentage for double stranded DNA resolution

Agarose (% w/v)	Range size of DNA fragments (Kb)
0.8	1.4-4.0
1	0.3-1.4
2.5	0.1-0.3

Molecular weight markers were purchased from Invitrogen (1 kb plus DNA ladder DNA) and Fermentas (GeneRuler™Low Range DNA Ladder).

3.10 Extraction and purification of DNA fragments from agarose gels

The recovery of DNA from agarose gels was done using a razor blade to cut out a slice of the agarose containing the band of interest, and placing it into an Eppendorf tube. After excising the band, DNA fragments were purified with the QIAquick Gel Extraction kit according to the manufacturer's manual. DNA was eluted either in 20-30 µL double distilled H_2O or EB buffer and quantified and analysed using Nanodrop N-2000 UV-Vis spectrophotometer (Thermo Scientific). DNA was stored at -20 °C until use.

3.11 DNA modifications

Restriction enzyme cleavage, DNA dephosphorylation and ligation of DNA fragments were used to modify DNA, following the manual Molecular Cloning: A Laboratory Manual (Sambrook and Russel, 2001), and Current Protocols in Molecular Biology (Ausubel et al., 1998). Some modifications were introduced, taking into account the recommendations of the manufacturer for each of the enzymes used.

3.11.1 Restriction enzyme digestion

Most of restriction endonucleases used to digest double-stranded DNA were purchased from Roche Applied Science. Digestion reactions were performed according to particular enzyme and manufacturer's protocol, using the appropriate buffer and the appropriate incubation temperature. Incubations

were performed in a bath (Precisterm, Selecta) and bovine serum albumin was added if recommended by manufacturers.

To purify double-stranded DNA fragments from enzymatic reactions, either QIAquick PCR Purification Kit or QIAquick Gel Extraction Kit (QIAGEN®) were used according to the manufacturer's manual. Purified DNA was eluted in double distilled water, EB or directly in restriction buffer if another enzymatic digestion was required. The eluted DNA was storage at -20 °C until use.

3.11.2 DNA dephosphorylation

Terminal 5'-phosphate groups were removed to suppress self-ligation and circularization of plasmid DNA. Dephosphorylation of the digested plasmid was carried out using the enzyme alkaline phosphatase (Roche Diagnostics GmbH) according to manufacturer's manual. To determine the amount of enzyme necessary for the reaction, the number of pmol 5' terminal was calculated using BioMath Calculators, dsDNA: Micrograms to Picomoles Converter (http://www.promega.com/a/apps/biomath/index.html?calc=pmolends). A 1 pmol 5' terminal phosphorylated DNA fragment was incubated with 1-2 units of alkaline phosphatase. The reaction assay was adjusted with 1/10 volume 10x dephosphorilation buffer and incubated at 37 °C for 60 min. Dephosphorylated DNA was stored at -20 °C until use.

3.11.3 Ligation of DNA fragments

DNA ligases used in molecular cloning reactions were T4 DNA ligase included in pGEM®-T Easy vector system (Promega) for the cloning of PCR products, T4 DNA ligase for build-up the plant transformation vectors (Fermentas). In both cases the ratio vector: insert, was considered by calculating the amount of insert DNA required to give a 1:3 ratio of vector: insert for pGEM®-T Easy Vector system, or 1:3 and 1:5 vector: insert for T4 DNA ligase from Fermentas. The reactions were performed according to manufacturer's manual and incubated at room temperature for 1 h or overnight at 4 °C (pGEM®-T Easy vector system) or overnight at room temperature (T4 DNA

ligase from Fermentas) to achieve the maximum number of transformants. The enzyme was inactivated incubating at 65 °C for 10 min. The ligation product was stored at -20 °C until transformation of competent *E. coli* cells.

3.12 DNA sequencing

Sequencing was performed on an automated sequencer ABI PRISM 377 (Applied Biosystems) at the Sequencing facility from CRAG (Barcelona, Spain).

Gene sequence database used in this work were: GenBank (http://www.ncbi.nlm.nih.gov/genbank/), EMBL (http://www.embl.org/), SWISS-PROT (http://web.expasy.org/docs/swiss-prot_guideline.html) reusing the program BLAST server through National Center for Biotechnology Information web site (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The sequence alignments were performed using the Clustal W program available at http://www.ebi.ac.uk/clustalw/.

To identify *cis*-acting regulatory elements in the promoter of rice genes, the Databases of Plant Cis-acting Regulatory DNA Elements (Higo et al., 1999, PLACE www.dna.affrc.go.jp/PLACE/) and PlantCare (Lescot et al., 2002, http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) were used.

3.13 Construction of plant expression vectors

Constructs for the expression of a variety of antimicrobial peptides in rice plants were prepared. For this purpose, the entire coding region of AMP derivatives (BP134 and Cec A derivatives) was amplified by PCR and the DNA fragments were purified and cloned into the pGEM® T-Easy vector using T/A cloning technology (a 3' adenosil extension is generated by Taq DNA in the PCR product). Subsequently, each gene cassettes consisting of a seed-specific promoter, the coding region of synthetic *AMP* gene and the *Nos* terminator sequences were introduced into the multiple cloning site of the pCAMBIA1300 plant expression vector. The steps followed in this work for the preparation of various plant expression vectors are indicated below.

***18 kDa oleosin promoter::18 kDa oleosin_PRS_Cec A* construct**

To obtain the *18 kDa oleosin_PRS_Cec A* gene, three consecutive rounds of PCR reactions were made using the AK243015 clone containing the *18 kDa oleosin* cDNA as the starting material. The entire process is presented in **Fig. 3.1**. The strategy followed to clone the *oleosin*_PRS_*Cec A* fusion gene into the binary vector for expression in rice seeds is detailed in **Fig. 3.2**.

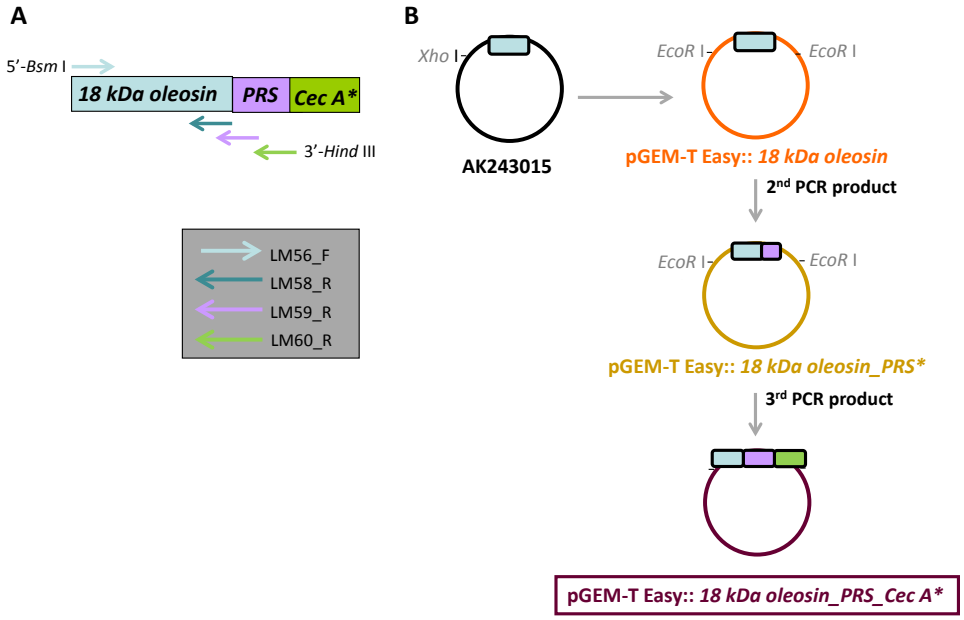
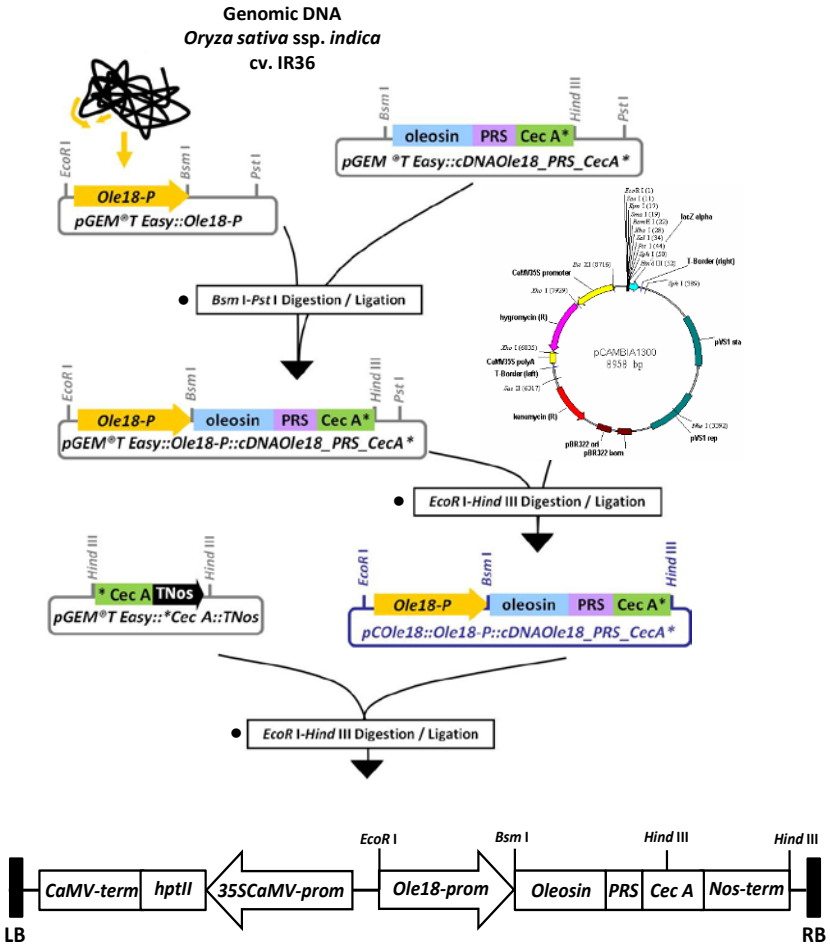


Fig. 3.1 Preparation of *18 kDa oleosin_PRS_Cec A* gene. (A) Schematic representation indicating the primers used on the successive PCR reactions to amplify the entire coding sequence of the *18 kDa oleosin_PRS* gene. LM60_R includes the first twelve base pairs of the *Cec A* coding sequence. Asterisk in *Cec A* indicates partial DNA sequence. Restriction sites relevant for cloning are indicated. (B) Construction of synthetic *18 KDa oleosin_TEV_Cec A** gene by third consecutive rounds of PCR and subsequent subcloning into pGEM®T-Easy cloning vector. *Xho* I restriction site was used to linearize the recombinant AK243015 plasmid before PCR reaction, and *EcoR* I restriction site were used to excise the target DNA to be used as a template in consecutive PCR reaction. The third-round of PCR product was subcloned into a definitive pGEM® T- Easy vector which was used for further cloning process. PRS, TEV protease recognition site.

A ***pCOle18::Ole18_PRS_Cec A***



B ***pCOle18::Ole18_PRS_Cec A***

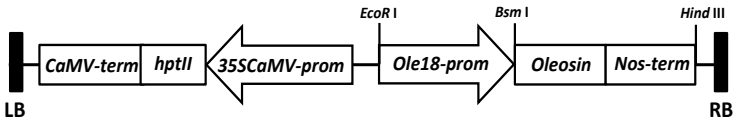


Fig. 3.2 Schematic diagrams showing the plant expression vectors prepared in this work for the expression of either the *oleosin_PRS_Cec A* fusion gene coding sequence or the *oleosin* into the rice oil bodies. (A) Cloning strategy to obtain the plant expression vector *pCOle18::Ole18_PRS_CecA::Nos-T*, used for rice transformation. (•) restriction enzyme cleavage (and when necessary, DNA dephosphorylation) followed by the ligation of the DNA fragment into the plasmid vector. Asterisk indicates partial *Cec A* gene sequence. (A) Plant expression vector

comprising the *oleosin_PRS_Cec A* fusion gene driven by the *Ole18* promoter and *Nos* terminator. (B) Plant transformation vector for the expression of the *18 kDa oleosin* gene in the rice oil bodies. *Oleosin* gene was subcloned into the *Bsm* I-*Pst* I site of *pGEM T®-Easy::Ole18-prom*. *Nos-T* sequence was fused downstream of *oleosin* gene using *Hind* III-*Pst* I restriction site. The complete cassette was cloned into *EcoR* I-*Pst* I-digested pCAMBIA1300 plant expression vector. 35SCaMV-prom, cauliflower mosaic virus 35S promoter; CaMV-term, 35S terminator from cauliflower mosaic virus. The *hptII* gene encoding resistance to hygromycin was used as a selectable marker during rice transformation. LB, left border; RB, right border.

Seed-specific promoter::AMP derivative **construct**

The entire coding region of the *BP178_KDEL*, *BP188_KDEL*, *BP192_KDEL*, *Cec A_KDEL* and *Cec A* synthetic genes was amplified using specific primers (one PCR reaction).

The various seed-specific promoters used in this work were amplified from genomic DNA obtained from rice leaves. In the case of the endosperm-specific rice promoters, in addition to the promoter region, the PCR-amplified DNA fragment also included the nucleotide sequence encoding the N-terminal signal peptide of the corresponding storage protein (glutelin B-4, glutelin B-1 or 26 kDa α-globulin). The cloning strategies of the binary vectors prepared for expression of *AMP* genes in rice seed endosperm are detailed in **Figs. 3.3** to **3.6**. The nucleotide sequence for the *nopaline synthase* terminator was also obtained by PCR using the *pCUbi::Cec A_KDEL::Nos-T* plasmid (Coca et al., 2006) as the template.

pCGluB-1::BP134der_KDEL::Nos-T

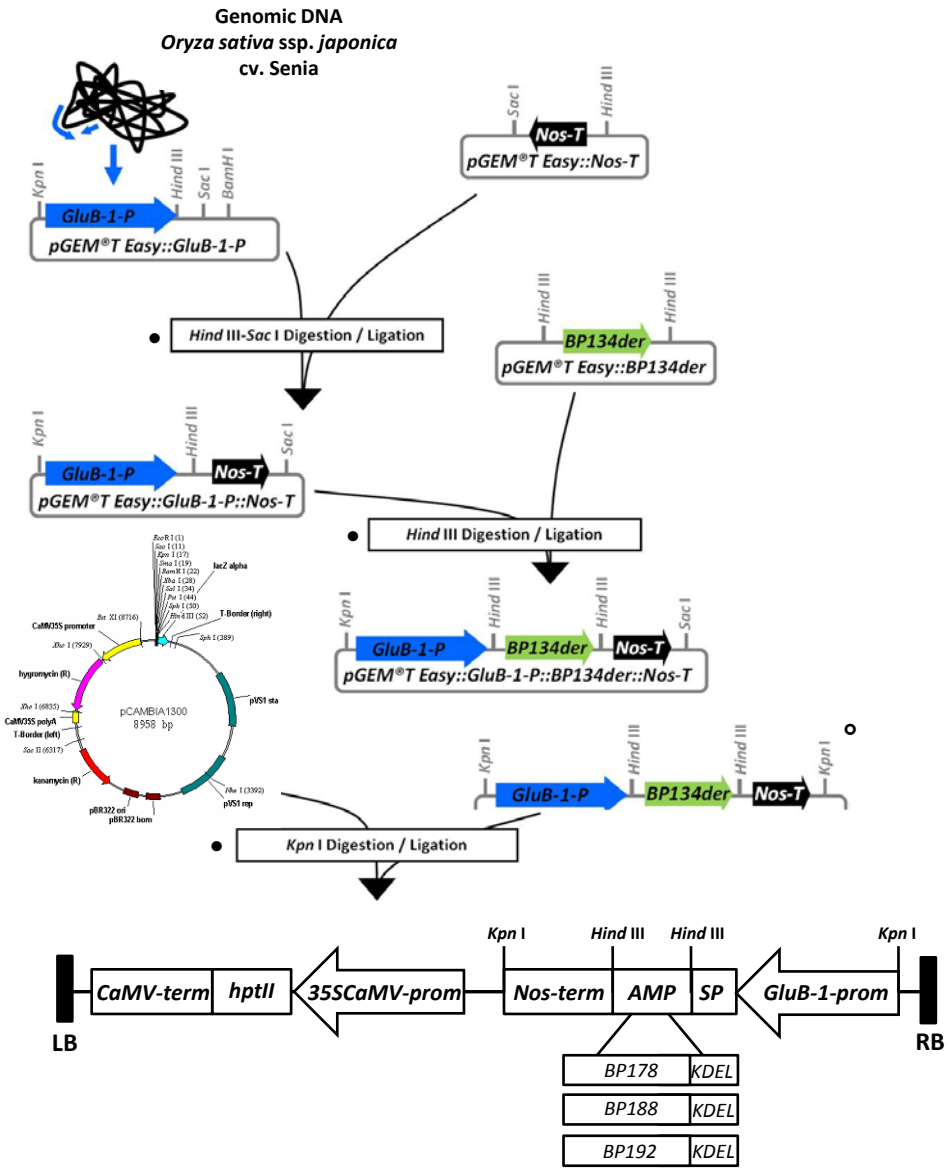


Fig. 3.3 Schematic representation of the plant expression vectors prepared in this work for the expression of the *BP134ders* in the rice endosperm. Each synthetic peptide was cloned between the *GluB-1* promoter (containing the signal peptide coding gene of the GluB-1 seed storage protein, SP) and the *nopaline synthase* terminator (*Nos-T*). Cloning strategy to obtain the plant expression vectors *pCGluB-1::BP134der::Nos-T*, used for plant transformation. The entire cassette was cloned into the *Kpn* I-digested pCAMBIA1300 vector. (•), restriction enzyme cleavage (and when necessary, DNA dephosphorylation) followed by the ligation of the DNA fragment into the plasmid vector. (°), *Kpn* I restriction site was added to 3' terminal end of the cassette in those constructs bearing KDEL sequence. (KDEL retention signal includes a native *Sac* I restriction site in

its sequence). *35SCaMV-prom*, cauliflower mosaic virus 35S promoter; *CaMV-term*, 35S terminator from cauliflower mosaic virus. The *hptII* gene encoding resistance to hygromycin was used as a selectable marker during rice transformation. LB, left border; RB, right border.

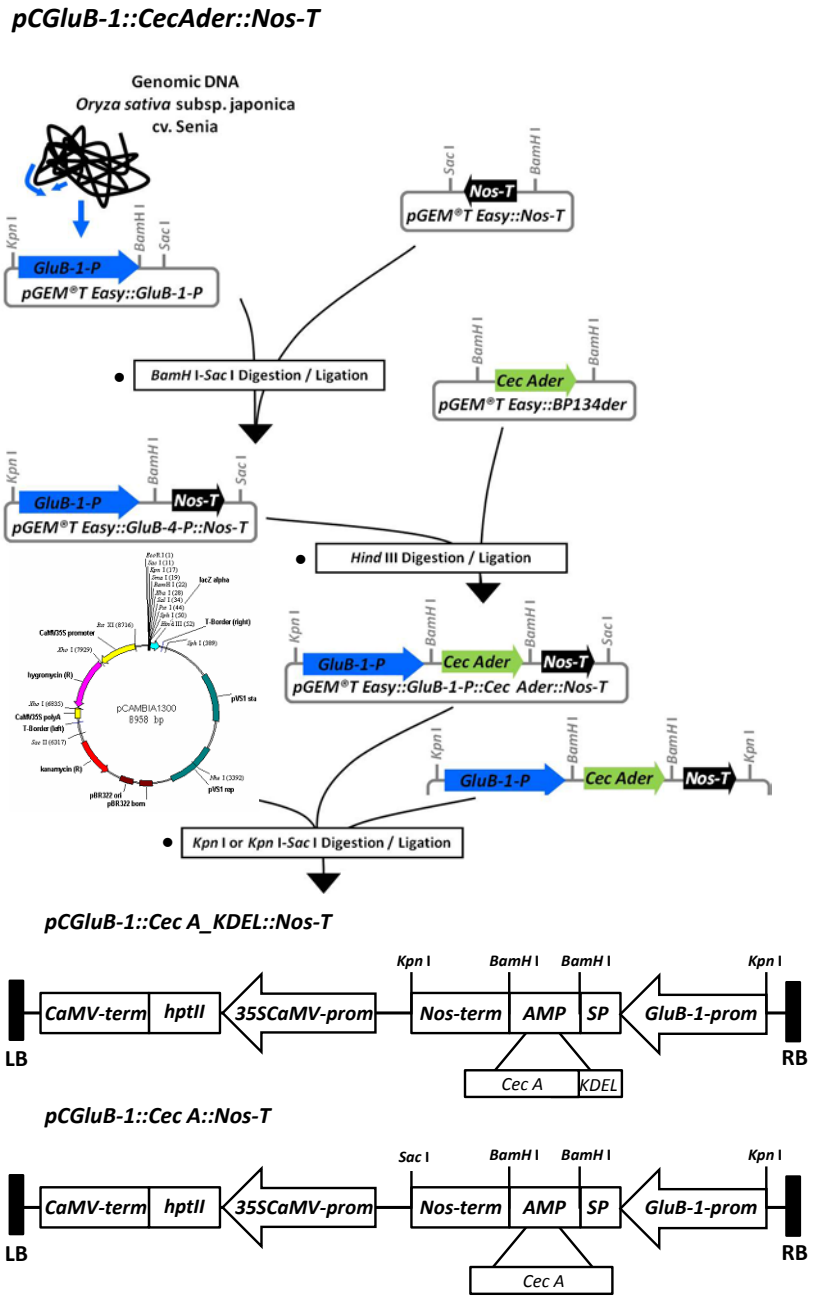


Fig. 3.4 Schematic representation of the plant expression vectors prepared in this work for the expression of the *Cec A derivatives* in the rice endosperm. Each synthetic peptide was cloned between the *glutelin B-1* promoter (containing the signal peptide coding gene of the GluB-1 seed storage protein, SP) and the *nopaline synthase* terminator (*Nos-T*). Cloning strategy to obtain the plant expression vectors *pCGluB-1::Cec Ader::Nos-T*, used for plant transformation. The complete cassette was cloned into *Kpn* I-*Sac* I or *Kpn* I-digested-pCAMBIA 1300. Those entire cassettes containing the *BP134 derivatives* and *Cec A_KDEL* genes were cloned into *Kpn* I digested

pCAMBIA1300, while the construct harbouring the *Cec A* peptide gene, was cloned into *Kpn* I-*Sac* I digested-pCAMBIA. (•), restriction enzyme cleavage (and when necessary, DNA dephosphorylation) followed by the ligation of the DNA fragment into the plasmid vector. (°), *Kpn* I restriction site was added to 3' terminal end of the cassette in those constructs bearing KDEL sequence. (KDEL retention signal includes a native *Sac* I restriction site in its sequence). *35SCaMV-prom*, cauliflower mosaic virus 35S promoter; *CaMV-term*, 35S terminator from cauliflower mosaic virus. The *hptII* gene encoding resistance to hygromycin was used as a selectable marker during rice transformation. LB, left border; RB, right border.

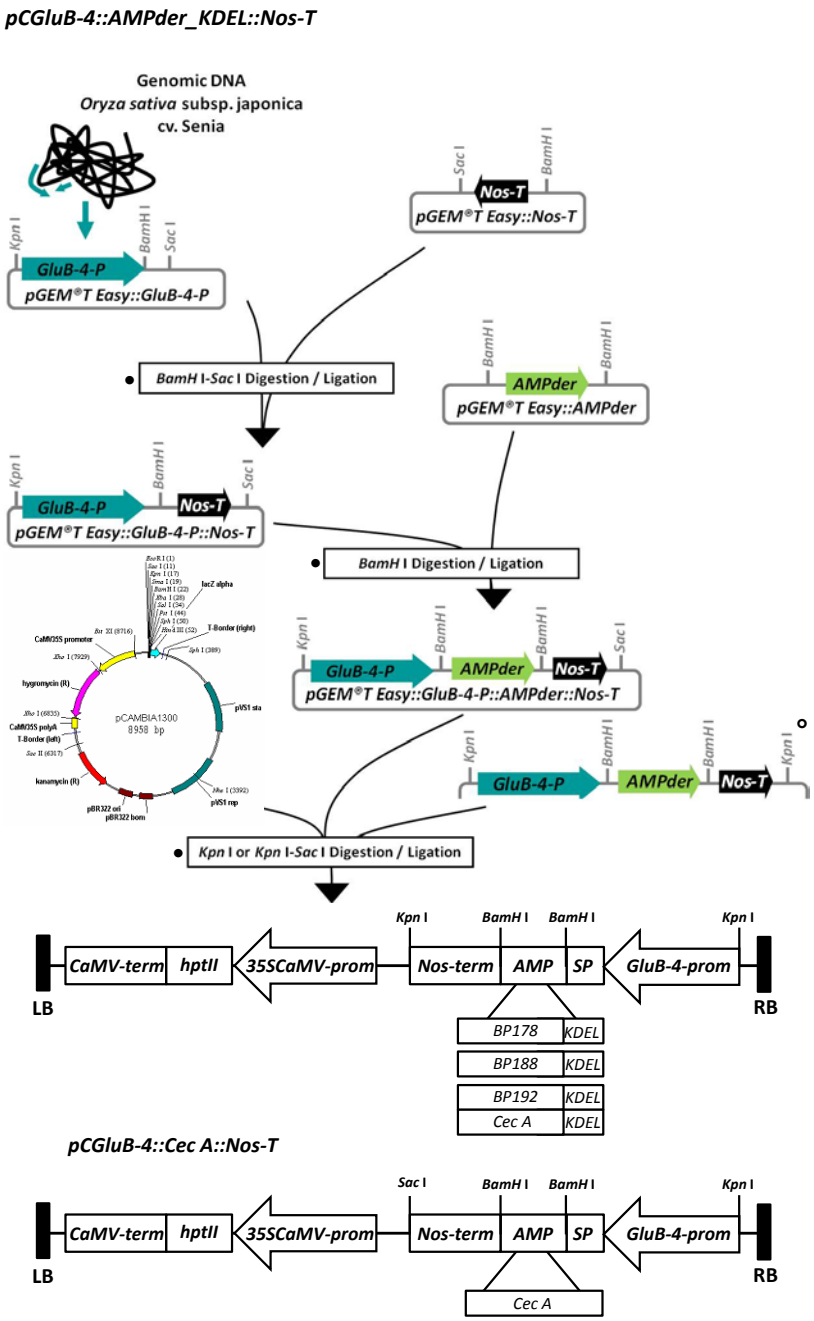


Fig 3.5 Schematic representation of the plant expression vectors prepared in this work for the expression of the *AMPs* in the rice endosperm. Each synthetic peptide was cloned between the *glutelin B-4* promoter (containing the signal peptide coding gene of the GluB-4 seed storage protein, SP) and the *nopaline synthase* terminator (*Nos-T*). Cloning strategy to obtain the plant expression vectors *pCGluB-4::AMPder::Nos-T*, used for plant transformation. The expression of *BP134 derivatives* and *Cec A_KDEL* was driven by the *GluB-4* promoter and the *nopaline synbthase* terminator. Those entire cassettes containing the *BP134 derivatives* and *Cec A_KDEL* genes were

cloned into *Kpn* I digested p-CAMBIA 1300, while the construct harbouring the *Cec A* peptide gene, was cloned into *Kpn* I-*Sac* I digested-pCAMBIA. (•), restriction enzyme cleavage (and when necessary, DNA dephosphorylation) followed by the ligation of the DNA fragment into the plasmid vector. (°), *Kpn* I restriction site was added to 3' terminal end of the cassette in those constructs bearing KDEL sequence. (KDEL retention signal includes a native *Sac* I restriction site in its sequence). *35SCaMV-prom*, cauliflower mosaic virus 35S promoter; *CaMV-term*, 35S terminator from cauliflower mosaic virus. The *hptII* gene encoding resistance to hygromycin was used as a selectable marker during rice transformation. LB, left border; RB, right border.

pCGlb-1::BP134der::Nos-T

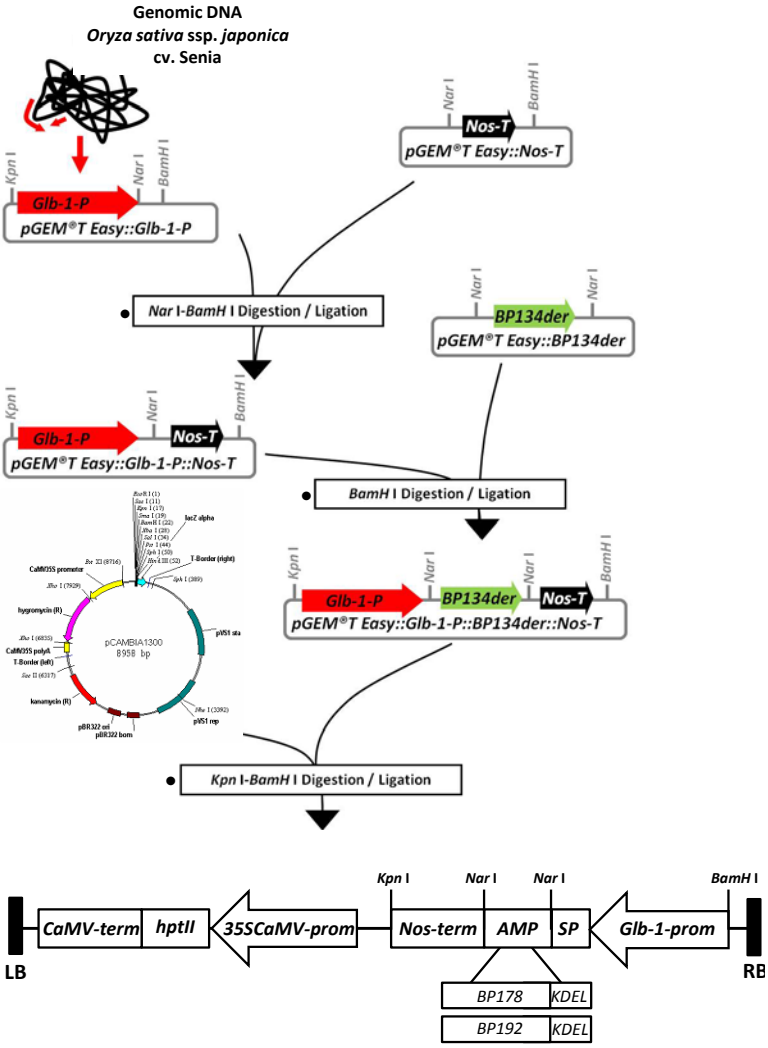


Fig. 3.6 Schematic representation of the plant expression vectors prepared in this work for the expression of the *AMPs* in the rice endosperm. Each synthetic peptide was cloned between the *26 kDa α-globulin* promoter (containing the signal peptide coding gene of the Glb-1 seed storage protein, SP) and the *nopaline synthase* terminator (*Nos-T*). Cloning strategy to obtain the plant expression vectors *pCGlb-1-P::BP134der::Nos-T,* used for rice transformation. The complete cassette was cloned into *Kpn* I-*BamH* I-digested pCAMBIA1300 plant expression vector. Those entire cassettes containing the *BP134 derivatives* and genes were cloned into *Kpn* I digested pCAMBIA 1300. (•), restriction enzyme cleavage (and when necessary, DNA dephosphorylation) followed by the ligation of the DNA fragment into the plasmid vector. *35SCaMV-prom*, cauliflower mosaic virus 35S promoter; *CaMV-term*, 35S terminator from cauliflower mosaic virus. The *hptII* gene encoding resistance to hygromycin was used as a selectable marker during rice transformation. LB, left border; RB, right border.

3.14 Stable transformation of *Oryza sativa*

Currently, particle bombardment (Christou, 1997) and *Agrobacterium*-mediated gene transfer (Hiei et al., 1994, 1997) are the most widely used methods for genetic transformation of rice plants. Advantages and limitations are reported for one or another method. Particle bombardment includes frequent integration of multiple transgene copies into a single or multiple integration sites, facilitating rearrangements of the transgenes (Shou et al., 2004). In contrast, *Agrobacterium*-mediated transformation results in lower copies of the transgene integrated in the rice genome (usually, 1-3 copies). One concern using the *Agrobacterium*-mediated transformation relies in the potential insertion of vector sequences (e.g. plasmid backbone, outside the T-DNA region) into the host genome (Ramanathan and Veluthambi, 1995). In this work, the *Agrobacterium*-mediated transformation was used to generate transgenic rice plants carrying AMP genes.

A. tumefaciens is a soil bacterium and the causal agent of crown gall (plant tumor) disease in dicotyledonous plants. *Agrobacterium*-mediated transformation had been well established for many dicotyledonous plants. This method allows the transfer of DNA with minimal rearrangements, the insertion of small numbers of copies of genes into plant chromosomes, and high quality and fertility of transgenic plants generated (Hiei et al. 1997). However, monocotyledonous plants have proven to be recalcitrant to *Agrobacterium*-mediated transformation. Hiei and co-workers (1994) were the first to achieve an efficient method to transfer DNA to rice cells, and demonstrated its stable integration in rice chromosomes, as well as, the Mendelian transmission of the DNA to the progeny. Moreover, they found that the callus cultures initiated from scutellum of mature seeds are the optimum starting material to co-cultivate with *Agrobacterium* (Hiei et al. 1994). In this work, the method used is based on the choice of starting material (callus) and a modified co-culture and selection procedure for transformation of embryogenic rice calli, following the protocol described by Sallaud et al. (2003).

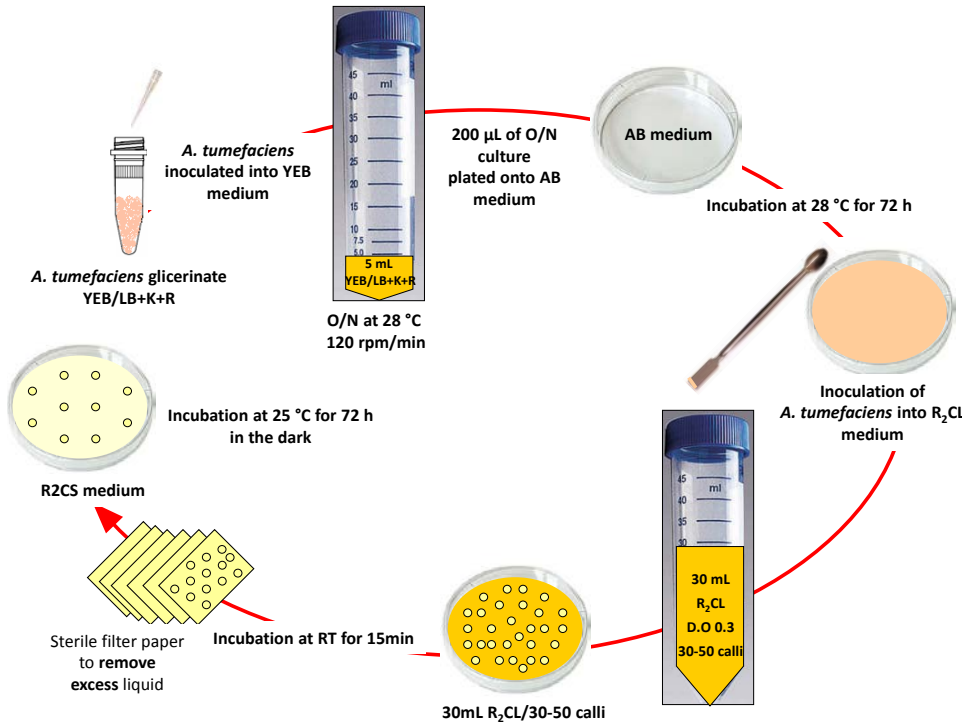
3.14.1 Disinfection of rice seeds

Mature seeds of *Oryza sativa* ssp. *japonica* cv. Senia were used for stable transformation of rice. Dehulled seeds were surface sterilized in 70% ethanol for 1 min and left in 30% of sodium hypochlorite solution containing Tween®20 as wetting agent. The seeds were placed in a rotary shaker for 40 min. After surface sterilization, seeds were rinsed with distilled water several times, including at least 1 h soaking step.

3.14.2 *Agrobacterium tumefaciens*-mediated transformation of embryogenic calli

Stable transformation of *Oryza sativa* ssp. *japonica* cv. Senia was performed as described by Sallaud et al. (2003), with some modifications to adapt the procedure to the local cultivar Senia. The plant material was grown in a Versatile Environmental Test Chamber MLR-351H (Sanyo (currently Panasonic), Japan).

Surface sterilized seeds were placed in NB agar medium and incubated at 28 °C in the dark for four weeks. 0.5-1 mm embryogenic nodular units released from primary callus derived from scutellum were transferred onto fresh NB agar and incubated for an additional 17 days. 30-50 compact embryogenic nodular units which presented 3 to 5 mm-long, rough surface, spherical shape and pale yellow colour were selected for the transformation process. Selected embryogenic nodular units were co-cultivated with the recombinants *A. tumefaciens* as follows:



Seven to ten un-contaminated calli were transferred onto R2S agar selection (including hygromycin for selection of transformed nodules and cefataxime and ticarcillin for eliminating *Agrobacterium*) and incubated at 28 °C for 14 days in the dark. During this period, the first transgenic white nodular units appeared, while the rest of the primaries calli became necrotic. Six calli were transferred and incubated on NBS agar selection at 28 °C in the dark. After seven days of incubation, the protuberances arising from primary necrotic calli were separated around the NBS plate using forceps and were incubated at 28 °C for 14 days in the dark. During these two weeks the putative transgenic calli had evolved into round shape, compact consistency and presented an opaque pale yellow colour. Calli were placed on PRAG agar regeneration (containing hygromycin, cefataxime and ticarcillin) and incubated for a further week. Seven creamy-white and dry appearance calli were transferred to RN agar regeneration, incubated at 28 °C for two days in the dark following by an incubation at 28 °C for 3-4 weeks using a 12 h/12 h (day/night) photoperiod under a photon flux of 110-150 µmol. m^{-2}. s^{-1}. Those calli which acquired green colour and were differentiated into shoots were transferred to test tubes containing P agar for 3 weeks to promote tiller and

root development before being acclimated in the greenhouse. The primary transgenic and further seed progenies were grown to maturity in the greenhouse.

3.14.3 Seed production under greenhouse conditions

The substrate used to grow rice plants consisted of 225 L of Floratorf substrate (Floragard), 125 L of num. 3 vermiculite (Europerlita Española), 250 g CO_3Ca and 375 g of Osmocote (11+11+13+2MgO, Scotts). Once the rice seed (or transgenic seedlings) were planted in substrate, the soil surface was kept moist all time, avoid to add large amounts of water to young plants (plants starting tilling). When rice plants had reached larger size, the soil was kept well watered at all times by underwatering and by adding water to the soil surface using automated dripping. Rice plants were fertilized, every two weeks, with a mixture consisting of 100 g Kristalon (Yara), 3 g Sequestrene (Syngenta) and 3 g Hortrilon (Compo) diluted 1:50 just before use. When necessary, plants were treated with pesticides to guard against insect or pathogen infestations. For fungicidal treatment Topas (Syngenta) and Draconil (Masso) was used, while Chas (Agrodan) and Confidor (Bayer) were used for insecticide treatment. All products used for the rice seed production were from Spanish manufacturers.

For greenhouse rice seed production environmental conditions are detailed as follows:

Table 3.5 Greenhouse environmental conditions used in this study

	Day	**Night**
Lighting	12 hours	0 hours
Humidity	60%	80%
Temperature	28°C	20°C

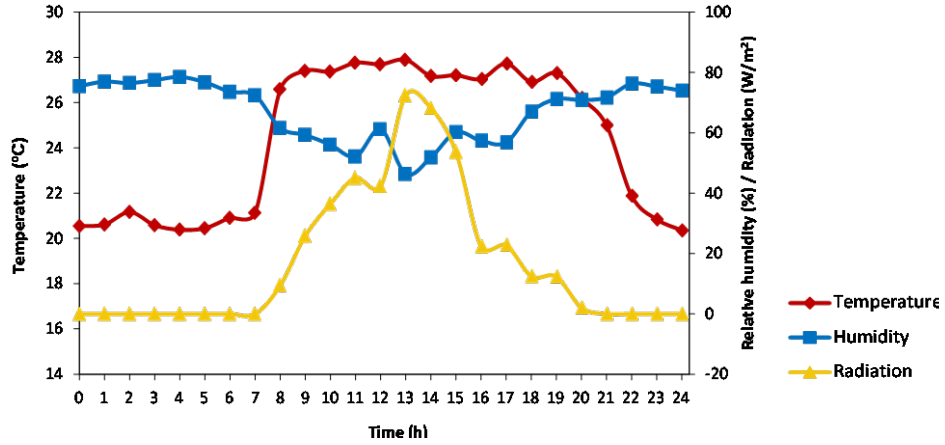


Fig.3.7 Greenhouse environmental conditions.

All mature seeds were harvested stored at 20 °C in dry and cool conditions until use.

3.14.4 Obtention of homozygous transgenic rice lines

To obtain homozygous lines expressing a foreign gene, the selection method was based on Mendelian segregation of the gene marker *hpt*II. The offspring homozygous for *hpt*II gene was identified in the T2 lines. Subsequently T3 lines were obtained from the homozygous T2 lines.

Twenty-four dehulled T1 rice seeds of transformants were surface-sterilized **(Section 3.14.1, Methods)** and placed on a Petri dish containing 50 mg/L of hygromycin (previously T1 seeds were pre-germinated in double distilled H_2O for 24 h). Seeds were incubated at 28 °C for seven to ten days using a 12 h/12 h (day/night) photoperiod under a photon flux of 110-150 µmol. m^{-2}. s^{-1} (Versatile Environmental Test Chamber MLR-351H, Sanyo (currently Panasonic), Japan). Nine hygromycin resistant seedlings were potted and grown to maturity in greenhouse. Homozygous T2 seeds were selected for *htp*II gene in the same way as T1 seeds and planted in substrate to obtain T3 homozygous lines for the gene marker *htp*II and for the transgene.

3.14.5 Screening of transgenic rice plants for transgene integration

To confirm the insertion of the peptide/fusion protein into the genome hygromycin-resistant plants, two independent PCRs were performed using two sets of oligonucleotides as primers. A set was designed to amplify the DNA fragment that expands from the promoter region to terminator region of the transgenes. The second pair of primers was designed to amplify the coding region for the peptide sequence. As neither the complete DNA fragment (Promoter to Terminator) nor the peptide coding sequence exist in wild-type transgenic plants, positive amplification should appear only in the transgenic plants.
Total genomic DNA of transgenic plants (including plants transformed with the empty vector) and non transgenic plants was isolated from rice plant leaves and analysed by PCR (**Sections 3.7 and 3.8, Methods)**. In all the cases negative controls were included: Senia or plants transformed with the empty pCAMBIA1300 vector and samples with no DNA added to the PCR reaction.

3.14.6 Determination of transgene copy number by Quantitative Real-Time PCR assay (qRT-PCR)

To estimate the transgene copy number in the transgenic rice plants, a qRT-PCR assay was performed. The *actin* gene was used as the endogenous reference gene.

Quantitative RT-PCR quantifies the amplification product by fluorescence. The chemical detection based on fluorescence can be performed by using either intercalating agents or specific probes labelled with fluorochromes. In this work, SYBR®Green I was used for chemical fluorescence detection. SYBR®Green I is the most commonly used dye and emits at 520 nm for non-specific detection. It is a double-stranded intercalating dye, which fluoresces once bounds to the DNA. The amount of dye incorporated is proportional to the amount of generated product. Fluorescence is monitored during each PCR cycle and analysis takes the Ct (threshold cycle) value, which means the cycle in which there is the first detectable significant increase in fluorescence. Every product generated has a different dissociation

temperature, so it is possible to check the number of products amplified, considering that a single pair of primers should produce a unique and well defined melt curve.

To perform the qRT-PCR analysis specific oligonucleotides were designed for the detection and quantification of the different transgenes using the Beacon Designer 7.0 software (Premier Biosoft International, Palo Alto, CA, USA). The oligonucleotides were designed to have a maximum length of 24 bases and to produce an amplicon between 100-200 pair bases.

Quantitative Real Time-PCR was carried out in a fluorometric thermal cycler (7300 Real-Time PCR System, Applied Biosystems®) by use of Mix SYBR®Green PCR Master Mix (Applied Biosystems). The total reaction volume of this PCR reaction was 20 µl and the reaction mixture is show in **Table 3.6**.

Table 3.6 Reagents composition for DNA amplification by qPCR using SYBR green

Stock reagents		**X1 (µL)**
2xSYBR®Green PCR Master Mix		10.0
2-6-18 µM Reverse primer*		1.0
2-6-18 µM Forward primer*		1.0
100 ng DNA		2
ddH_2O	adjust to	20.0

*the concentration of the primers was different for different set of primers.

Reaction conditions are based on an initial denaturation, amplification and quantification and a melting (dissociation) curve program to prevent the detection of artefacts **(Table 3.7)**.

Table 3.7 Termocycler conditions for standard qPCR using SYBR green

Step	**Temperature (°C)**	**Duration**	**PCR cycles**
Initial denaturation	95	10 min	1
Denaturation	95	15 s	50
Annealing_Extension	60	1 min	50
Dissociation curve analysis	95	15 s	1
	60	1 min	1
	95	15 s	1
	60	15 s	1

The *actin* housekeeping gene and the *hygromycin* gene were used in this work as previously described by Nadal et al., (2012). For each gene system, the concentration of the primer pair was optimized to prevent non specific reactions or artefacts that could hide the real result. For this optimization, the combinations of concentration tested were 100/100 nM, 300/300 nM and 900/900 nM using as a template 10^6 molecules DNA for reaction.

Once determined the concentration of primers to be used, the performance of the calibration curve was tested. The establishment of the standard curve is one of the key steps in determining the copy number of a selected target sequence. The calibration curve was performed with different amounts of DNA molecules of the target sequence per reaction tube. Serial dilutions of 10^6, 10^5, 10^4 and 10^3 molecules/reaction were generated for each gene. Three replicates per sample were performed and qRT-PCR conditions were the same as those used during the optimization of the qPCR primers. No template controls were also included for each assay.

After performing the qRT-PCR analysis, a standard curve corresponding to the relationship between the number of molecules and Ct was obtained. Standard curves with a correlation efficiency equal to negative one (-1) are desirable. The efficiency of the standard calibration curve is calculated with slope by means of the following formula:

$E= (10^{(-1/Slope)}-1) \times 100$ (Rasmussen, 2001), where E is the efficiency of the amplification reaction.

The slope of a standard curve provides an indication of the efficiency of the amplification and should be close to -3.22, corresponding to an efficiency of 100% and indicating that each qRT-PCR cycle doubles the amount of DNA.

The Ct values obtained for each event were used to estimate the copy number of the endogenous reference (*actin*), the selectable marker (*hygromycin*) and the *AMP* transgene, and these result were used to calculate the ratios of transgene and *hygromycin* genes (relative to the *actin* gene for all events analysed).

When working in heterozygote transgenic plants, the ratio of transgene to actin copy number should be close to 0.5 for a single copy insertion (one copy of transgene in contrast to the two copies of actin gene). If the transgenic plants were homozygous, the ratios will be close to 1 for a single copy insertion (two copies of both trangene and actin gene).

3.15 RNA extraction from rice seeds

Plant material (immature seeds, 17-21 days after flowering) was ground to a fine powder in liquid nitrogen with the Tissuelyzer II system (Qiagen). Total RNA was extracted from rice immature seeds using TriZol® Plus RNA Purification Kit (Invitrogen) according to the manufacturer's manual. 200 mg of seed powder was mixed with 1 mL of Trizol® reagent and incubated for 5 min at room temperature. After adding 200 µL of chloroform and shaking the tube vigorously by hand, the sample was incubated for 2-3 min at room temperature. Then, the tubes were centrifuged at 12000 x g for 15 min at 4 °C, and the upper phase containing the RNA was transferred to a fresh RNA free tube containing an equal volume of 70% ethanol. Following vortexing, 700 µL of sample was transferred to a spin cartridge to bind, wash and elute the RNA. The RNA was eluted in 35 µL of RNAse free water and quantified using Nanodrop N-2000 UV-Vis spectrophotometer (Thermo Scientific). Total RNA was stored al -80 °C until use.

3.15.1 DNAse treatment of RNA

Following a TriZol® extraction protocol, RNA samples were routinely subjected to DNAse treatment (Turbo DNA-free™ Kit, Applied Biosystems) to remove any contaminant DNA. 0.1 volumes of 10X Turbo DNase free buffer and 1 µL Turbo DNAse were added to 33 µL of purified RNA. After vortexing, the reaction was incubated for 30 min at 37 °C. Next, 0.1 volumes of DNAse Inactivation reagent was added and the reaction was incubated at room temperature for 2 min (mix occasionally). The supernatant (DNA-free RNA) was recovered by centrifuging at 10000 x g for 1.5 min at room temperature. The RNA was quantified using Nanodrop N-2000 spectrophotometer. Total DNA-free RNA was stored al -80 °C until use.

3.15.2 cDNA synthesis

Complementary DNA (cDNA) was generated from immature seed RNA (17-21 DAF) by using reverse transcriptase (GeneAmp® RNA PCR). The reaction mixture is detailed in **Tables 3.8** and **3.9**, respectively.

Table 3.8 RNA denaturation

Stock reagents	**X1 (µL)**
50 µM oligodT	1.5
RNA (2 µg)	x
Double distilled H_2O adjust to	**14.0**

(x) µL required to denature 2 µg of RNA

After heating to 75 °C for 5 min using T300 thermocycler (Biometra, Germany), the reaction was chilled for 5 min in ice.

Table 3.9 Reverse transcription

Stock reagents	**X1 (µL)**
Denatured RNA	**14.0**
10x PCR buffer	3.0
50 mM $MgCl_2$	1.5
10 mM dNTPs	3.0
20 U/µL MuLVRT	1.5
Double distilled H_2O adjust to	30.0

Reverse transcription was performed in a T300 thermocycler (Biometra, Germany) using the cycle conditions described in **Table 3.10**.

Table 3.10 Termocycler conditions for standard retrotranscription reactions

Step	Temperature (°C)	Duration (min)	PCR cycles
Reverse transcription	42	15	1
Denaturation	99	5	1
Cool	5	5	1
Hold	4	∞	1

The cDNA was stored at -20 °C until use.

3.16 Subcellular fractionation of rice seeds and protein extraction

In this study, transgenic plants were designed to target the expression of recombinant antimicrobial peptides or fusion proteins to subcellular compartments such as protein bodies (seed endosperm) or oil bodies (embryo). Each organelle has its characteristics density, size and shape that allow them to be isolated from the homogenate. Thus, two different centrifugation techniques were performed based on organelles characteristics.

3.16.1 Purification of protein bodies from rice seeds

Protein bodies are high density organelles that can be recovered from rice seed homogenate using differential centrifugation. A homogenate subjected to a series of increasing centrifugal forces cycles will form a series of pellet with decreasing sedimentation particles rate. Fifteen dehulled mature rice seeds from each transformant and the non-transformant control were soaked in water for 2 h and the endosperm was excised from the embryo. Endosperms were homogenized at 4 °C in 1.5 mL of grinding buffer using a mortar and pestle. After centrifugation at 200 x g for 2 min at 4 °C, the upper phase containing protein bodies was transferred to a fresh tube and centrifuged at 2000 x g for 10 min at 4 °C. The pellet was suspended in 70 µL of grinding buffer or TT1X. The enriched protein body fraction was stored at -20 °C until use.

3.16.2 Purification of oil bodies from rice seeds

Oil bodies and associated proteins (e.g. oleosins) are easily isolated from rice seed homogenate components by flotation centrifugation as they form an immiscible “fat pad” in the upper supernatant. Oil bodies were isolated from

rice seeds by sucrose gradient density as described by Chuang et al. (1996) with some modifications. Five dehulled mature rice seeds from each transformant and the non-transformant control were soaked in water for 2 h and homogenized at 4 °C in 0.9 mL of grinding buffer using a mortar and pestle. 750 µL of homogenate were transferred to the bottom of a 2 mL tube, and then 750 µL of flotation buffer was layered on top. After centrifugation at 10000 x g for 20 min at 4 °C, the oil bodies were collected at the top of the supernatant phase. Oil bodies were placed into 2 mL fresh tube and resuspended in 750 µL of grinding buffer containing 2 M NaCl. 750 µL of floating buffer was layered on top and after centrifugation at 10000 x g for 20 min at 4 °C, oil bodies were recovered again by removing infranatant phase. Oil body fraction was suspended in 100 µL of grinding buffer, TM1X or TEV protease buffer and stored at -20 °C until use.

3.16.3 Protein precipitation procedures

3.16.3.1 Protein precipitation from protein bodies

The enriched protein body fraction was pelleted after centrifugation at 2000 x g for 10 min at 4 °C. Proteins were precipitated by adding ice cold-acetone containing 15% (v/v) TCA and 20 mM DTT to the pelleted protein bodies. After incubation for 60 min at -20 °C the precipitate was recovered by centrifugation at 13000 x rpm for 23 min at 4 °C. The pellet was washed by adding ice-cold acetone and centrifuging at 13000 x rpm for 15 min at 4 °C. After air-drying, the pellet was resuspended in double distilled H_2O or TT1x, depending on their subsequent use. This protocol was followed to obtain precipitated proteins free of cell debris or starch.

3.16.3.2 Protein precipitation from oil bodies

When necessary, oil body proteins were precipitated by adding 10% of TCA. After incubation for 30 min on ice the precipitate was recovered by centrifugation at 13000 x rpm for 40 min at 4 °C. The pellet was washed by adding ice-cold acetone and centrifuging at 13000 x rpm for 5 min at 4 °C. This wash was repeated once. After air-dried, the pellet was resuspended in

double distilled H_2O, TM1x or TEV protease buffer, depending on their subsequent use.

3.16.4 Digestion of fusion proteins with tobacco Etch Virus (TEV) protease

One of the strategies followed to produce antimicrobial peptides in rice embryo was based on the production of AMPs as a fusion protein with rice oleosin (van Roojien and Moloney, 1995). A sequence encoding a TEV cleavage site was included between the oleosin and the peptide coding sequence. TEV protease recognize a linear epitope of E-X-X-Y-X-Q-(G/S) and cleaves between Q and G/C. The most frequently used cleavage sequence is ENLYFQG/S (Dougherty et al., 1988; Dougherty and Parks, 1989).

3.16.4.1 TEV activity from different commercial suppliers on a synthetic peptide as the substrate

A standard peptide, BP251, which included the TEV processing sequence, was synthesized (LIPPSO) and used to test the performance of three different commercial TEV preparations. Thus, AcTEV™Protease (Invitrogen), TEV Protease (Sigma-Aldrich) and ProTEV plus (Promega) were used according to the recommendations of the manufacturer for each of the enzymes tested. In all cases, 20 μg of fusion protein were used as a TEV substrate. Best results were obtained using the ProTEV plus.

3.16.4.2 Digestion of recombinant fusion protein with the TEV protease

Three types of oil body samples were subjected to protease digestion: (1) crude oil body fraction **(Section 3.16.1.2, Methods)**, (2) oil body protein extracts **(Section 3.16.2.2, Methods)** and (3) fusion protein purified from SDS-PAGE gels **(Section 3.17.3, Methods)**. After suspending oil body samples in buffer, the fusion proteins were routinely digested under reducing conditions using ProTEV plus protease (Promega) according to the manufacturer's manual as follows:

Table 3.11 Mixture of proteolytic digestion with TEV

Stock reagents		X1 (µL)
20x ProTEV buffer		5.0
100 mM DTT		1.0
Fusion protein		20 µg
5 U/µL ProTEV Plus		2.0
Double distilled H_2O	adjust to	100.0

To set up optimal conditions for proteolytic digestion, samples were incubated at 30 °C or 4 °C and 20 µL aliquots were taken from the digestion mixture at 0, 1, 2, 4, 6 h and overnight, and stored at -20 °C until use. If necessary, TT1X sample buffer was added and analysed by Tricine SDS-PAGE **(Section 3.17.2, Methods)**.

3.17 Denaturing protein electrophoresis

Proteins were routinely separated using SDS-polyacrylamide gel electrophoresis (SDS-PAGE). In this study, SDS-PAGE was used (1) to separate proteins from protein mixtures and subsequent detection of the protein/peptide of interest by immunoblot analysis, and (2) for preparative purposes (e.g. purification of rice oleosin and oleosin_Cec A fusion protein). For gel electrophoresis, the Mini-Protean® Tetra Cell system (Bio-Rad) or the Midi Format 1-D Electrophoresis System, Criterion™ cell (Bio-Rad) were used following the manufacture's instructions with minor modifications.

3.17.1 SDS-PAGE

SDS-PAGE was used to separate proteins with a mass >10 kDa. Stock solutions and buffers were performed following the SDS-PAGE Buffer System described by Laemmly (1970). Protein samples in solution were mixed with TM2X sample buffer to a final concentration of TM1X, while lyophilized protein extracts were resuspended directly in TM1X. After boiling at 100 °C for 5 min in a block heater (Stuart, UK), protein samples were loaded into each well of the SDS-PAGE gel (composition described in **Tables 3.12** and **3.13**, respectively). The gel was run at 60 V constant for 30 min until the samples entered the resolving gel and 100 V constant till the bromophenol blue front reached the end of the gel.

Table 3.12 Composition of the stacking gel (5.2% acrylamide)

Stock reagents	**X1**
Acrylamide/Bis-acrylamide 40% (37.5:1)	3.13 mL
Lower buffer	2.50 mL
Double distilled H_2O	4.38 mL
Ammonium persulfate 10% (w/v)	40 µL
TEMED	5 µL

Table 3.13 Composition of the resolving gel (12.5% acrylamide)

Stock reagents	**X1**
Acrylamide/Bis-acrylamide 40% (37.5:1)	0.50 mL
Upper buffer	1.00 mL
Double distilled H_2O	2.25 mL
Ammonium persulfate 10% (w/v)	40 µL
TEMED	6 µL

Alternatively, when the Criterion™ cell midi electrophoresis was performed, the Criterion™ TGX™ Precast Gels were used (Bio-Rad).

3.17.2 Tris-Tricine SDS-PAGE

Tris-Tricine SDS-PAGE was used for the optimal separation of peptides and small proteins in the size range of 1 to 10 kDa. A higher resolution and separation is achieved by slowing the migration rate of peptide-SDS complex due to the strongly differing pK values of glycine and tricine groups. Stock solutions and buffers were prepared following the Tricine-SDS-PAGE protocol described by Schägger (2006). Protein samples were prepared as indicated above **(Section 2.16.1, Methods)** using TT2x instead of TM2x protein loading buffer. The gel was run at 80 V constant for 60 min until the samples entered the resolving gel and 100 V constant till the Coomassie G-250 front reached the bottom of the gel. The composition of Tricine SDS-PAGE gel is detailed in **Tables 3.14** and **3.15**.

Table 3.14 Composition of the Tris-Tricine SDS-PAGE stacking gel(9.7% acrylamide)

Stock reagents	**X1**
Double distilled H_2O	1.0 mL
Tris-HCl 3M pH 8.5	0.5 mL
Acrylamide/Bis-acrylamide 40% (37.5:1)	0.5 mL
SDS 20%	8 µL
TEMED	5 µL
APS 15% (w/v)	6 µL

Table 3.15 Composition of the Tris-Tricine SDS-PAGE resolving gel (18% acrylamide)

Stock reagents	**X1**
Double distilled H_2O	0.45 mL
Glycerol	1.20 mL
Tris-HCl 3M pH 8.5	3.35 mL
Acrylamide/Bis-acrylamide 40% (37.5:1)	4.85 mL
SDS 20%	50 µL
TEMED	5 µL
APS 15% (w/v)	50 µL

Alternatively, when the Criterion™ cell midi electrophoresis was performed, the Criterion Tris-Tricine Precast Gels were used (Bio-Rad).

3.18 Coomassie blue staining

After electrophoretic separation, the proteins were visualized by Coomassie blue staining. Depending on the subsequent use, Coomassie G-250 or Coomassie R-250 were used as described bellow.

3.18.1 Colloidal Coomassie staining

Colloidal Blue G-250 staining was used to visualize those bands which were later excised from de polyacrylamide gel. The advantage of Colloidal blue stain lies in the reduction of free dye in solution requiring only water to distain. Moreover, allows detection of 8-50 ng of protein (three times more sensitive than Coomassie R-250) and is compatible with mass spectrometry reagents. The staining and distaining protocols were performed according to Neuhoff et al. (1988) with some modifications. After overnight incubation in the staining solution with gentle shaking, the gel was washed in double distilled H_2O. The gel was stored in double distilled H_2O until use.

3.18.2 Coomassie R-250 staining

Coomassie R-250 staining was routinely used to visualize proteins in gels. Gels were stained for at least 30 min in staining R-250 solution followed by distaining with distaining R-250 solution until the background became clear. The gels were stored in double distilled H_2O.

3.19 Preparative SDS-PAGE

Proteins from polyacrylamide gel were purified according to the protocol described by Chuang et al. (1996). After colloidal Coomassie staining **(Section 3.18.1, Methods)**, proteins of interest were excised from the gel using a razor blade. After excising the band, the gel slice was homogenized in protein gel band extraction buffer using a mortar and a pestle. After centrifugation at 10000 x g for 10 min at 4 °C the supernatant was retained and the pellet was re-extracted twice with the extraction buffer. Combined supernatants were precipitated by adding an equal volume of ice-cold acetone. After overnight incubation at -20 °C, purified protein was recovered by centrifuging at 10000 x rpm for 30 min at 4 °C. After air-drying, the pellet was resuspended in double distilled H_2O or protein gel band extraction buffer depending on their subsequent use.

3.20 Immunodetection of fusion protein and peptides

Peptides and proteins were transferred onto 0.2 µm PVDF membranes (Immobilon-P^{SQ} transfer membrane, Millipore) designed for protein sequencing and immunodetection of proteins <20 kDa from electrophoresis gels. In this study, immunodetection was used either to detect the presence of recombinant proteins or to obtain semiquantitative data on immunodetected proteins.

3.20.1 Preparation of a polyclonal antibody against the 18 kDa rice oleosin

Polyclonal antibodies against 18 kDa rice oleosin were produced in rabbits. For this, the 18 kDa oleosin was purified from oil bodies **(Sections 3.16.1.2 and 3.16.2.2, Methods)**. Following SDS-PAGE **(Section 3.17.1, Methods)**, the

18 kDa oleosin was purified from the gel **(Section 3.17.3, Methods)**. Three triweekly injections (50 µg in each injection) were carried out, and the rabbits were bled 8 days after the last injection. Preimmune serum was collected from rabbits 1 week before immunization. Polyclonal antibodies were produced at the IBMB-CSIC centre (Barcelona, Spain).

3.20.2 Western blot analysis: Transfer and immunodetection of proteins and peptides on membranes

Following electrophoretic separation of protein mixtures by SDS-PAGE or Tris-Tricine SDS-PAGE, the proteins/peptides were transferred onto membranes using Mini Trans-Blot Cell Assembly (Bio-Rad), according to the manufacturer's instructions with minor modifications. For this, the resolving gel was removed from the gel cassette and soaked in transfer buffer for at least 30 min (with filter paper, fiber pads and membrane) The PVDF membrane is highly hydrophobic, so it is prewetted with 100% methanol and rinsed with double distilled water. A blotting sandwich was prepared from equilibrated gel, membrane, filter paper and fiber pads according to the transfer unit manual. Electrophoretic transfer was done at 100 V constant for 1 h at room temperature or at 30 V constant for overnight transfer at 4 °C. Next, the membrane was washed twice in water to remove any contaminating transfer buffer and blocked in 20 mL of blocking buffer to prevent nonspecific binding of antibodies to the membrane surface. After incubation for 80 min at room temperature with gently shaking, the membrane was washed three times in 20 mL of PBST for 5 min each step. The primary antibody was used as a dilution of 1:2000 or 1:3000 (depending on the primary antibody) in blocking solution. The membrane was incubated in the primary antibody solution for 1 h at room temperature or overnight at 4 °C with shaking. After six wash steps each in 20 mL of PBST for 5 min, the membrane was incubated in the secondary antibody solution for 1 h at room temperature with shacking, followed by another six wash steps. Depending on the detection method, two types of detection systems were used in this work: the chromogenic method using an Anti-Rabbit IgG (Fc) Alkaline Phosphatase conjugate (Promega) as the secondary antibody (1:7500 dilution in blocking solution), and the chemiluminescent method using an

Anti-Rabbit IgGHRP linked (GE Healthcare) secondary antibody (1:100000 dilution). The Western Blue®Stabilized substrate served for visualization of Alkaline Phosphatase activity (Promega). Luminata™ Forte Western HRP Substrate (Millipore) was used for chemiluminescent detection of immunological reactions that were viewed using the ChemiDoc™ XRS+ System (Bio-Rad, USA).

3.21 Protein and peptide chromatography

Chromatography was used to either purify or detect the protein or peptide of interest taking advantage of its specific binding properties to an immobilized ligand.

3.21.1 Solid Phase purification of peptides using the Weak Cation Exchange-Reverse Phase (Strata X-CW)

To purify recombinant peptides produced by the transformed rice seed, chromatography method was based on their high positive charge. Thus, Weak Cation Exchange-Reverse Phase (Strata X-CW, Phenomenex) was used following the manufacturer's protocol, adjusting it to strong cationic peptide properties. Seed protein extracts **(Section 3.16.1, Methods)** were either precipitated **(Section 3.16.2, Methods)** or directly loaded into Strata X-CW columns. In both cases, double distilled H_2O containing 0.1% TFA was used as a loading buffer. Standard peptides and protein extracts from non-transformed seed plants were included in each purification process. The optimized conditions for purifying cationic peptides from rice protein extracts were compiled as follows:

Table 3.16 Solid Phase extraction procedure using acetic acid as elution solvent

Step	Solvent/action	# Recovered fraction
Conditioning	2 mL methanol	-
Equilibration	2 mL double distilled H_2O	-
Load	1,5 mL protein extract resuspended in double distilled H_2O:0,1% TFA	-
Wash 1	2 mL double distilled H_2O:0,1% TFA	1
Wash 2	2 mL methanol	2
Dry	1 min (N_2gas)	
Elution 1	2 mL 50% acetic acid	3
Elution 2	2 mL 50% acetic acid	4
Elution 3	1 mL 100% acetic acid	5

Table 3.17 Solid Phase extraction procedure using acetic acid followed by methanol as elution solvents

Step	Solvent/action	# Recovered fraction
Conditioning	2 mL methanol	-
Equilibration	2 mL double distilled H_2O	-
Load	1,5 mL protein extract resuspended in double distilled H_2O:0,1% TFA	-
Wash	2 mL double distilled H_2O:0,1% TFA	1
Dry	1 min (N_2gas)	-
Elution 1	2 mL 50% acetic acid	2
Elution 2	2 mL methanol	3

All recovered fractions were lyophilized and resuspended in double distilled H_2O. Then, samples were analysed on Tris-Tricine SDS-PAGE gel and Western blot for detection of the expected recombinant protein or peptide. Appropriate fractions were stored at -20 °C until further processing.

3.21.2 Analytical High Performance Liquid Chromatography (HPLC)

Standard peptides and selected fractions eluted from Strata X-CW columns were analysed in an Agilent 1200 HPLC system (Agilent Technologies, Palo Alto, CA) using an analytical reverse-phase C_{18}-Kromasil 100 column (40x4.6 mm, 3.5 µm particle size). Peptides were eluted using a linear gradient of 0-100% acetonitrile in 0.1% TFA at flow rate of 1.0 mL/min over 6 min. The injection volumes ranged from 10 to 100 µL depending on the analysis. UV

detector was used to monitor the UV absorbance of samples at 220 nm and the purity was estimated integrating the area under elution peaks.

Calibration curves were generated using a 10-point standard over a concentration range of 10 to 100 µM for BP134 standard peptide, allowing to check the performance obtained from Strata X-CW eluted fractions.

3.22 Mass spectrometry

LC-MS/MS identification of the AMPder peptides was carried out by Esther Izquierdo and M. Rossignol (Mass Spectrometry Proteomics Platform-MSPP, Laboratoire de Protéomique Fonctionnelle, INRA, Montpellier, France) within the framework of the European Trilateral Plant KBBE project SEPSAPE (Safe and Efficient Plant Systems for Antimicrobial PEptide production).

3.22.1 Preparation of protein body proteins from transgenic rice seeds

Protein body enriched fractions **(Section 3.16.1.1, Methods)** were diluted in 15 volumes of ultrapure water and centrifuged at 18000 x g for 20 min. Pelleted protein bodies and standard peptides (1, 10 and 100 ng) were solubilized in TT1X buffer (in agitation for 2 h) and separated by Tris-Tricine SDS-PAGE gel (16.5%). Electrophoresis was performed at 30 V for 1 h and at 80 V for 1 h. Samples were separated in only 3 cm in order to keep the proteins concentrated in small gel pieces. Gels were stained using Coomasie blue and three bands were excised from the gel between the dye front to the 20 kDa marker. For in-gel digestion, gel bands were cut in small squares of 2 mm and washed sequentially in double distilled water, 25 mM NH_4HCO_3, pH 8.0, 50 % 25 mM NH_4HCO_3/50% acetonitrile and finally in acetonitrile. Acetonitrile was removed and after gel pieces were dry (at room temperature), 60 µL of trypsin solution (0.0125 µg/µL, Promega) dissolved in 25 mM NH_4HCO_3 were added to each sample and incubated overnight at 37°C for protein digestion. Peptides were extracted with 50 µL of 2% (v/v) formic acid and dried.

3.22.2 Preparation of oil body proteins from transgenic rice seeds

Oil bodies **(Section 3.16.1.2, Methods)** were diluted in 6 volumes 25 mM NH_4HCO_3, 10 µL of trypsin (0.1 µg/µL, Promega) and incubated overnight at 37 °C. Increasing amounts of synthetic Cec A were treated similarly (1, 10 and 50 ng). After digestion, 10 µL of 2% (v/v) formic acid and samples were desalted on 100 mg Sep-Pak cartridges (Waters) by three washes of 0.1% formic acid and eluted with 30% ddH_2O:70% acetonitrile containing 0.1% formic acid. Eluted samples were dried and LC-MS analysis was performed as described for protein body samples.

3.22.3 Mass spectrometry analysis

For LC-MS analysis, peptide samples were dissolved in 0.1% (v/v) of formic acid and each sample was analysed using an Agilent 1200 nanoflow HPLC coupled to a 4000 QTRAP LC-MS/MS hybrid triple quadruple/linear ion trap mass spectrometer using a microSpray source (AB/MDS Sciex). Peptide samples were concentrated using a cartridge trap column (Algilent Zorbax 300 SB-C18; 0.3 mm id, 5 mm length, 3.5 mm particles) and separated on a nano-flow reverse-phase column (Aligent Zorbax 300 SB-C18; 0.075 mm id, 150 mm length, 3.5 mm particles) eluted with a linear gradient from 10 to 50% acetonitrile in 0.1% formic acid (42 min; 400 nL/min).

For each peptide, at least three transitions were selected, according to their relative intensity and m/z greater than the precursor m/z. Collision energy (CE) and declustering potential (DP) were optimized for maximum sensitivity of each transition.

Table 3.18 Selected transitions and parameters for the peptides produced in rice seeds

Peptide	Precursor ion (m/z)	Charge	DP (V)	Product ion (m/z)	CE (V)
Cec A	741.0	+2	140	816.4	33
				915.5	37
				1085.0	36
BP192	390.8	+2	80	504.5	22
				617.4	22
				659.4	22
BP178	474.0	+2	100	783.4	23
				670.3	26
				599.3	29
				542.3	27

3.23 Antimicrobial activity of the *in planta* produced AMPs

Antimicrobial peptides produced by transgenic plants were evaluated for their antibacterial activity against fungal and bacterial phytopathogens. Also, transgenic rice seeds were tested for resistance to pathogen infection.

3.23.1 Antibacterial activity of protein extracts from oil bodies of transgenic plants

Antibacterial activity assays were carried out with recombinant peptides produced in transgenic plants. Here, seed protein extracts were assayed for both antibacterial and bactericidal activity against the bacterial phytopathogen *Dickeya* sp. 1552.10.1. Protein extracts from non-transformed and pCAMBIA1300 empty vector seeds were used as controls.

3.23.1.1 Antibacterial activity of the Cec A produced in oil bodies of transgenic rice seeds

The antimicrobial activity of the *in planta* produced Cec A was assessed against the *Dickeya* sp. 1552 10.1 and compared to that of synthetic Cec A. Oil bodies were purified from transgenic and control (empty vector) seeds and proteins were recovered by precipitation as described in **Sections 3.16.1.2 and 3.16.2.2, Methods**. Subsequently, the fusion protein was

digested using the TEV protease technology **(Section 3.16.3, Methods)**. Synthetic Cec A was solubilized in double distilled H_2O to a final concentration of 1 mM, filter sterilized through a 0.2 µm pore filter (Whatman®) and subsequently diluted to obtain a final concentration of 0.097, 0.195, 0.250, 0.390, 0.500, 0.780, 1.000 and 1.560 µM. Twenty-five µL of each synthetic peptide dilution or OB protein extracts (digested and not digested-OB) were mixed in a microtiter plate with 75 µL of the suspension of bacterial indicator at 10^4 cfu/mL.
Bacterial cell concentration in suspension was automatically determined by optical density measurement at 600 nm (Bioscreen C, Labsystem, Helsinki, Finland). Microplates were incubated at 25 °C with 20 s shaking before hourly absorbance measurement for 24 h. Each experiment was repeated twice.

3.23.1.2 Bactericidal activity of the *in planta* produced Cec A peptide

The bactericidal activity of the plant-derived Cec A peptide was assessed by means of the contact time exposure test of the pathogenic bacteria to a peptide solution in water and compared to that of synthetic Cec A peptide. Oil bodies were purified from transgenic and control (empty vector) seeds and proteins were precipitated **(Sections 3.16.1.2 and 3.16.2.2, Methods)**. Subsequently, fusion protein was digested using the TEV protease technology. Synthetic Cec A was solubilized in double distilled H_2O to a final concentration of 1 mM, filter sterilized through a 0.2 µm pore filter (Whatman®) and subsequently diluted to obtain a final concentration of 0.097, 0.195, 0.250, 0.390, 0.500, 0.780, 1.000 and 1.560 µM.
Twenty-five µL of each synthetic peptide dilution or OB protein extracts (digested and not digested-OB) were mixed in a microtiter plate with 75 µL of the suspension of bacterial indicator at 10^4 cfu/mL. Microplates were incubated at 25 °C with shaking. The minimal killing concentration (MKC) was evaluated by the viable plate count method 30 min after contact exposure. Three replicates of each strain, peptide and concentration were used as well as positive controls containing water instead of peptide and negative controls containing peptide without bacterial suspension.

3.23.2 Direct seed antimicrobial activity bioassays

A seed bioassay consisting of seed inoculation with a bacterial plant pathogen was optimized to evaluate resistance to infection. Five plant pathogenic bacterial strains were tested in order to identify the most suitable one for the rice seed infection assays. Fifty-four dehulled wild type seeds (Senia) were surface sterilized and placed in groups of three in 24-well culture chambers (Sigma) with increasing concentrations of the pathogenic strains (10^6, 10^7 and 10^8) **(Fig. 3.8)**. The plant pathogenic bacterial strains assayed in this work are shown in **Table 3.19**.

Fig. 3.8 Scheme representing a typical seed activity bioassay. This experiment was performed for each plant pathogenic bacteria (and concentration assayed).

Table 3.19 Bacterial strains used to set up the seed activity bioassay

Code	Bacterialstrain
B1	*Dickeya* sp. 1552.10.1
B2	*Dickeya* sp. 1428 1a
B3	*Pectobacterium carotovorum* 578.3
B4	*Pectobacterium carotovorum* 3902.3
B5	*Pectobacterium carotovorum* 558-6-b

Bacterial strains were provided by M.López, Instituto Valenciano de Investigaciones Agrarias, Valencia, Spain.

Inoculation was done by adding 500 µL of the bacterial suspension to the seeds in each well. After seed inoculation, the 24-well culture chambers

were subjected under vacuum infiltration and incubated at 28 °C for seven days using a 12 h/12 h (day/night) photoperiod under a photon flux of 110-150 µmol. m^{-2}. s^{-1} (Versatile Environmental Test Chamber, Sanyo (currently Panasonic) Japan). The development of infected seedlings was determined using a semi-quantitative scale **(Table 3.19 and Fig. 3.6).** Control seed incubated with water were used to check seed germination.

Table 3.20 Scale used to asses resistance to bacterial infection in the seed activity bioassay

Scale	Description of seedling growth
0	Not germinated seed
0.5	Germinated seed. The seedling length was less than 25% of control seedling
1.0	Germinated seed. The seedling length was less than 50% of control seedling and presented an altered root or shoot development
1.5	Germinated seed. The seedling length was less than 75% of control seedling and presented an altered root or shoot development
2.0	Germinated seed. The seedling presented the same morphology as the control seedling

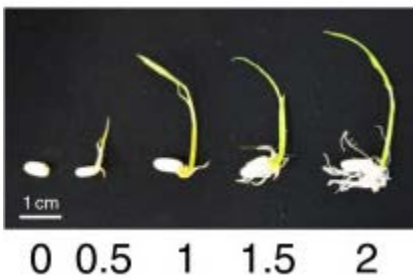


Fig. 3.9 Representative examples of seedling growth after inoculation with *Dickeya* sp. 1552.10.1. Picture was taken from Nadal et al., 2012.

After selection of the pathogenic strain and proper bacterial concentration (cfu/mL), seed infection bioassays were performed with transgenic, Senia and empty vector seeds. The incidence and the severity of the disease were determined. Resistance was compared to that of Senia and empty vector-

transformed plants. Control seeds inoculated with water were used to check seed germination.

To test resistance of transgenic rice seeds to the seed-borne pathogen *Fusarium verticillioides*, seed from transgenic and control plants (WT and empty vector plants) were pregerminated for 6 hours on sterile filter-watered paper. Subsequently, seed were inoculated with 50 µl of a suspension of *F. verticillioides* spores at a concentration of 10^3 spores/ml (or double distilled water as germination control) on MS medium without sucrose. Seeds allowed to continue germination. Inhibition of germination was evaluated after inoculation with fungal suspension and compared with that of control seeds (inoculated with double distilled water), seven days after inoculation. *F. verticillioides* isolate was obtained from rice plants in Spain and provided by the Facilities of Plant Protection Service of the Generalitat de Catalunya.

Fusarium infection bioassays were carried out by M. Coca (CRAG Centre, Barcelona, Spain) within the framework of the European Trilateral Plant KBBE project SEPSAPE (Safe and Efficient Plant Systems for Antimicrobial PEptide production).

4. RESULTS

4.1 Design and characterization of BP100 derivatives for plant expression

In the present section, different peptides (BP134 derivatives) were designed and evaluated for antimicrobial activity against phytopathogens. When designing these peptides, the specific requirements for the expression of the corresponding *AMP* genes in transgenic rice were considered (i.e. multimerization, presence/absence of targeting signals for AMP accumulation in specific subcellular compartments, *AMP* fusion genes containing a cleavage site for specific protease). The aim of these studies was to identify the best AMP candidates for subsequent production in transgenic rice, while maintaining or improving their antimicrobial activity and minimizing potential toxicity.

4.1.1 Strategy of peptide design for plant expression

The strategy of peptide design for plant expression arises from the information gained in previous studies for the antimicrobial peptide BP100 (Badosa et al., 2007). This peptide belongs to a chemical library of linear undecapeptides (cecropin A (1-7)-melittin (2-9) hybrids) obtained by combinatorial chemistry (named CECMEL11). The CECMEL11 peptide library comprises 125 compounds that exhibit antimicrobial activity against phytopathogenic bacteria and are able to control infection in host plants. They exhibit low cytotoxicity and moderate susceptibility to proteolytic degradation (Ferre et al., 2006; Badosa et al., 2007).

Thus, the present work was based on the leader peptide BP100 (amidated form), which has been adapted for biotechnological production using the non-amidated form BP134 (H-KKLFKKILKYL-OH).

As previously mentioned, the heterologous expression of such type of peptide requires overcoming several challenges to allow its expressability in plants, such as the need to achieve high expression levels, to avoid loss of

activity due to degradation by host proteases and the need to have a minimum size. Our design strategy aimed to solve these problems by increasing the size of the peptide (elongation, n-merization) and including sequences for stabilization or retention into the endoplasmic reticulum of the plant cell. However, any change or addition in the basic sequence of BP134 may affect its spectrum and antimicrobial activity or its potential phytotoxic effects or hemolytic activity (Badosa et al., 2007). In order to determine whether the aforementioned adjustments have an impact on the properties of this peptide, a library of 52 peptides based on BP134 was synthesized at the LIPPSO Laboratory (University of Girona, Spain) (In some cases peptides were synthesized by CASLO Laboratory ApS (Lyngby, Denmark)) within the framework of the European Trilateral Plant KBBE project SEPSAPE (Safe and Efficient Plant Systems for Antimicrobial PEptide production **(Fig. 4.1)**.

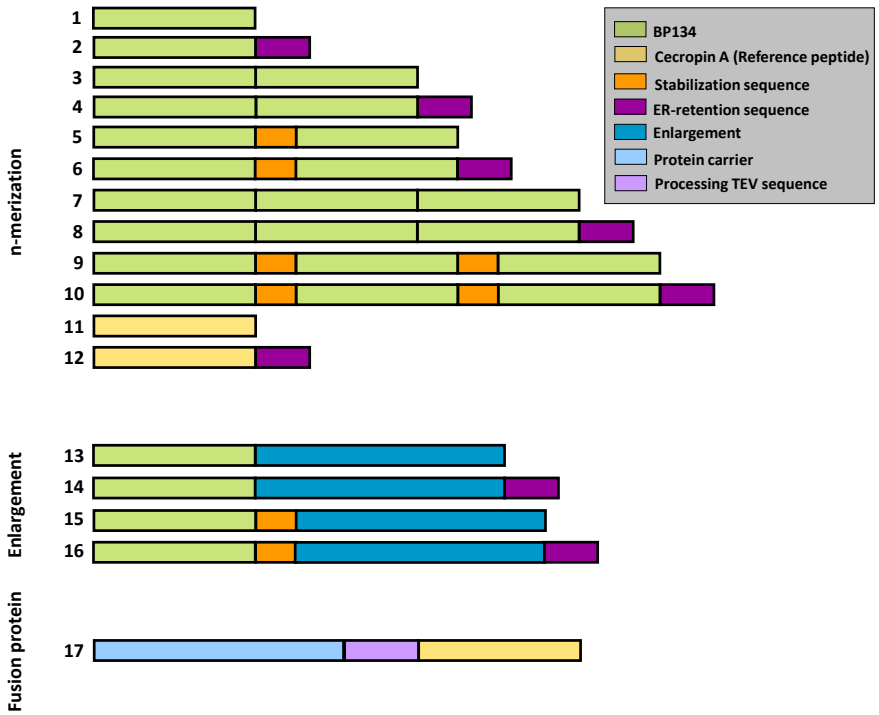


Fig. 4.1 Schematic representation of the peptide design for plant expression in rice plants. BP134, KKLFKKILKYL-OH a cecropin A (1-7)-melittin (2-9) hybrid; cecropin A, KWKLFKKIEKVGQNIRDGIIKAGPAVAVVGQATQIAK-OH; cecropin A stabilization sequence, AGPA (hinge sequence); ER retention signal, KDEL; Protein carrier, 18 kDa rice oleosin (GenBank entry, AAD10240); Processing TEV sequence, PTTENLYFQS. For details, see **Table 4.1**.

4.1.1.1 Endoplasmatic reticulum (ER) retention signal sequence (KDEL)

A series of BP134-derived peptides (BP134ders) were prepared in which the C-terminal KDEL ER retention signal sequence was included. Several reports indicate that the KDEL tetrapeptide increases accumulation of certain fusion proteins and reduces the risk of toxicity to plant cells. As an example, Takagi and co-workers (2005a) obtained high amounts of a peptide vaccine using the KDEL retention signal. In addition, Coca and co-workers (2006) observed that the omission of the retention signal resulted in a reduction in cecropin A levels compared with those obtained using the ER retention signal in transgenic rice. Whereas no effect on plant phenotype was observed in rice plants harbouring the *cecropin A_KDEL* gene, most of the rice lines expressing the *cecropin A* gene with no KDEL extension were non-fertile. Furthermore, when expressing a cecropin A hybrid peptide without ER retention signal (called Pep 3) in rice plants under the control of a constitutive promoter, the peptide was found to be toxic to the rice plants (San Segundo, unpublished results).

Thus, based on the above mentioned reports, in this work two different strategies were designed based on the absence or presence of the KDEL ER retention signal at the C-terminal end of the BP134 peptide **(Fig. 4.1, peptides 1 and 2)**.

On the other hand, it is well known that the entry of proteins into the ER endomembrane system is dependent on the presence of a transient N-terminal peptide. A unique feature of plants is that they are able to store proteins in the ER in addition to other endomembrane compartments. In this work, the N-terminal signal peptide of either the GluB-1, GluB-4 or Glb-1 rice storage proteins were used to allow translocation of candidate peptides into the ER system of the plant cell. This issue will be further developed in **Section 4.2, Results**.

4.1.1.2 Stabilization sequence (AGPA)

Cec A is an insect antimicrobial peptide in which an N-terminal basic, amphipathic α-helix (residues 5 to 21) is linked by an AGPA hinge to a shorter and more hydrophobic C-terminal helix (residues from 24 to 37)

(Bulet and Stöcklin, 2005). As discussed previously, de novo designed peptides were synthesized departing from BP134, a cecropin-melittin hybrid. BP134 sequence comprises seven residues of cecropin A, a cationic antimicrobial peptide composed entirely of ordinary 37 L-amino acids. Because the stability of the α- helix in enlarged BP134 derivatives may be affected, the AGPA hinge region was included between each monomer subunit of the designed peptides **(Fig. 4.1**, **peptides 5 and 6; 9 and 10; 15 and 16).**

4.1.1.3 Sequence enlargement

Several reports have shown the difficulty to express genes encoding peptides less than 50 residues in length in plant cells.
Therefore, for expression of genes encoding BP134 derivatives in rice, the length of this peptide was increased. Three strategies have been explored: (1) fusion of the peptide sequence to a rice protein (carrier), namely the rice oleosin **(Fig. 4.1, peptide 17)** (2) peptide tandem repeats **(Fig. 4.1, peptides 3-8)**, and (3) chimeric enlargement **(Fig. 4.1, peptides 13-16)**.

4.1.1.4 Cleavage site for the Tobacco Etch Virus (TEV) protease

The TEV protease specifically recognises the seven amino acid sequence ENLYFQS/G, cleaving between the Q and S/G residues. In the present work PTTENLYFQS, the recognition site for TEV (plus three extra amino acids at its N-terminal end) was used to fuse the cecropin A peptide to the oleosin protein (oleosin_TEV_cecropin A) **(Fig. 4.1, Fusion protein, 17)**. The PTTENLYFQS sequence corresponds to the naturally occurring processing site in the TEV polyprotein that is cleaved more efficiently than the canonical recognition site (ENLYFQS). *A priori* the presence of the cleavage site for the TEV protease will allow the peptide to be released from the *in planta* produced fusion protein by an enzymatic treatment after extraction and purification of the fusion product (Dougherty et al., 1988).

4.1.2 Candidate peptide sequences

The peptide library was designed based on the ideal α-helical wheel diagram of H-KKLFKKILKYL-OH (BP134) and comprised 52 peptides. The sequences of the *de novo* designed peptides are presented in **Table 4.1**.

Table 4.1 Amino acid sequences of BP100 derivative candidates for plant expression

Code	Sequence[a]	Number of Aa residues
BP100	KKLFKKILKYL-NH_2	11
BP134	KKLFKKILKYL-OH	11
BP134 derivatives and n-mers		
BP184	K**L**FKKILSKLFKILSKILKYL-OH	21
BP185	K**W**FKKILSKLFKILSKILKYL-OH	21
BP186	K**L**FKKILSKLFKILSKILKYL - **KDEL**-OH	25
BP187	K**W**FKKILSKLFKILSKILKYL - **KDEL**-OH	25
BP199	KKLFKKILKYL- **AGPA** –OH	15
BP214	KKLFKKILKYL- **KDEL-**OH	15
BP203	KKLFKKILKYL-KKLFKKILKYL-OH	22
BP202	KKLFKKILKYL-**AGPA-**KKLFKKILKYL-OH	26
BP193	**G**-KKLFKKILKYL - **AGPA -**KKLFKKILKYL-OH	27
BP195	**S**-KKLFKKILKYL- **AGPA -**KKLFKKILKYL-OH	27
BP200	KKLFKKILKYL- **AGPA -**KKLFKKILKYL- **AGPA**-OH	30
BP213	KKLFKKILKYL- **AGPA -**LYKLIKKFLKK - **KDEL** – OH	30
BP198	KKLFKKILKYLKKLFKKILKYL - **KDEL**-OH	26
BP197	KKLFKKILKYL - **AGP** - KKLFKKILKYL - **KDEL**-OH	29
BP192	KKLFKKILKYL - **AGPA** - KKLFKKILKYL - **KDEL**-OH	30
BP194	**G** - KKLFKKILKYL - **AGPA** - KKLFKKILKYL - **KDEL**-OH	31
BP196	**S** - KKLFKKILKYL - **AGPA** - KKLFKKILKYL - **KDEL**-OH	31
BP204	KKLFKKILKYL - KKLFKKILKYL - KKLFKKILKYL-OH	33
BP201	KKLFKKILKYL - **AGPA** - KKLFKKILKYL - **AGPA** - KKLFKKILKYL-OH	41
BP216	KKLFKKILKYL - **AGPA** - KKLFKKILKYL - **AGPA** - KKLFKKILKYL-**KDEL-**OH	45

Table 4.1 (continued)

Code	Sequence[a]	Number of Aa residues
BP134 - Melittin (amino acids 10 to 19) chimeras		
BP169	KKLFKKI**K**KYL - TTGLPALISW – OH	21
BP205	**G**- KKLFKKI**K**KYL - TTGLPALISW – OH	22
BP206	**S**- KKLFKKI**K**KYL - TTGLPALISW – OH	22
BP170	KKLFKKILKYL - TTGLPALISW – OH	21
BP171	KKLFKKILKYL - **AGPA** - TTGLPALISW-OH	25
BP207	**G**- KKLFKKILKYL - **AGPA** - TTGLPALISW-OH	26
BP208	**S**- KKLFKKILKYL - **AGPA** - TTGLPALISW-OH	26
BP172	KKLFKKILKYL - TTGLPALISW - **KDEL**-OH	25
BP173	KKLFKKILKYL - **AGPA** - TTGLPALISW - **KDEL**-OH	29
BP134- Melittin (amino acids 1 to 13) chimeras		
BP189	KKLFKKILKYL - GIGAVLKVLTTGL - **KDEL**-OH	28
BP191	GIGAVLKVLTTGL - KKLFKKILKYL - **KDEL**-OH	28
BP134 - Melittin (amino acids 10 to 18) (inverted dimer) chimeras		
BP217	KKLFKKILKYL-TTGLPALIS-**AGPA-**SILAPLGTT-LYKLIKKFLKK-**KDEL**	48
BP134- Magainin (amino acids 4 to 10) chimeras		
BP180	KKLFKKILKYL - KFLHSAK-OH	18
BP134- Magainin (amino acids 4 to 10) chimeras		
BP181	KKLFKKILKYL - **AGPA** - KFLHSAK-OH	22
BP211	**G-**KKLFKKILKYL **- AGPA -** KFLHSAK-OH	23
BP212	**S-** KKLFKKILKYL - **AGPA** - KFLHSAK-OH	23
BP182	KKLFKKILKYL - KFLHSAK - **KDEL**-OH	22
BP134- Magainin (amino acids 4 to 10) chimeras		
BP183	KKLFKKILKYL - **AGPA** - KFLHSAK - **KDEL**-OH	26
BP134- Magainin (amino acids 1 to 10) chimeras		
BP174	KKLFKKIL**L**YL - GIGKFLHSAK-OH	21
BP175	KKLFKKILKYL - **AGPA** - GIGKFLHSAK-OH	25

Table 4.1 (continued)

Code	Sequence[a]	Number of Aa residues
BP209	**G-**KKLFKKILKYL - **AGPA** - GIGKFLHSAK-OH	26
BP210	**S-**KKLFKKILKYL - **AGPA** - GIGKFLHSAK-OH	26
BP176	KKLFKKILKYL - GIGKFLHSAK-OH	21
BP177	KKLFKKILKY**A** - GIGKFLHSAK - **KDEL**-OH	25
BP178	KKLFKKILKYL - **AGPA** - GIGKFLHSAK - **KDEL**-OH	29
BP179	KKLFKKILKYL - GIGKFLHSAK - **KDEL**-OH	25
BP134- Cecropin A (amino acids 25 to 37) chimeras		
BP188	KKLFKKILKYL - AVAVVGQATQIAK - **KDEL**-OH	28
BP215	KKLFKKILKYL - **AGPA** -VAVVGQATQIAK - **KDEL**-OH	31
BP190	AVAVVGQATQIAK - KKLFKKILKYL - **KDEL**-OH	28
Cecropin A and modifications		
Cec A	KWKLFKKIEKVGQNIRDGIIKAGPAVAVVGQATQIAK-OH	37
Cec A_KDEL	KWKLFKKIEKVGQNIRDGIIKAGPAVAVVGQATQIAK - **KDEL-**OH	41

(a) Boldfaced letters indicate the modifications introduced into the BP134 sequence.
Boldfaced underlined residues indicate the non-native Ser or Gly residue retained on N-terminal BP sequence after TEV protease digestion.

4.1.3 Antibacterial activity of the BP134 derivatives

We tested the antimicrobial activity of the newly synthesized BP134 derivatives by using *in vitro* growth inhibition assays. For these studies three plant pathogenic bacteria were used, namely *Xanthomonas axonopodis* pv. vesicatoria*, Pseudomonas syringae* pv. syringae and *Erwinia amylovora*. The MIC values were assessed by using increasing concentrations of peptide (0.6, 1.25, 2.5, 5.0, 7.5, 10, 12.5 and 20 μM). As shown in **Table 4.2**, the modifications incorporated in the BP134 derivatives strongly influenced the antimicrobial activity of the resulting peptides. Most BP134 derivatives (80.8%) exhibited higher antibacterial activity against one or another pathogen and 63.5% of *de novo* designed peptides had higher antibacterial activity against the three pathogen bacteria tested, compared to BP134. Interestingly, these peptides presented MIC values in the range of 0.6 to 10 μM. In general, the incorporation of stabilizing and ER retaining sequences,

as well as, its elongation or n-merization, resulted in a similar or even higher antibacterial activity compared to BP134. The BP134_melittin and magainin derivatives, cecropin A and cecropin A_KDEL were the most active peptides, whereas BP216 was the less active peptide **(Table 4.2)**.

Table 4.2 Antibacterial activity of BP134 derivatives

		MIC intervals (μM)[a]		
Code	**Sequence**	***Xav*[b]**	***Pss*[b]**	***Ea*[b]**
BP 100	H-KKLFKKILKYL-NH_2	2.5-5.0	5.0-7.5	2.5-5.0
BP134	KKLFKKILKYL-OH	12.5-25	6.25-12.5	6.25-12.5
BP134 derivatives and n-mers				
BP184	K**L**FKKILSKLFKILSKILKYL-OH	10-20	10-20	10-20
BP185	K**W**FKKILSKLFKILSKILKYL-OH	10-20	10-20	10-20
BP186	K**L**FKKILSKLFKILSKILKYL - **KDEL**-OH	10-20	7.5-10	10-20
BP187	K**W**FKKILSKLFKILSKILKYL - **KDEL**-OH	5-7.5.0	10-20	10-20
BP199	KKLFKKILKYL- **AGPA** –OH	2.5-5	10-20	10-12.5
BP214	KKLFKKILKYL- **KDEL-**OH	1.25-2.5	2.5-5.0	2.5-5.0
BP203	KKLFKKILKYL-KKLFKKILKYL-OH	1.25-2.5	5.0-7.5	10-20
BP202	KKLFKKILKYL-**AGPA-**KKLFKKILKYL-OH	0.6-1.25	2.5-5.0	2.5-5.0
BP193	**G -** KKLFKKILKYL - **AGPA -**KKLFKKILKYL-OH	10-20	5.0-10	10-20
BP195	**S -**K KLFKKILKYL- **AGPA -**KKLFKKILKYL-OH	10-20	5.0-10	10-20
BP200	KKLFKKILKYL- **AGPA -**KKLFKKILKYL- **AGPA**-OH	0.6-1.25	7.5-10	>20
BP213	KKLFKKILKYL- **AGPA -**LYKLIKKFLKK - **KDEL** – OH	1.25-2.5	2.5-5.0	2.5-5.0
BP198	KKLFKKILKYLKKLFKKILKYL - **KDEL**-OH	10-20	10-20	10-20
BP197	KKLFKKILKYL - **AGP** - KKLFKKILKYL - **KDEL**-OH *	5.0-10	5.0-10	5.0-10
BP192	KKLFKKILKYL - **AGPA** - KKLFKKILKYL - **KDEL**-OH	5.0-10	5.0-10	5.0-10
BP194	**G** - KKLFKKILKYL - **AGPA** - KKLFKKILKYL - **KDEL**-OH	10-20	5.0-10	10-20
BP196	**S** - KKLFKKILKYL - **AGPA** - KKLFKKILKYL - **KDEL**-OH	10-20	5.0-10	10-20
BP204	KKLFKKILKYL - KKLFKKILKYL - KKLFKKILKYL-OH	2.5-5.0	10-20	>20
BP201	KKLFKKILKYL - **AGPA** - KKLFKKILKYL - **AGPA** - KKLFKKILKYL-OH	1.25-2.5	7.5-10	>20
BP216	KKLFKKILKYL - **AGPA** - KKLFKKILKYL - **AGPA** - KKLFKKILKYL-**KDEL**-OH	10-20	>20	>20
BP134 - Melittin (amino acids 10 to 19) chimeras				
BP169	KKLFKKI**K**KYL - TTGLPALISW – OH	2.5-5.0	<2.5	2.5-5.0
BP205	**G**- KKLFKKI**K**KYL - TTGLPALISW – OH	0.6-1.25	1.25-2.5	1.25-2.5
BP206	**S**- KKLFKKI**K**KYL - TTGLPALISW – OH	0.6-1.25	1.25-2.5	1.25-2.5
BP170	KKLFKKILKYL - TTGLPALISW – OH	<2.5	2.5-5.0	2.5-5.0
BP171	KKLFKKILKYL - **AGPA** - TTGLPALISW-OH	2.5-5.0	<2.5	2.5-5.0
BP207	**G-** KKLFKKILKYL - **AGPA** - TTGLPALISW-OH	0.6-1.25	1.25-2.5	1.25-2.5
BP208	**S**- KKLFKKILKYL - **AGPA** - TTGLPALISW-OH	0.6-1.25	1.25-2.5	1.25-2.5
BP172	KKLFKKILKYL - TTGLPALISW - **KDEL**-OH	<2.5	2.5-5.0	2.5-5.0
BP173	KKLFKKILKYL - **AGPA** - TTGLPALISW - **KDEL**-OH	5.0-7.5	2.5-5.0	5.0-7.5

Table 4.2 (continued)

Code	Sequence	MIC intervals (μM)[a]		
		Xav[b]	*Pss*[b]	*Ea*[b]
BP134- Melittin (amino acids 1 to 13) chimeras				
BP189	KKLFKKILKYL - GIGAVLKVLTTGL - **KDEL**-OH	2.5-5.0	2.5-5.0	5.0-7.5
BP191	GIGAVLKVLTTGL - KKLFKKILKYL - **KDEL**-OH	2.5-5.0	>20	>20
BP134 - Melittin (amino acids 10 to 18) (inverted dimer) chimeras				
BP217	KKLFKKILKYL-TTGLPALIS-**AGPA**-SILAPLGTT-LYKLIKKFLKK-**KDEL**	<0.6	2.5-5.0	2.5-5.0
BP134- Magainin (amino acids 4 to 10) chimeras				
BP180	KKLFKKILKYL - KFLHSAK-OH	2.5-5.0	2.5-5.0	2.5-5.0
BP181	KKLFKKILKYL - **AGPA** - KFLHSAK-OH	2.5-5.0	<2.5	2.5-5.0
BP211	**G-**KKLFKKILKYL **- AGPA -** KFLHSAK-OH	0.6-1.25	2.5-5.0	1.25-2.5
BP212	**S-** KKLFKKILKYL - **AGPA** - KFLHSAK-OH	0.6-1.25	2.5-5.0	2.5-5.0
BP182	KKLFKKILKYL - KFLHSAK - **KDEL**-OH	<2.5	<2.5	2.5-5.0
BP183	KKLFKKILKYL - **AGPA** - KFLHSAK - **KDEL**-OH	5.0-7.5	7.5-10	2.5-5.0
BP134- Magainin (amino acids 1 to 10) chimeras				
BP174	KKLFKKI**LL**YL - GIGKFLHSAK-OH	5.0-7.5	10-20	5.0-7.5
BP175	KKLFKKILKYL - **AGPA** - GIGKFLHSAK-OH	1.25-2.5	2.5-5.0	5.0-7.5
BP209	**G-**KKLFKKILKYL - **AGPA** - GIGKFLHSAK-OH	0.6-1.25	2.5-5.0	1.25-2.5
BP210	**S-**KKLFKKILKYL - **AGPA** - GIGKFLHSAK-OH	0.6-1.25	2.5-5.0	1.25-2.5
BP176	KKLFKKILKYL - GIGKFLHSAK-OH	2.5-5.0	2.5-5.0	5.0-7.5
BP177	KKLFKKILKY**A** - GIGKFLHSAK - **KDEL**-OH	5.0-7.5	2.5-5.0	7.5-10
BP178	KKLFKKILKYL - **AGPA** - GIGKFLHSAK - **KDEL**-OH	2.5-5.0	2.5-5.0	2.5-5.0
BP179	KKLFKKILKYL - GIGKFLHSAK - **KDEL**-OH	2.5-5.0	2.5-5.0	5.0-7.5
BP134- Cecropin A (amino acids 25 to 37) melittin				
BP188	KKLFKKILKYL - AVAVVGQATQIAK - **KDEL**-OH	2.5-5.0	<2.5	2.5-5.0
BP215	KKLFKKILKYL - **AGPA** -VAVVGQATQIAK - **KDEL**-OH	1.25-2.5	2.5-5.0	2.5-5.0
BP190	AVAVVGQATQIAK - KKLFKKILKYL - **KDEL**-OH	5.0-7.5	>20	>20
Cecropin A and modifications				
Cec A	KWKLFKKIEKVGQNIRDGIIKAGPAVAVVGQATQIAK-OH	<1	<1	<1
Cec A_KDEL	KWKLFKKIEKVGQNIRDGIIKAGPAVAVVGQATQIAK-**KDEL-**OH	0.6-1.25	0.6-1.25	0.6-1.25

(a) Minimal inhibitory concentrations were determined with a concentration of target bacterial cells of 10^8 cfu/mL. (b) Antibacterial activity of BP134 derivatives was determined against *Erwinia amylovora* (*Ea*), *Pseudomonas syringae* pv. syringae (*Pss*) and *Xanthomonas axonopodis* pv. vesicatoria (*Xav*).

4.1.4 Hemolytic activity

In order to estimate the toxicity of BP134 derivatives to eukaryotic cells, the hemolytic activity of those AMPs showing MIC ≤ 10 μM (for the three pathogenic bacteria) was investigated. Their toxicity was determined as the ability to lyse human erythrocytes and compared to that of melittin (a highly hemolytic natural peptide) (Ferre et al., 2006; Badosa et al., 2007). The percent hemolysis of peptides at a final concentration of 150 μM after 1 h of incubation is shown in **Table 4.3**. Of interest, ten peptides showed lower hemolytic activity than BP134. Among them, BP181 and BP211 (BP134_melittin), BP178 (BP134_magainin), BP215 (BP134_cecropin A), Cec A and Cec A_KDEL exhibited the lowest hemolytic activity in human erythrocytes. The remaining peptides (67.6%) displayed high hemolytic activity compared to BP134. In fact, eight peptides exhibited >50% hemolysis at 150 μM.

Table 4.3 Hemolytic activity of BP134 derivatives with relevant antimicrobial activity (MIC ≤ 10 μM)

Code	Sequence	Hemolysis (%)[a]
BP 100	H-KKLFKKILKYL-NH_2	22.0 ± 2.8
BP134	KKLFKKILKYL-OH	7.8 ± 2.8
BP134 derivatives		
BP214	KKLFKKILKYL- **KDEL-**OH	6.0 ± 0.8
BP202	KKLFKKILKYL-**AGPA-**KKLFKKILKYL-OH	93.0 ± 3.3
BP213	KKLFKKILKYL- **AGPA -**LYKLIKKFLKK - **KDEL** – OH	92.0 ± 4.6
BP197	KKLFKKILKYL - **AGP** - KKLFKKILKYL - **KDEL**-OH	74.3 ± 2.4
BP192	KKLFKKILKYL - **AGPA** - KKLFKKILKYL - **KDEL**-OH	68.8 ± 4.3
BP134 - Melittin (amino acids 10 to 19)		
BP169	KKLFKKI**K**KYL - TTGLPALISW – OH	30.3 ± 1.2
BP205	**G**- KKLFKKI**K**KYL - TTGLPALISW – OH	13.0 ± 0.9
BP206	**S**- KKLFKKI**K**KYL - TTGLPALISW – OH	20.0 ± 1.2
BP170	KKLFKKILKYL - TTGLPALISW – OH	93.0 ± 3.8
BP171	KKLFKKILKYL - **AGPA** - TTGLPALISW-OH	15.8 ± 0.4
BP207	**G-** KKLFKKILKYL - **AGPA** - TTGLPALISW-OH	36.0 ± 1.1
BP208	**S**- KKLFKKILKYL - **AGPA** - TTGLPALISW-OH	47.0 ± 1.2
BP172	KKLFKKILKYL - TTGLPALISW - **KDEL**-OH	48.5 ± 4.7
BP173	KKLFKKILKYL - **AGPA** - TTGLPALISW - **KDEL**-OH	16.3 ± 6.5

Table 4.3 (continued)

Code	Sequence	Hemolysis (%)[a]
BP134- Melittin (amino acids 1 to 13)		
BP189	KKLFKKILKYL - GIGAVLKVLTTGL - **KDEL**-OH	60.2 ± 5.2
BP134 - Melittin (amino acids 10 to 18) (inverted dimer)		
BP217	KKLFKKILKYL-TTGLPALIS-**AGPA-**SILAPLGTT-LYKLIKKFLKK-**KDEL**	100.0 ± 3.9
BP134- Magainin (amino acids 4 to 10)		
BP180	KKLFKKILKYL - KFLHSAK-OH	54.4 ± 2.6
BP181	KKLFKKILKYL - **AGPA** - KFLHSAK-OH	<1
BP211	**G-**KKLFKKILKYL **- AGPA -** KFLHSAK-OH	1.0 ± 0.2
BP212	**S-** KKLFKKILKYL - **AGPA** - KFLHSAK-OH	5.0 ± 0.4
BP182	KKLFKKILKYL - KFLHSAK - **KDEL**-OH	38.1 ± 6.5
BP183	KKLFKKILKYL - **AGPA** - KFLHSAK - **KDEL**-OH	4.2 ± 1.3
BP134- Magainin (amino acids 1 to 10)		
BP175	KKLFKKILKYL - **AGPA** - GIGKFLHSAK-OH	14.3 ± 2.5
BP209	**G-**KKLFKKILKYL - **AGPA** - GIGKFLHSAK-OH	13.0 ± 0.7
BP210	**S-**KKLFKKILKYL - **AGPA** - GIGKFLHSAK-OH	17.0 ± 1.8
BP176	KKLFKKILKYL - GIGKFLHSAK-OH	43.7 ± 9.3
BP177	KKLFKKILKY**A** - GIGKFLHSAK - **KDEL**-OH	4.6 ± 0.4
BP178	KKLFKKILKYL - **AGPA** - GIGKFLHSAK - **KDEL**-OH	3.4 ± 4.14
BP179	KKLFKKILKYL - GIGKFLHSAK - **KDEL**-OH	12.3 ± 0.25
BP134- Cecropin A (amino acids 25 to 37)		
BP188	KKLFKKILKYL - AVAVVGQATQIAK - **KDEL**-OH	23.8 ± 4.0
BP215	KKLFKKILKYL - **AGPA** -VAVVGQATQIAK - **KDEL**-OH	3.0 ± 1.1
Cecropin A		
Cec A	KWKLFKKIEKVGQNIRDGIIKAGPAVAVVGQATQIAK-OH	<1
Cec A_KDEL	KWKLFKKIEKVGQNIRDGIIKAGPAVAVVGQATQIAK- **KDEL-**OH	<1

(a) Hemolytic activity at 150 μM peptide concentration with confidence interval for $\alpha = 0.05$.

4.1.5 Selection and further characterization of candidate peptides to be produced in rice plants

Taking into account the specific challenges that involves the expression of small peptides in plant systems, those peptides showing improved antimicrobial activity profile, lower cytoxicity and increased length, were selected as candidates for preparation of the corresponding synthetic genes. The ER retention signal and stabilization sequences were also considered at

selecting the candidate peptides **(Table 4.4)**. The following three BP134 derivative peptides were selected:

1) BP192, a dimer of BP134

2) BP178, a BP134_melittin hybrid

3) BP188, a BP134_Cec A hybrid which does not include the AGPA stabilizing sequence

The cecropin A and cecropin A_KDEL peptides were included in this study and used as references for peptide production in rice plants because they have been used in previous reports in constitutive expression strategies (Coca et al., 2006).

Table 4.4 Antibacterial and hemolytic activity of the selected candidates

		MIC intervals (µM)			
Code	**Sequence**	***Xav***	***Pss***	***Ea***	**Hemolysis[a]**
BP192	KKLFKKILKYL - **AGPA** - KKLFKKILKYL - **KDEL**-OH	5.0-10	5.0-10	5.0-10	68.8 ± 4.3
BP178	KKLFKKILKYL - **AGPA** - GIGKFLHSAK - **KDEL**-OH	2.5-5.0	2.5-5.0	2.5-5.0	3.4 ± 4.1
BP188	KKLFKKILKYL - AVAVVGQATQIAK - **KDEL**-OH	2.5-5.0	<2.5	2.5-5.0	23.8 ± 4.0
Cec A	KWKLFKKIEKVGQNIRDGIIKAGPAVAVVGQATQIAK-OH	<1	<1	<1	<1
Cec A_KDEL	KWKLFKKIEKVGQNIRDGIIKAGPAVAVVGQATQIAK- **KDEL-**OH	0.6-1.25	0.6-1.25	0.6-1.25	<1

(a) Hemolytic activity is given in % at 150 µM peptide concentration with confidence interval for α = 0.05.

4.1.5.1 Phytotoxicity

Further experiments were carried out to investigate whether the selected peptides show toxicity to plant cells. Peptide phytotoxicity was evaluated in tobacco leaves (*Nicotiana tabacum*) by infiltrating the leaf mesophyle with a peptide solution (100 µL, at 50, 100 or 150 µM each peptide, or water as a control). The highly cytotoxic peptide, melittin, was used as a positive control. Representative results are presented in **Fig. 4.2** and **Table 4.5.**

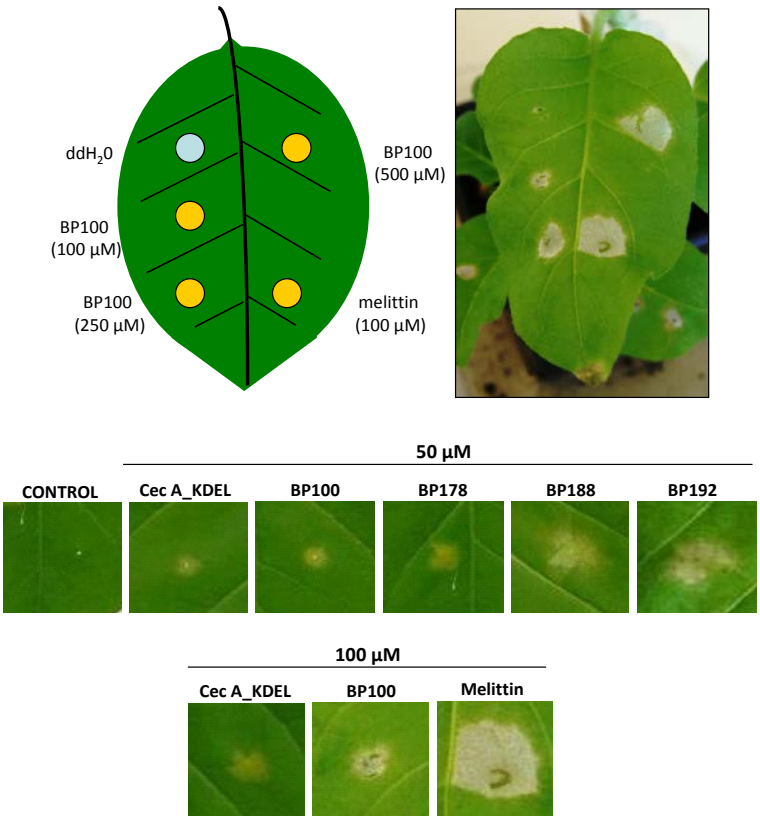


Fig. 4.2 Phytotoxicity of BP134 derivatives in *Nicotiana tabacum* leaves (concentrations 5 to 100 fold their MIC were used). Pictures were taken at 48 h after infiltration with the various peptides at the indicated concentrations. Representative images are shown.

The results obtained with the selected peptides were essentially in agreement with the hemolytic assay. At 48 h after infiltration, no necrotic lesions were observed in the water control while lesions appeared in the leaf areas where melittin was inoculated. The various peptides here assayed produced lesions in a concentration-dependent manner **(Fig. 4.2 and Table 4.5)**. The most phytotoxic peptide was BP192, which also exhibited the highest hemolytic activity among the selected peptides, whereas the less toxic were BP134, Cec A_KDEL and BP100. BP178 and BP188 were intermediate **(Table 4.3)**. No significant differences among the selected peptides were observed when the phytotoxicity was assessed at 150 µM ($p > 0.05$). By contrast, when phytotoxicity was tested using peptide concentrations in the range of 50-100 µM, differences in lesion areas were clearly observed ($p< 0.05$).

Table 4.5 Phytotoxicity in tobacco leaves (lesion diameter, cm) of the selected peptides

	Peptide concentration (µM)					
Peptide	**50**		**100**		**150**	
water control	0.00 ± 0.00	a	0.00 ± 0.00	a	0.00 ± 0.00	a
BP134	0.47 ± 0.03	b	0.50 ± 0.06	b	0.87 ± 0.07	b
Cec A_KDEL	0.47 ± 0.03	b	0.73 ± 0.09	bcd	1.07 ± 0.07	b
BP100	0.57 ± 0.07	bc	0.57 ± 0.03	bc	0.93 ± 0.03	b
BP178	0.67 ± 0.07	bcd	0.77 ± 0.03	cd	1.03 ± 0.14	b
BP188	0.77 ± 0.07	cd	0.80 ± 0.06	cd	1.23 ± 0.09	b
BP192	0.86 ± 0.10	d	0.97 ± 0.06	d	1.15 ± 0.07	b

Phytotoxicity values are given as the mean of the lesion diameter (cm) after peptide solution infiltration. The confidence interval for $\alpha = 0.05$ is indicated. Letters indicate significant differences according to Tukey one-way ANOVA ($\alpha = 0.05$).

4.1.5.2 Bactericidal activity

Bactericidal activity of selected peptides was assessed by a contact time exposure test of pathogenic bacteria to a peptide solution. This method is in fact a killing assay and differs from continuous exposure (bacteriostatic assay) used in screening of activity **(growth inhibition, Section 4.1.3, Results).** The selected peptides **(Table 4.6)** were assessed for bactericidal activity against *E. amylovora* and *Dickeya* sp. 1552.10.1. The minimal killing concentration (MKC) values are shown in **Tables 4.7**, and **4.8**, respectively. The MKC was dependent on the concentration of bacteria used for the assay and the exposure time. *E. amylovora* and *Dickeya* sp. 1552.10.1 did not survive when treated with BP192, BP178 or Cec A at a concentration <1.56 µM (30 min exposure period). Adding the KDEL retention signal either to the cecropin A (Cec A_KDEL) or to the BP134_Cec A (BP188) peptide decreased the bactericidal activity against *E. amylovora* (MKCs values of 1.56 to 12.5 µM, and 1.56 to 6.25 µM, respectively). By contrast, Cec A_KDEL showed a slight increase of activity against *Dickeya* sp. 1552.10.1. (1.56 to 6.25 µM), whereas BP188 still was less active against the pathogen.

Table 4.6 Amino acid sequence of the selected peptides tested for their bactericidal activity against *Erwinia amylovora* and *Dickeya* sp. 1552.10.1.

Peptide	Sequence
AMP models	
BP134	KKLFKKILKYL-OH
Magainin	GIGKFLHSAKKFGKAFVGEIMNS-OH
Cec A	KWKLFKKIEKVGQNIRDGIIKAGPAVAVVGQATQIAK-OH
AMPs selected in this work	
BP178_KDEL (BP134-magainin)	KKLFKKILKYL - AGPA - GIGKFLHSAK - KDEL-OH
BP188_KDEL (BP134-cecropin A)	KKLFKKILKYL - AVAVVGQATQIAK - KDEL-OH
BP192_KDEL (BP134-BP134)	KKLFKKILKYL - AGPA - KKLFKKILKYL - KDEL-OH
Cec A_KDEL	KWKLFKKIEKVGQNIRDGIIKAGPAVAVVGQATQIAK- KDEL-OH
Cec A	KWKLFKKIEKVGQNIRDGIIKAGPAVAVVGQATQIAK-OH

Table 4.7 Bactericidal activity (MKC, µM) of the selected peptides against *Erwinia amylovora*

Peptide	Bacterial concentration (CFU/mL)			
	10^3		10^4	
	Exposure time (min)		Exposure time (min)	
	30	120	30	120
BP192	<1.56	<1.56	<1.56	<1.56
BP188	1.56-3.12	<1.56	3.12-6.25	1.56-3.12
BP178	<1.56	<1.56	<1.56	<1.56
Cec A_KDEL	1.56-3.12	1.56-3.12	6.25-12.5	6.25-12.5
Cec A	<1.56	<1.56	<1.56	<1.56

Table 4.8 Bactericidal activity (MKC, µM) of the selected peptides against *Dickeya* sp. 1552.10.1

Peptide	Bacterial concentration (CFU/mL)			
	10^3		10^4	
	Exposure time (min)		Exposure time (min)	
	30	120	30	120
BP192	<1.56	<1.56	<1.56	<1.56
BP188	12.5-25.0	3.12-6.25	12.5-25.0	12.5-25.0
BP178	<1.56	<1.56	<1.56	<1.56
Cec A_KDEL	<1.56	<1.56	3.12-6.25	3.12-6.25
Cec A	<1.56	<1.56	<1.56	<1.56

Taking into account the results obtained in these studies, the BP134 derivatives BP192, BP188, BP178 and Cec A/Cec A_KDEL were selected to be produced in the rice seed platform. These peptides exhibited optimized features such as strong antibacterial and bactericidal effects, and low hemolytic and phytotoxic activity.

4.2. Construction of plant expression vectors

For expression of *AMP* genes in transgenic rice, seed specific promoters were initially obtained which were then used to drive expression of the selected AMP genes. Strategies to optimize transgene expression and to direct the transgene product to specific subcellular compartments, namely the protein bodies (endosperm) or oil bodies (embryo) were evaluated.
Transgenes were carefully designed taking into account the results obtained in the characterization of BP134 derivatives and for the peptides BP178_KDEL, BP188_KDEL and BP192_KDEL **(Section 4.1, Results)**.

4.2.1 Strategies for the expression of antimicrobial genes in rice seed tissues

As previously described **(Section 4.1, Results)**, transgene expression level is highly dependent on the regulatory elements used to control its expression, promoters and terminators. Generally, the stronger the promoter is, the higher protein yield can be expected (unless the transgene transcripts are unstable or recombinant protein is unstable or toxic to the plant cell). In order to hinder the possible negative effects of AMPs on plant growth and development (in our case, AMPs might display certain phytotoxicity) an approach for seed-specific expression of *AMPs* was assessed by using different seed-specific promoters to drive transgene expression. Four rice promoters were chosen on the basis of their activity and pattern of expression in the rice seed (Qu and Takaiwa, 2004). They were: *glutelin B-1* (*GluB-1*), *glutelin B-4* (*GluB-4*) and *26 kDa* α-*globulin* (*Glb-1*) promoters (active in the rice endosperm during seed germination), and the *18 kDa oleosin* (*Ole18*) promoter (active in the rice embryo and aleurone tissues). Moreover, the Ole18 protein was used as a carrier to produce the AMP of interest as an oleosin_AMP fusion protein in the seed oil bodies (van Roojien and Moloney, 1995). Towards this end, an *oleosin_AMP* fusion gene was obtained and expressed in rice plants under the control of the *18 kDa oleosin* promoter. The activity and strength of the seed specific promoters used in this work during rice seed development, are presented in **Table 4.9** according to Qu and Takaiwa (2004).

Table 4.9 Activity of the selected seed-specific promoters during seed development, as revealed by GUS activity of *promoter::GUS* fusion genes (according to Qu and Takaiwa, 2004).

Promoter	Expression pattern	Promoter strengh	Description
GluB-1, GluB-4	Outer endosperm	Moderate, high	Expression is observed in the aleurone layer and subaleurone tissues 7 days after flowering. Progressively spreads into the inner endosperm during maturation
26 kDa Glb-1	Inner endosperm	High	Expression in the inner starchy endosperm. The expression pattern does not change during seed development.
Ole 18	Embryo	Low	Expression is first detected in the aleurone layer 7 days after flowering, and in the embryo at later stages of maturation

The universal promoter *ubiquitin* is included in the low GUS activity group.

In the present work, the signal peptide sequence of the rice endosperm storage protein either GluB-1, GluB-4 or 26 kDa Glb-1 was fused to the N-terminal end of the AMP peptide sequence to direct its internalization into the ER. During translocation into the ER, the signal peptide is cleaved off by a signal peptidase (von Heinje, 1983, 1985; Günter Blobel 1999-Novel price in Physiology or Medicine: *"for the discovery that proteins have intrinsic signals that govern their transport and localization in the cell"*). When necessary, the ER retention signal (KDEL tetrapeptide) was added to the C-terminal end of the peptide**.** In all the cases, the *nopaline synthase* terminator (*Nos-T*) derived from *A. tumefaciens* was used to control transcriptional termination. The components of the different plant expression vectors prepared in this work are presented in **Fig. 4.3**.

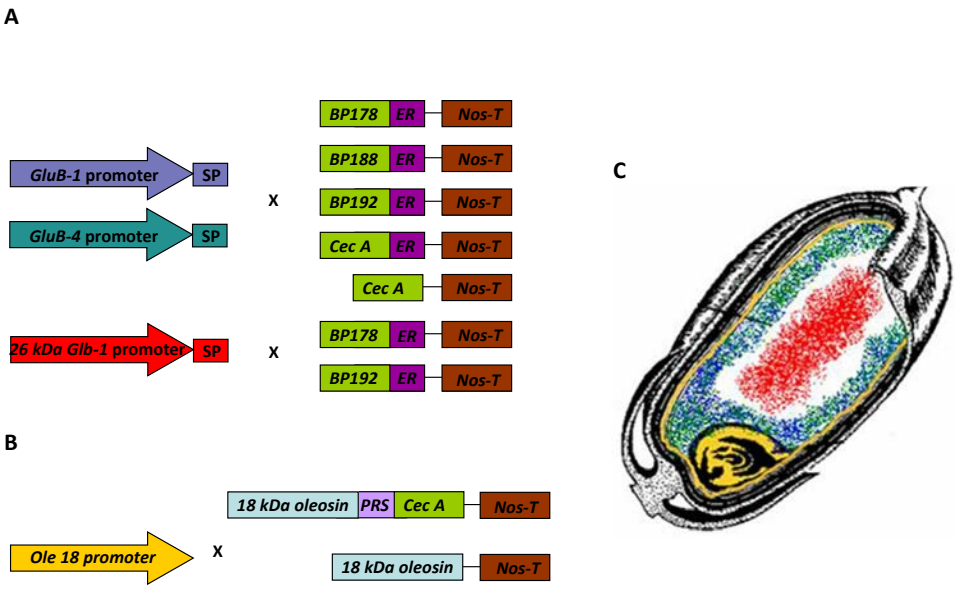


Fig. 4.3 (A) Schematic representation of the chimeric gene constructs used for expression of *AMP* genes in rice seed tissues, either endosperm (A) or embryo (B). Arrows indicate the orientation of the sequence. SP, signal peptide; ER, endoplasmatic reticulum retention signal (KDEL); PRS, processing site (TEV protease recognition sequence). (C) Expression pattern of rice promoters which direct embryo-specific or endosperm-specific expression. Highlighted areas indicate: blue, *GluB-1* promoter; green, *GluB-4* promoter; red, *Glb-1* promoter; yellow, *Ole18* promoter.

4.2.2 Synthesis and cloning of *AMP* genes

Synthetic genes were designed based on the codon usage bias in *O. sativa*. The codon usage pattern has been correlated with the gene expression level (Liu et al., 2004). Lysine residues in the peptide genes are preferentially encoded by AAG, isoleucine by ATC, valine by GTC, threonine by ACC, glycine by GGC or phenylalanine by TTC, among others (Liu et al., 2004). Similarly, the synthetic Cecropin A coding sequence was previously obtained by changing 20 out of the 37 codons of the native Cecropin A coding sequence (Coca et al., 2006).

Based on the amino acid sequence of BP192, BP178, BP188, Cec A_KDEL and Cec A, the corresponding codon-optimized *AMPs* genes were chemically synthesized. Synthetic *BP192* and *Cec A* genes were fully synthesized using recursive PCR (Prodromou and Pearl, 1992). This technique is used to generate new synthetic genes from overlapping complementary oligonucleotides without the need for precursor template DNA. *BP192* and

cecropin A coding sequences were synthesized by Maria Montero (University of Girona) and Maria Coca (CRAG, Barcelona), respectively. The coding sequence of BP178 and BP188 was chemically synthesized and cloned into the pUC57 plasmid (GenScrip).The full-length cDNA for the *18 kDa oleosin* gene (AK243015) was purchased from the Rice Genome Resource Center (National Institute of Agrobiological Science, Japan, http://www.nias.affrc.go.jp).

The nucleotide sequences of the synthetic genes prepared in the present work and deduced amino acid sequences are presented in **Table 4.10** and **4.11**, respectively.

Table 4.10 Nucleotide sequences of the synthetic *BP134 derivatives* and *Cec A* genes prepared in this work. The rice promoter used to drive the expression of these genes is indicated.

Promoter	Peptide	Peptide sequence	bp
GluB-1	BP178	aag aag ctt ttc aag aag atc ctc aag tac ctc gcc ggc cca K K L F K K I L K Y L A G P gcc ggc atc ggc aag ttc ctc cac tcc gcc aag aag gac gag A G I G K F L H S A K K D E ctc **tag taat** L	94
	BP188	aag aag ctt ttc aag aag atc ctc aag tac ctc gcc gtc gcc K K L F K K I L K Y L A V A gtc gtc ggc caa gcc acc caa atc gcc aag aag gac gag ctc V V G Q A T Q I A K K D E L **tag taat**	91
	BP192	aag aag ctt ttc aag aag atc ctc aag tac ctc gcc ggc cca K K L F K K I L K Y L A G P gcc aag aag ctc ttc aag aag atc ctc aag tac ctc aag gac A K K L F K K I L K Y L K D gag ctc **tag taat** E L	97
	Cec A_KDEL	aag tgg aag ctt ttc aag aag atc gag aag gtc ggc caa aac K W K L F K K I E K V G Q N atc cgc gac ggc atc atc aag gcc ggc cca gcc gtc gcc gtc I R D G I I K A G P A V A V gtc ggc caa gcc acc caa atc gcc aag aag gac gag ctc **tga** V G Q A T Q I A K K D E L **taat**	130
	Cec A	aag tgg aag ctt ttc aag aag atc gag aag gtc ggc caa aac K W K L F K K I E K V G Q N atc cgc gac ggc atc atc aag gcc ggc cca gcc gtc gcc gtc I R D G I I K A G P A V A V gtc ggc caa gcc acc caa atc gcc aag **tga taat** V G Q A T Q I A K	118

Table 4.10 (continued)

Promoter	Peptide	Peptide sequence	bp
GluB-4	BP178	aag aag ctc ttc aag aag atc ctc aag tac ctc gcc ggc cca K K L F K K I L K Y L A G P gcc ggc atc ggc aag ttc ctc cac tcc gcc aag aag gac gag A G I G K F L H S A K K D E ctc **tag taat** L	94
	BP188	aag aag ctt ttc aag aag atc ctc aag tac ctc gcc gtc gcc K K L F K K I L K Y L A V A gtc gtc ggc caa gcc acc caa atc gcc aag aag gac gag ctc V V G Q A T Q I A K K D E L **tag taat**	91
	BP192	aag aag ctc ttc aag aag atc ctc aag tac ctc gcc ggc cca K K L F K K I L K Y L A G P gcc aag aag ctc ttc aag aag atc ctc aag tac ctc aag gac A K K L F K K I L K Y L K D gag ctc **tag taat** E L	97
	Cec A_KDEL	aag tgg aag ctt ttc aag aag atc gag aag gtc ggc caa aac K W K L F K K I E K V G Q N atc cgc gac ggc atc atc aag gcc ggc cca gcc gtc gcc gtc I R D G I I K A G P A V A V gtc ggc caa gcc acc caa atc gcc aag aag gac gag ctc **tga** V G Q A T Q I A K K D E L **taat**	130
	Cec A	aag tgg aag ctt ttc aag aag atc gag aag gtc ggc caa aac K W K L F K K I E K V G Q N atc cgc gac ggc atc atc aag gcc ggc cca gcc gtc gcc gtc I R D G I I K A G P A V A V gtc ggc caa gcc acc caa atc gcc aag **tga taat** V G Q A T Q I A K	118
Glb-1	BP178	aag aag ctc ttc aag aag atc ctc aag tac ctc gcc ggc cca K K L F K K I L K Y L A G P gcc ggc atc ggc aag ttc ctc cac tcc gcc aag aag gac gag A G I G K F L H S A K K D E ctc **tag taat** L	94
	BP192	aag aag ctc ttc aag aag atc ctc aag tac ctc gcc ggc cca K K L F K K I L K Y L A G P gcc aag aag ctc ttc aag aag atc ctc aag tac ctc aag gac A K K L F K K I L K Y L K D gag ctc **tag taat** E L	97

Single nucleotide changes were introduced in peptide coding sequences to create specific restriction sites. Restriction sites introduced for cloning purposes are underlined; nucleotides in red indicate the modifications made in the peptide coding sequence to introduce restriction sites; sequences in bold indicate the stop codons (**tga, taa**). Amino acid sequence: in green, BP134; in blue, (1-10) magainin; in pale brown, (25-37) cecropin A; in yellow, cecropin A; in orange, AGPA hinge; in purple, ER retention signal and in black, the last two amino acids of GluB-4 signal peptide that were added during subsequent cloning.

Table 4.11 Nucleotide sequence encoding the 18 kDa oleosin and 18 kDa oleosin_Cec A fusion genes

Promoter	Fusion protein	Peptide sequence	bp
Ole 18	Ole18-Cec A	cgttgcattcatcttgctagctagcttagca **atg** gcg gat cgc gac cgc M A D R D R gcc ggg cag tac tac cag cag cag aga ggg cag gtg ggg gag A G Q Y Y Q Q Q R G Q V G E acg gtg aag ggg atc ctg ccg gag aag gcg ccg tcg gcg tcg T V K G I L P E K A P S A S cag gcg ctg acg gtg gcg acg ctg ttc ccg ctg ggt ggg ctg Q A L T V A T L F P L G G L ctg ctc gtg ctg tcc ggg ctg gcg ctg gcg gcg tcc gtg gtg L L V L S G L A L A A S V V ggg ctc gcc gtc gcc acg ccg gtg ttc ctg atc ttc agc ccg G L A V A T P V F L I F S P gtg ctc gtc ccg gcc gcg ctg ctc atc ggg ctc gcc gtc gcc V L V P A A L L I G L A V A ggc ttc ctc acc tcc ggc gcg ctg ggc ctc ggc ggg ctg tcg G F L T S G A L G L G G L S tcg ctc acc ttc ctc gcc aac acg gcg cgc cag gcg ttc cag S L T F L A N T A R Q A F Q cgc acg ccc gac tac gtc gag cag gcg cgg cgc agg atg gcc R T P D Y V E Q A R R R M A gag gcc gcc gcg cac gcc ggc cac aag acg gcg cag gcc ggc E A A A H A G H K T A Q A G cac gcc atc cag ggc agg gcc gac cag gcc ggc acc ggc gcc H A I Q G R A D Q A G T G A ggc gcc ggc ggt ggc gcc ggc acc aag aca tcc tcg ccc acc G A G G G A G T K T S S P T acc gag aac ctc tac ttc cag agc aag tgg aag ctt ttc aag T E N L Y F Q S K W K L F K aag atc gag aag gtc ggc caa aac atc cgc gac ggc atc atc K I E K V G Q N I R D G I I aag gcc ggc cca gcc gtc gcc gtc gtc ggc caa gcc acc caa K A G P A V A V V G Q A T Q atc gcc aag **tga taat** I A K	695
	Ole18	cgttgcattcatcttgctagctagcttagca **atg** gcg gat cgc gac cgc M A D R D R gcc ggg cag tac tac cag cag cag aga ggg cag gtg ggg gag A G Q Y Y Q Q Q R G Q V G E acg gtg aag ggg atc ctg ccg gag aag gcg ccg tcg gcg tcg T V K G I L P E K A P S A S cag gcg ctg acg gtg gcg acg ctg ttc ccg ctg ggt ggg ctg Q A L T V A T L F P L G G L ctg ctc gtg ctg tcc ggg ctg gcg ctg gcg gcg tcc gtg gtg L L V L S G L A L A A S V V ggg ctc gcc gtc gcc acg ccg gtg ttc ctg atc ttc agc ccg G L A V A T P V F L I F S P gtg ctc gtc ccg gcc gcg ctg ctc atc ggg ctc gcc gtc gcc V L V P A A L L I G L A V A ggc ttc ctc acc tcc ggc gcg ctg ggc ctc ggc ggg ctg tcg G F L T S G A L G L G G L S tcg ctc acc ttc ctc gcc aac acg gcg cgc cag gcg ttc cag S L T F L A N T A R Q A F Q cgc acg ccc gac tac gtc gag cag gcg cgg cgc agg atg gcc R T P D Y V E Q A R R R M A gag gcc gcc gcg cac gcc ggc cac aag acg gcg cag gcc ggc E A A A H A G H K T A Q A G cac gcc atc cag ggc agg gcc gac cag gcc ggc acc ggc gcc H A I Q G R A D Q A G T G A ggc gcc ggc ggt ggc gcc ggc acc aag aca tcc tcg **tga** G A G G G A G T K T S S **taat**	556

Restriction sites used for cloning purpose are underlined; sequences in bold indicate the start codon (**atg**) and the stop codons (**tga, taa**). Amino acid sequence: in blue 18 kDa rice oleosin; in purple TEV protease recognition site; in yellow cecropin A.

The synthetic gene sequences were then amplified by PCR using specific primers containing the appropriate restriction sites for subsequent cloning steps **(Fig. 4.4 A, B)**. When needed, the signal peptide of the seed storage protein, GluB-1, GluB-4 or Glb-1, was fused to the N-terminal end of the AMP sequence. For this, the nucleotide sequence encoding the signal peptide was PCR-amplified from rice (*O. sativa* cv. Senia) genomic DNA together with the promoter DNA sequence of the corresponding gene (*GluB-1*, *GluB-4* or *Glb-1*). Next, the DNA fragment encompassing the promoter and the signal peptide sequence was fused to the synthetic *AMP* gene of interest. All the PCR products were gel-purified using a DNA purification column **(Fig 4.4 C)** and cloned into the pGEM ®T-Easy plasmid.

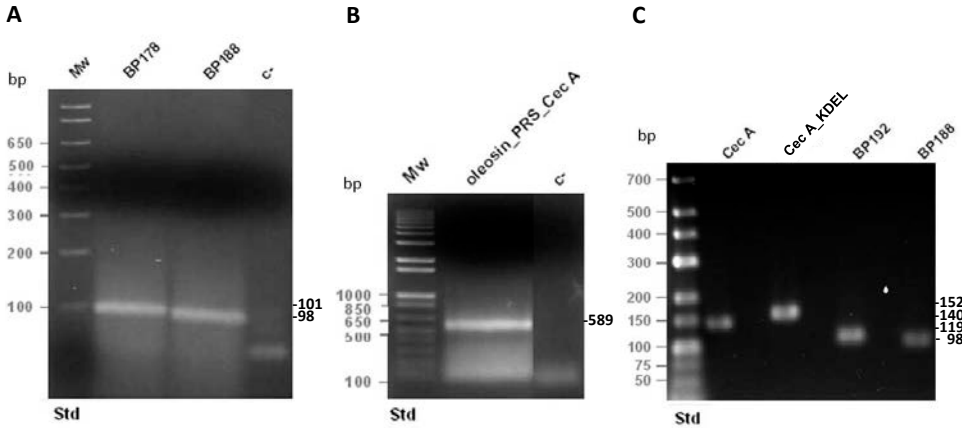


Fig. 4.4 Analysis of PCR products on agarose gel stained with ethidium bromide. (A) Product amplification corresponding to *BP178_KDEL* (101 bp) and *BP188_KDEL* (98 bp) peptide genes designed to be expressed under the control of *GluB-1* endosperm promoter. (B) PCR products corresponding to *oleosin_PRS_Cec A* fusion gene (589 bp) designed to be expressed under the control of *Ole18* embryo promoter. (C) Purified DNAs encoding *Cec A* (140 bp), *Cec A_KDEL* (152 bp), *BP192_KDEL* (119 pb) and *BP188_KDEL* (98 pb). PRS, TEV protease recognition sequence; bp, base pairs. The PCR products here presented contain restriction sites added for cloning purposes, as well as some extra nucleotides from the corresponding signal peptide sequence. Std, molecular standard (1 Kb, Invitrogen) plus. Numbers in the left side indicate the size of the DNA fragments of the molecular weight DNA markers.

The insertion of the DNA fragment of interest into the plasmid was verified by colony PCR (using primers to amplify the insert), followed by restriction enzyme analysis.

After screening for recombinant plasmids, those colonies harbouring the plasmid vector containing the cloned fragment were isolated and the recombinant plasmids were digested using the appropriate restriction enzyme (for which restriction sites were inserted by PCR) to release the cloned insert.

The complete nucleotide sequence of the cloned genes was confirmed by DNA sequence analysis performed on an automated sequencer ABI PRISM 377 (Applied Biosystems) at the DNA Sequencing Facility from CRAG (Barcelona, Spain). In this way, sequences encoding the various *BP134* and *Cec A* gene derivatives, as well as the oleosin and oleosin fusion sequences were obtained.

4.2.3 Cloning of seed-specific rice promoters and N-terminal signal peptide sequences

Four rice promoters that are known to be active in particular tissues of the rice seed were isolated by PCR using rice genomic DNA as the template (*Oryza sativa* ssp*. japonica* cv. Senia, and *Oryza sativa* ssp. *indica* cv. IR36). The oligonucleotide primer sequences were designed to amplify a DNA fragment encompassing the promoter, the 5'-untranslated region (UTR) and the coding region for the N-terminal signal peptide of the corresponding seed storage protein genes. Oligonucleotide primers were designed according to the nucleotide sequences available at the EMBL/GenBank/DDBJ nucleotide sequence database with accession numbers AY427569 (*GluB-1*), AY427571 (*GluB-4*), AY427575 (*Glb-1*) and AY427563 (Ole18) (Qu et al., 2004). The signal peptide encoding-region of the *GluB-1*, *GluB-4* and *Glb-1* genes is described in **Table 4.12** (the 18 kDa oleosin does not contain N-terminal signal peptide sequence).

Table 4.12 Nucleotide sequences of the N-terminal signal peptide region of the *GluB-1*, *GluB-4* and *Glb-1* genes from *Oryza sativa*

Promoter	Signal peptide sequence	#bp
GluB-1	**atg** gcg agt tcc gtt ttc tct cgg ttt tct ata tac ttt tgt M A S S V F S R F S I Y F C gtt ctt cta tta tgc cat ggt tct atg gcc aag aagctt gagctc ggatcc V L L L C H G S M A K K L	72
GluB-4	**atg** gcg acc ata gct ttc tct cgg tta tct atc tac ttt tgt M A T I A F S R L S I Y F C gtt ctt ctc cta tgc cat gga tcc gagctc atg gcc V L L L C H G S M A	72
Glb-1	**atg** gct agc aag gtc gtc ttc ttc gcg gcg gcg ctc atg gcg M A S K V V F F A A A L M A gcc atg gtg gcc atc tcc ggc gcc ggatcc A M V A I S G A	66

Restriction sites introduced for cloning purpose are underlined; nucleotides in red indicate the modifications made in the signal peptide coding sequence to introduce restriction sites; sequences in bold indicate the start codon (**atg**). Amino acid sequence: in black, signal peptide; in green, amino acids from BP134 derivatives; in grey, the last two amino acids of the GluB-4 signal peptide which were incorporated during cloning steps. #bp indicates the number of pair bases of the DNA fragment encoding the signal peptide.

The *GluB-1*, *GluB-4* and *Glb-1* promoters were isolated from genomic DNA from *Oryza sativa* ssp. *japonica* cv. Senia, while genomic DNA from *Oryza sativa* ssp. *indica* cv. IR36 was used to obtain the *Ole18* gene promoter **(Fig. 4.5)**. Attempts were made to isolate the *Ole18* promoter from the *japonica* ssp., with no success.

The DNA fragments containing the promoter and N-terminal signal peptide region (included in endosperm promoters) were as follows: (1) *GluB-1* promoter (2292 bp) containing the 5'UTR region (43 bp) and the N-terminal signal peptide sequence (72 bp), (2) *GluB-4* promoter (1448 bp) containing the 5'UTR region (26 bp) and the signal peptide sequence (72 bp), (3) *Glb-1* promoter (902 bp) containing the 5'UTR region (73 bp) and the N-terminal signal peptide sequence (66 bp), and (4) *18 kDa oleosin* promoter (1139 bp) containing the 5'UTR region (61 bp). A more detailed information on rice seed promoters is compiled in **Fig 4.6**.

The PCR-amplified DNA fragments were cloned into a pGEM ®T-Easy vector (*Eco*R I restriction enzyme site) and the ampicillin-resistant clones were screened by PCR and restriction digestion analysis. Cloning of the correct DNA fragments was confirmed by DNA sequencing. The nucleotide sequences of the isolated rice promoters containing the signal peptide

coding region are shown in **Annex 8.1**. Nucleotide changes were observed in the *GluB-1*, *Glb-1* and *Ole18* promoter sequences compared to the DNA sequences deposited in GenBank sequences which might be explained by differences among rice varieties (Nipponbare variety, GenBank database; Senia and IR36, present work).

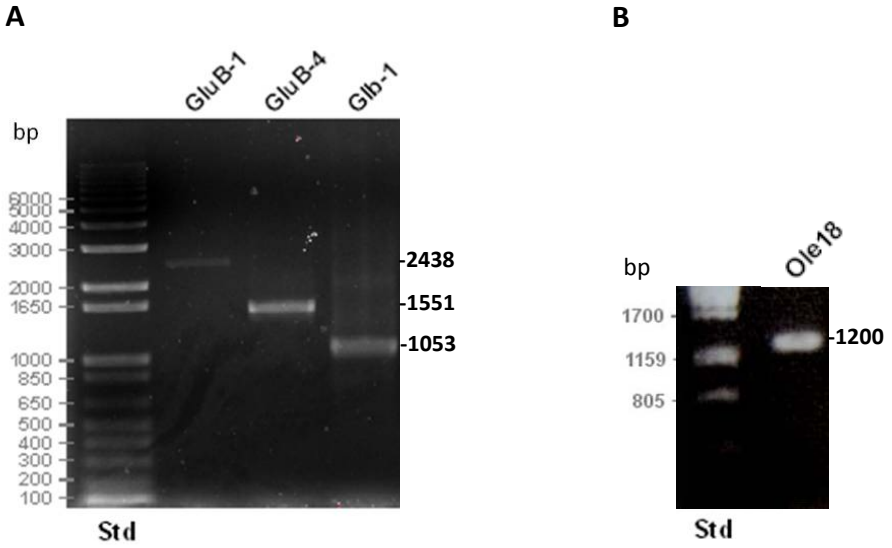


Fig. 4.5 Analysis of the PCR products for the obtention of the *GluB-1* (2438 bp), *GluB-4* (1551 bp), *Glb-1* (1053 bp) and *Ole18* (1200 bp) promoters. (A) PCR products of the endosperm specific promoters including the N-terminal signal peptide coding region. (B) DNA fragment corresponding to the *18 kDa oleosin* promoter. Ethidium bromide-stained agarose gel (1%). Std, molecular standard 1 Kb plus (Invitrogen). Numbers in the left side indicate the size of the DNA fragments of the molecular weight DNA markers. Bp, base pairs.

4.2.3.1 Promoter analysis

It is generally accepted that the seed specificity of storage protein gene expression is primarily regulated at the transcriptional level, although post-transcriptional processes can modulate the final amount of gene products. *Cis*-regulatory sequences, or *cis*-regulatory elements are recognized by distinct transcription factors and/or other regulatory molecules that are needed to activate and sustain transcription. Takaiwa and co-workers (1996) described that the AACA and GCN4 motifs are required for the high level of seed-specific expression of the glutelin multigene family in rice. Specifically, the combination of GCN4, AACA and ACGT motifs is shown to be sufficient to

confer detectable endosperm expression of the *GluB-1* gene (Washida et al., 1999; Wu et al., 2000). Furthermore, Wu, L. S. and co-workers (1998) found three conserved sequences (ATAGTTTAG, ATG/A CGGCC and CTCCC…$_{10bp}$...CACCC **(Fig. 4.6 D)** and two potential regulatory elements of cereal storage proteins (RY repeat and CATC tetranucleotide) in the flanking 5' region of *16* and *18 kDa* rice *oleosin* genes. Other common *cis*-regulatory elements involved in seed-specific regulation/expression are described in **Table 4.13**.

Table 4.13 Relevant plant *cis*-regulatory elements involved in seed specific expression in different plant species

Motif	Reported in	Sequence	Function
RY-element	*Helianthus annuus*	CATGCATG	Involved in seed-specific regulation
Skn-I_motif	*Oryza sativa*	GTCAT	Required for endosperm expression
CAATBOX1	*Pisum sativum*	CAAT	Responsible for the tissue specific promoter activity
PROLAMINBOX	*Oryza sativa*	TGCAAAG	Involved in quantitative regulation of *GluB-1* gene
CANBNNAPA	*Brassica napus*	CNAACAC	Required for seed (embryo and endosperm) specific transcription
PROBXBBNNAPA	*Brassica napus*	CAAACACC	Required for seed specific gene expression and ABA responsiveness
ACGTSEED3	*Tobacco*	GTACGTGGCG	Related to seed expression
ACGTOSGLUB1	*Oryza sativa*	GTACGTG	Required for endosperm expression of glutelin genes
AACACOREOSGLUB1	*Oryza sativa*	AACAAAC	Involved in controlling endosperm expression of glutelin genes
CEREGLUBOX2PSLEGA	*Pisum sativum*	TGAAAACT	Sequence homologous to the cereal glutenin gene control element
GCN4	*Oryza sativa*	CAAGCCA/TGA GTCA	Involved in endosperm expression

(Higo et al., 1999 (PLACE database); Lescot et al., 2002 (PlantCare database)).

In the present work, a search for the presence of known *cis*-regulatory elements in the sequence of the isolated rice promoters was carried out using the PLACE and PlantCare databases (Higo et al., 1999 (PLACE database); Lescot et al., 2002 (PlantCare database). Interestingly, in addition

to seed-specific regulation, *cis*-regulatory elements associated with plant hormonal regulation, light and stress responsiveness were identified **(Fig. 4.6).**

None of the changes found in the promoter regions of the *GluB-1* and *Glb-1* promoters (e. g. single nucleotide replacements, 1-4 nucleotide deletions/insertions) located at any of the known *cis*-elements controlling seed-specific gene expression **(Fig. 4.6 A, C)**.

To note, major nucleotide changes were identified in the sequence of the isolated *Ole18* promoter in comparison to the AY427563 sequence deposited in GenBank. In fact, the *Ole18* promoter isolated from genomic DNA of *indica* was 1200 nucleotides in length, while AY427563 (isolated from *japonica*, GenBank) comprises 1249 nucleotides. Nucleotide sequencing also revealed a 64-nucleotide fragment deletion and a 13 nucleotide fragment insertion. An additional GCN4 motif was identified in the nucleotide sequence of the isolated promoter (*indica* rice). Interestingly, the GCN4 motif was also found in the *GluB-1*, *GluB-4* and *Glb-1* promoter genes, but not in the *Ole18* AY427563 promoter sequence **(Fig. 4.6 D).**

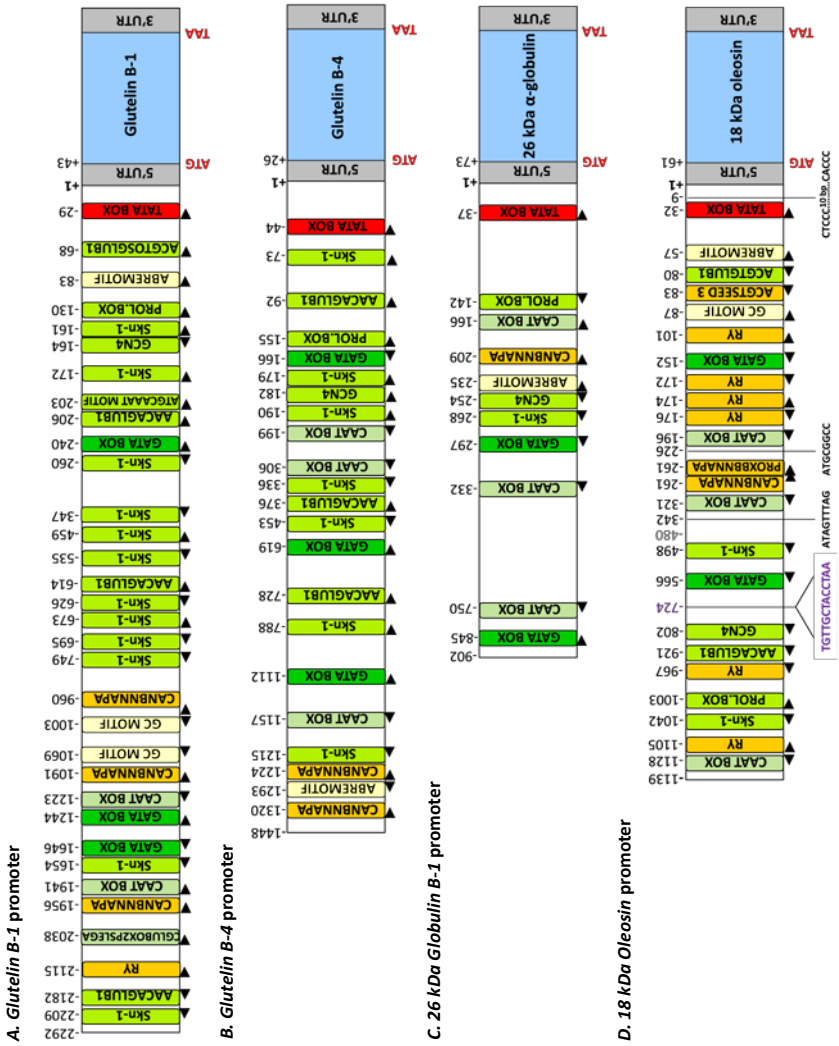


Fig. 4.6 Schematic diagram of *cis*-elements identified in the rice seed specific promoters from *Oryza sativa* ssp. *indica* cv. IR36 (for *18 kDa oleosin* promoter) or *Oryza sativa* ssp. *japonica* cv. Senia (for *GluB-1*, *GluB-4* and *Glb*-1 promoters). The transcriptional start site is indicated as **+1**. Nucleotides in red indicate the translational initiation (ATG) and termination codon (TAA). Negative numbers represent the position of nucleotides upstream the transcription start site (+1). Blue box denotes the seed protein coding region. The uncoloured area represents the promoter gene sequence. The location of the *cis*-acting regulatory elements identified in seed specific promoters are indicated as follows: *cis*-acting regulatory element required for tissue specific expression; *cis*-acting regulatory element required for endosperm expression; *cis*-acting regulatory element involved in seed specific regulation/expression; required for high level expression; other identified motifs; TATA box, the main recognition site for the binding of RNA polymerase in the promoter sequence. Arrowheads at the bottom indicate the direct (→) or reverse orientation (←) of the *cis*-acting element. In *18 kDa oleosin* promoter, the nucleotide sequence in bold indicate the three conserved sequences identified in promoter region of the rice *16* and *18 kDa* rice *oleosin* genes (Wu, L. S. et al., 1998), nucleotides in purple indicate a 13 nucleotide fragment insertion (TGTTGCTACCTAA) and position -480 indicate a 64-fragment deletion (CATTATAACTAGGTATTATGAGGTACCAAATTTACAATAGAAAAAATAGTACTTCATGGTACTT) compared to AY427563 (isolated from *japonica)*.

4.2.4 Cloning of the *nopaline synthase* terminator

The DNA sequence encoding the *nopaline synthase* terminator (*Nos-T*) from *A. tumefaciens* was isolated by PCR using the plasmid *pCUbi::Cec A_KDEL::Nos-T* (Coca et al., 2006) as the template **(Fig. 4.7)**. Plasmidic DNA was provided by Sonia Campo (CRAG, Barcelona, Spain). The amplified fragment was cloned into pGEM®-T Easy vector and its nucleotide sequence confirmed by DNA sequencing.

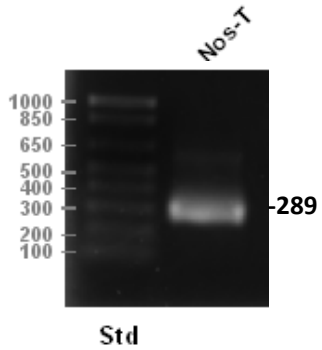


Fig. 4.7 Analysis of PCR products for the amplification of the *Nos* terminator (289 bp) by agarose gel stained with ethidium bromide. Std, molecular standard 1 Kb plus (Invitrogen). Std, molecular standard (1 Kb, Invitrogen) plus. Numbers in the left side indicate the size of the DNA fragments of the molecular weight DNA markers.

4.2.5 Construction of vectors for rice transformation

DNA fragments consisting of an endosperm-specific promoter, the signal peptide coding region, the coding region of the synthetic *AMP* gene and the *Nos* terminator sequence, were cloned into the pCAMBIA1300 plant expression vector. For this purpose, the *pGEM::Nos-T* and *pGEM:endosperm promoter-signal peptide* (*SP*) vectors **(Fig. 4.8)** were digested with the appropriate restriction enzymes. The *Nos* terminator sequence was cloned downstream of the *promoter-SP*, resulting in plasmids *pGEM::endosperm promoter-SP::Nos_T*. Then, the synthetic *AMP* gene was cloned into the *pGEM::endosperm promoter-SP::Nos-T* plasmid DNA, flanked by the promoter (containing the signal peptide coding region) and the terminator sequence. The complete cassette (*endosperm promoter-SP::AMP::Nos-T*) was cloned into the pCAMBIA1300 vector. All the constructs used for rice transformation were verified by nucleotide sequencing. The schematic

representations of the binary vectors are shown in **Figs. 4.8** and **4.9**. The expression vectors were transferred to *Agrobacterium tumefaciens* (EAH105 strain) (Hood et al., 1993). The parent pCAMBIA1300 vector already contains the *hptII* (hygromycin phosphotransferase) gene encoding hygromycin resistance in the T-DNA region.

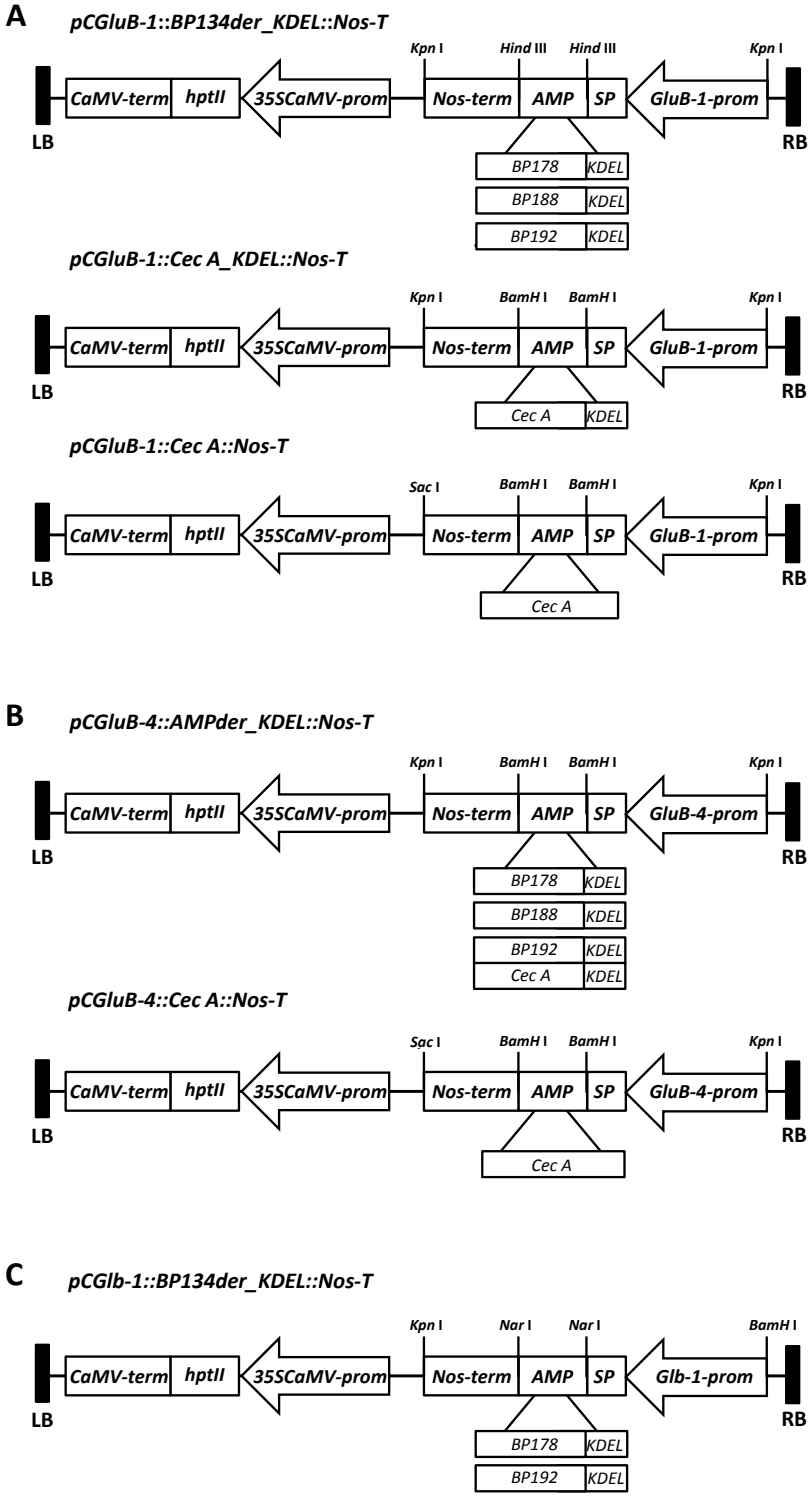


Fig. 4.8 Schematic representation of the plant expression vectors prepared in this work for the expression of *AMP* genes in the rice endosperm. Each synthetic peptide was cloned between an

endosperm-specific promoter (containing the signal peptide (SP) coding sequence for the corresponding gene and the *nopaline synthase* terminator (*Nos-T*). *35SCaMV-prom*, cauliflower mosaic virus 35S promoter; *CaMV-term*, 35S terminator from cauliflower mosaic virus; *hptII,* gene encoding resistance to hygromycin; LB, left border; RB, right border. Restriction sites used for cloning purposes are indicated.

Moreover, in the present work, the *Cec A* synthetic gene was expressed in rice as a *18 kDa oleosin_Cec A* fusion gene under the control of *18 kDa oleosin* promoter and the *Nos* terminator. The nucleotide sequence encoding the TEV protease recognition site (PRS) was placed between the *oleosin* and *Cec A* DNA sequences. The components of the T-DNA in the *pCOle18::oleosin_PRS_Cec A::Nos-T* construct are indicated in **Fig 4.9 A**. A construct consisting of the 18 kDa oleosin coding sequence under the control of its own promoter (*Ole18* promoter) and the *Nos* terminator was also obtained and used as a control of *oleosin* overexpression in transgenic rice **(Fig 4.9 B)**.

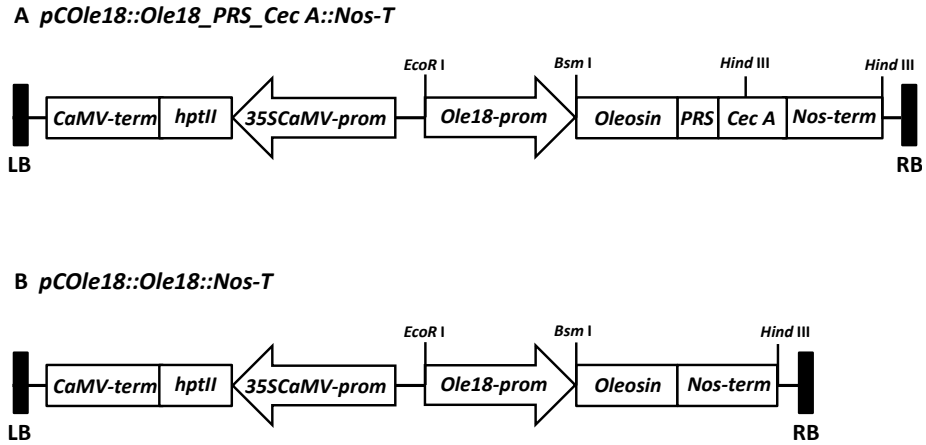


Fig. 4.9 Schematic representation of the plant expression vectors prepared in this work for the expression of the *oleosin_Cec A* fusion gene (A) and the *oleosin* gene (B) in the rice embryo/aleurone. Each transgene was cloned between the *18 kDa oleosin* promoter and the *nopaline synthase* terminator (*Nos-T*). *35SCaMV-prom*, cauliflower mosaic virus 35S promoter; *CaMV-term*, 35S terminator from cauliflower mosaic virus; *hptII,* gene encoding resistance to hygromycin; LB, left border; RB, right border. Restriction sites used for cloning purposes are indicated.

In summary, fourteen plant expression vectors containing a seed-specific promoter (with the N-terminal peptide sequence), the coding sequence of a synthetic *AMP* gene and the *Nos* terminator sequence were obtained and used for rice transformation **(Table 4.14)**.

Table 4.14 Plant expression vectors prepared in this work for transgenic expression of *AMPs* in the rice seed

Promoter	*AMP* gene	N-terminal signal sequence	Features	Expected tissue of expression
GluB-1	*BP178*	+	ER retention signal	Endosperm
	BP188	+	ER retention signal	Endosperm
	BP192	+	ER retention signal	Endosperm
	Cec A_KDEL	+	ER retention signal	Endosperm
	Cec A	+	-	Endosperm
GluB-4	*BP178*	+	ER retention signal	Endosperm
	BP188	+	ER retention signal	Endosperm
	BP192	+	ER retention signal	Endosperm
	Cec A_KDEL	+	ER retention signal	Endosperm
	Cec A	+	-	Endosperm
Glb-1	*BP178*	+	ER retention signal	Endosperm
	BP192	+	ER retention signal	Endosperm
Ole18	*18 kDa oleosin_PRS_Cec A*	-	Oleosin_AMP fusion	Embryo
	18 kDa oleosin	-	-	Embryo

4.3 Production and characterization of transgenic rice plants expressing an *AMP* gene in seed tissues

In this work, the *japonica* rice (*O. sativa*) cultivars Senia and Ariete were used for transformation. Transgenic rice plants were produced by *Agrobacterium*-mediated transformation, using the hygromycin resistance gene as the selectable marker (Sallaud et al., 2003). The various plant transformation vectors indicated in **Section 4.2.5 (Table 4.14)** were used. Rice transformation and phenotypic characterization was done in collaboration with B. San Segundo and M. Coca (CRAG-Barcelona) within the framework of the European Trilateral Plant KBBE project SEPSAPE (Safe and Efficient Plant Systems for Antimicrobial PEptide production). Thus, rice transformation was performed at the plant tissue culture facilities of CIDSAV (University of Girona, Girona, Spain), CRAG (Autonomous University of Barcelona, Barcelona, Spain) as well as at The Rice Functional Genomics platform (REFUGE) at Montpellier, http://www.refuge-platform.org/. In brief, embryogenic rice callus derived from the scutellum of mature seeds of japonica cv. Senia and Ariete **(Figs. 4.10 and 4.11 A)** were co-cultivated with *A. tumefaciens* (EHA105 strain) bearing a pCAMBIA1300-*seed promoter::AMP::Nos-T* binary vector. As a control, the pCAMBIA 1300 empty vector (including *hptII* selection gene and no *AMP* expression cassette) was used for rice transformation. The Senia cv. was mainly used as starting material to obtain the transgenic plants at the CIDSAV, while the Ariete cv. was used at the CRAG. Thirty embryogenic nodular units per construct were co-cultured. Seventy-two hours after transferring to R2C agar medium, the co-cultured callus pieces were transferred onto R2S agar first selection medium. Contaminated calli, resulting from an overgrowth of *A. tumefaciens* were discarded. About 14 days after transferring to selective medium, the primary calli become necrotic, while the first hygromycin resistant, white nodular units appear **(Fig. 4.11 B)**. Uncontaminated calli were placed onto NBS agar selection medium. The transformed nodules arising from primary necrotic callus were scattered around the medium surface **(Fig. 4.11 C)**.

4.3.1 Recovery of transformed events and plant regeneration

Independently transformed rice callus were selected by hygromycin resistance according to Sallaud et al., 2003. Seven to ten transformed calli (selected on the basis of hygromycin resistance) were transferred onto PRAG agar pre-regeneration medium and then to RN agar regeneration **(Fig. 4.11 D-F)**. The resistant calli (each one arising from a single co-cultured callus) were transferred and identified by the code number of the primary callus. The transformation events were identified by using the RX.Y code where X refers to the construct used for rice transformation (type of transgene) and Y indicates the primary callus from which each line derives. After 3-4 weeks, the hygromycin-resistant calli already differentiated into shoots and were transferred to test tubes containing P agar **(Fig. 4.11 G)**. After 3 weeks, plantlets presenting vigorous shoots and roots were transferred to the greenhouse for acclimatation **(Fig. 4.11 H)**. In this step a third number was added to the previous code (for instance, the plantlet RX 7.1 derives from the single co-cultured callus number 7, and from the secondary callus number 1).

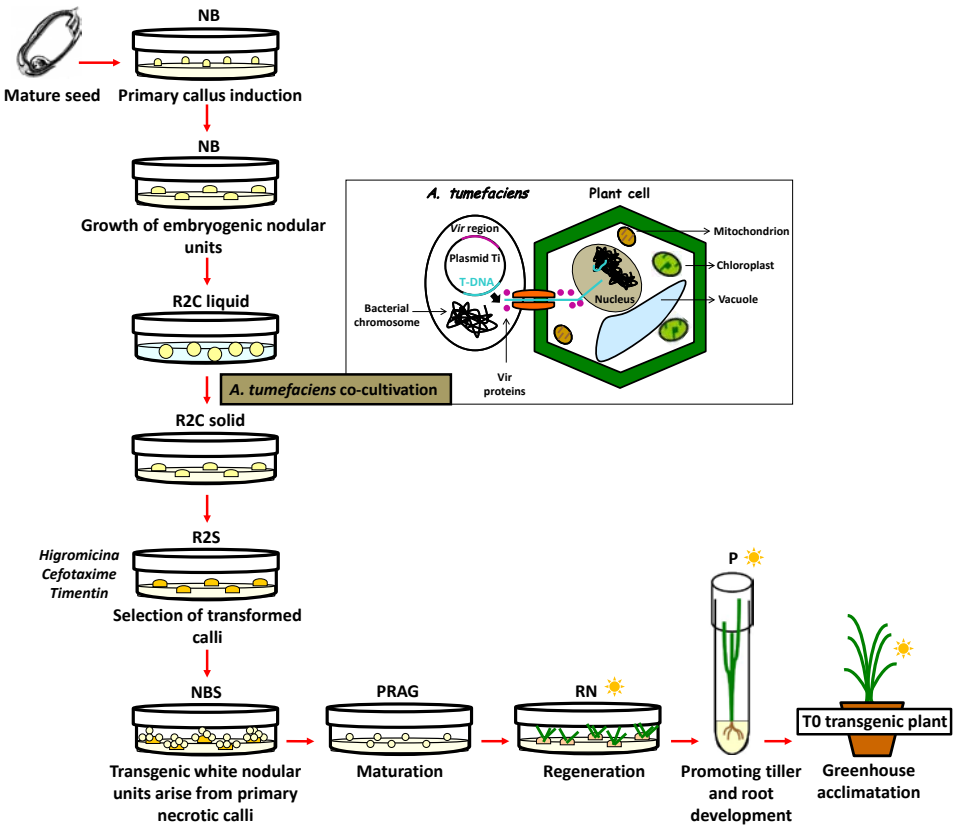


Fig. 4.10 Schematic diagram of *A. tumefaciens*-mediated transformation of *japonica* rice. Embryogenic calli were co-cultured with *A. tumefaciens* (EHA105 strain) carrying a binary vector for expression of an *AMPder* under the control of a rice seed-specific promoter. 50 mg/mL of hygromycin B was used for selection of transformed cells. ☀ Indicates 12/12 h (day/night) photoperiod under a photon flux of 110-150 µmol. m^{-2}. s^{-1}. The preceding steps are performed under dark conditions. The whole transformation process lasts approximately 17 weeks.

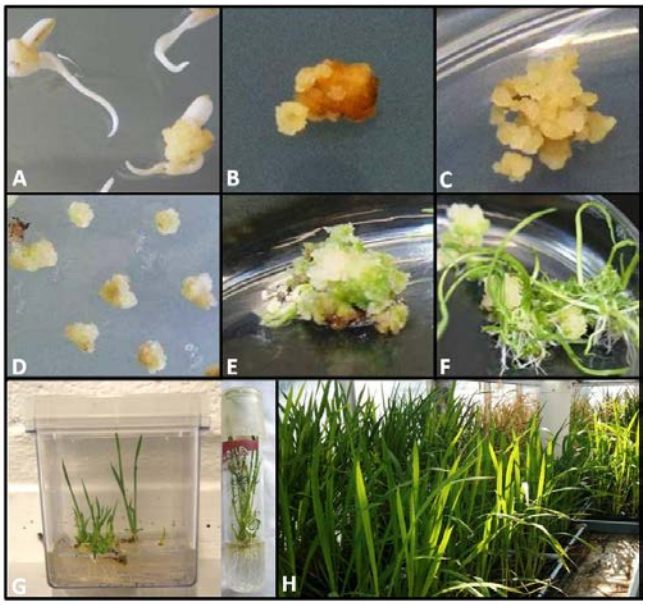


Fig. 4.11 The *Agrobacterium*-mediated transformation of rice. Embryogenic calli derived from the scutellum of mature seeds (*japonica* cv. Senia) are co-cultivated with *A. tumefaciens* EHA105 (harbouring either a pCAMBIA-derived plasmid for expression of an *AMP*, or the vector control plasmid). (A) Calli induced from mature seeds. Embryogenic nodular units produced by primary callus. (B) Transformed callus after 14 days incubation on R2S selection medium. First hygromycin-resistant transgenic white nodular units appear, while the primary callus became necrotic. (C) Development of hygromycin-resistant cell lines at the surface of the primary callus, seven days after being transferred to NBS selection medium. (D) Hygromycin-resistant cell lines maturating on the P pre-regeneration medium. (E, F) Shoots regenerating from hygromycin-resistant calli three weeks after being transferred onto RN regeneration medium. (G) 3-week-old regenerated plantlets ready for acclimation. (H) The primary transgenic plants (T0) are grown to maturity in the greenhouse.

Transformations made with the 14 different constructs prepared in this study **(Section 4.2.5, Results)** yielded hygromycin-resistant transgenic calli and also transgenic T0 plants **(Table 4.15)**. In *Agrobacterium*-mediated transformation, there is the possibility that more than one hygromycin-resistant calli arise from the same co-cultured primary callus. In this respect, Sallaud and co-workers (2003) demonstrated that transgenic calli mainly correspond to independent transformation events. On this basis, the transformation efficiency was determined in terms of T0 plants obtained per callus (starting from thirty embryogenic calli/construct), which was found to range between 20 and 103.3% under the greenhouse conditions operating at the CIDSAV facility **(Section 3.14.3, Methods)**.

Table 4.15 Production of transgenic rice plants (cultivars, Senia and Ariete)

			# plants regenerated		
Construct	**Cultivar**	**Code**	**T0 total**	**T0 fertile**	**Fertility (%)**
pC1300	Senia	EV	2	2	100.0
pCGluB-1::BP178_KDEL	Senia	S11	1	1	100.0
	Ariete	A11	2	1	50.0
pCGluB-1::BP188_KDEL	Senia	R7	6	6	100.0
	Ariete	A12	9	7	77.7
pCGluB-1::BP192_KDEL	Senia	R4	6	5	83.3
		S8	6	6	100.0
	Ariete	MC2	19	17	89.5
pCGluB-1::Cec A_KDEL	Senia	S3	2	1	50.0
	Ariete	MC3	13	9	69.2
		A3	1	1	100.0
pCGluB-1::Cec A	Ariete	A5	6	4	66.6
pCGluB-4::BP178_KDEL	Ariete	A14	6	6	100.0
pCGluB-4::BP188_KDEL	Senia	R8	31	31	100.0
	Ariete	A13	1	0	0.0
pCGluB-4::BP192_KDEL	Ariete	A7	13	11	84.6
pCGluB-4::Cec A_KDEL	Senia	S2	4	3	75.0
pCGluB-4::Cec A	Senia	R10	24	24	100.0
	Ariete	A1	9	4	44.4
pCGlb-1::BP178_KDEL	Senia	R11	25	23	92.0
pCGlb-1::BP192_KDEL	Senia	S6	4	4	100.0
pCOle18::Ole18_PRS_Cec A	Senia	R9	26	26	100.0
	Ariete	A16	16	16	100.0
pCOle18::Ole18	Ariete	MCX	n.d	n.d	n.d

n.d., not determined.

Most transgenic T0 plants were morphologically similar to the wild-type untransformed plants with respect to vegetative growth, flowering behaviour and seed yield. However, R8 (*pCGluB-4::BP188_KDEL*), R9 (*pCOle::Ole18_PRS_Cec A*) and R10 lines (*pCGluB-4::Cec A*) showed more vigorous root and shoot systems than wild-type plants, and their seed yield was higher than the wild-type (results not shown). By contrast, transgenic lines bearing the *BP188_KDEL* transgene under the control of the *GluB-1* promoter (R7) exhibited less shoot biomass and reduced seed yield **(Fig. 4.12 A)**. All independent transgenic R7 lines exhibited the same abnormal phenotype under controlled greenhouse conditions, excluding the possibility of an insertional effect caused by transgene integration into the rice genome.

Interestingly, although most of T1 seeds from primary regenerants (T0) showed wild-type-like morphology and viability, mature T1 seeds from some T0 plants exhibited different morphological traits while maintaining their viability. As an example, **Fig. 4.12 B and Table 4.16** show the morphological variability among T1 seeds from *pCOle::Ole18_PRS_Cec A* (R9 lines).

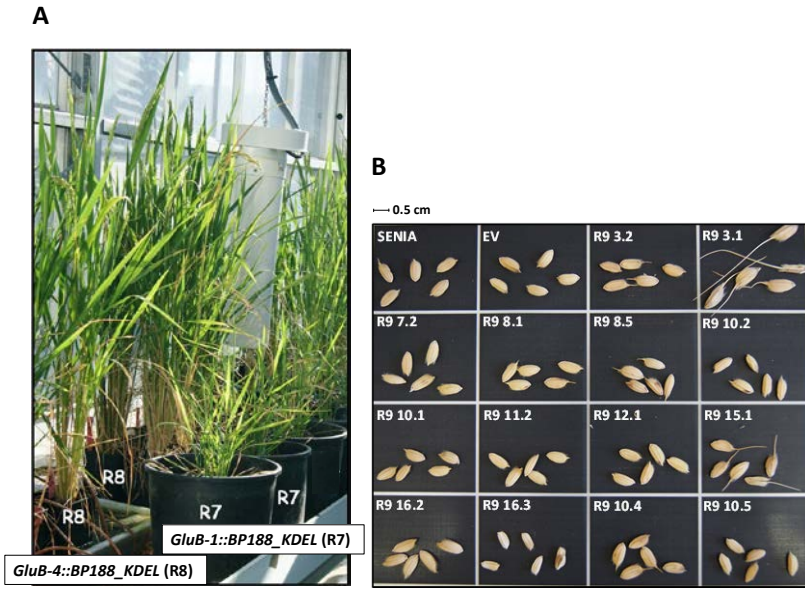


Fig. 4.12 Appearance of the transgenic rice plants designed to produce AMPder peptides. (A) Impact of BP188_KDEL on rice development. Transgenic plants carrying the *BP188_KDEL* gene driven by *GluB-1* (R7) and *GluB-4* (R8) promoter. Plants were grown in soil in the greenhouse. (B) Seed progeny from *Ole18::Ole18_PRS_Cec A* (R9) T0 transgenic rice plants. Seeds from control plants carrying the pCAMBIA1300 empty vector (EV) and wild-type (Senia). Grains are morphological different.

Table 4.16 Characteristics of T0 seeds from *Ole::Ole18_PRS_Cec A* (R9)-lines

Line	Length (mm)	Width (mm)	Awn length (mm)	Weight (mg)
Controls				
Senia	7.2 bcd	4.0 ab	0.8 a	29.1 bcd
EV	7.0 bcd	3.8 ab	0.0 a	29.1 bcd
***Ole18::Ole18_PRS_Cec A* (R9)**				
3.1	9.4 d	4.2 b	24.0 c	42.8 d
3.2	7.4 bcd	4.0 ab	2.2 ab	30.1 bcd
7.2	7.8 d	3.8 ab	0.0 a	32.0 cd
8.1	7.4 bcd	3.8 ab	1.4 ab	28.5 bcd
8.5	7.2 bcd	3.2 ab	2.2 ab	26.0 bcd
10.1	6.6 abc	3.8 ab	0.8 a	25.6 bcd
10.2	6.4 ab	3.2 ab	0.6 a	25.1 bcd
10.4	7.6 cd	3.8 ab	1.4 ab	29.0 bcd
10.5	7.4 bcd	3.2 ab	0.0 a	23.3 bc
11.2	7.0 bcd	4.0 ab	0.0 a	28.2 bcd
12.1	7.0 bcd	3.6 ab	0.0 a	28.8 bcd
15.1	7.2 bcd	3.8 ab	8.4 b	22.9 b
16.2	7.8 d	3.8 ab	0.0 a	32.1 cd
16.3	5.6 a	3.0 a	0.0 a	20.3 a

Data show the mean seed length, width, awn length, and weight form seeds from indicated independent lines from each transformation event (5 seeds per line). The confidence interval is also indicated. Letters indicate significant differences according to Tukey one-way ANOVA (α=0.05).

As most regenerated plants showed normal growth and development, toxicity of the antimicrobial peptides to the rice cells can not be expected. However, different transformation efficiency (T0 plants obtained/number of starting callus) appears to occur depending on the construct used for transformation. In addition to the type of the transgene, the promoter which drives its expression might well be responsible of the observed differences in transformation efficiencies. For instance, the R7 and R8 plants harbour the same *AMP* gene (*BP188_KDEL*) but differ in the promoter used to drive its expression (the *GluB-1* promoter in R7 lines and the *GluB-4* promoter in R8 lines). Whereas the R8 lines (*GluB-4::BP188_KDEL*) grew vigorously, the R7 lines exhibited reduced growth and abnormal phenotype. Thus, in the case of the R7 and R8 plants, the promoter appears to play a critical role in determining the impact of *BP188_KDEL* expression in rice performance. Further studies are, however, needed to understand why transgenic expression of the *BP188_KDEL* under the control of the endosperm-specific *GluB-1* promoter causes such unintended effects on rice growth.

4.3.2 Transgene detection and copy number assessment

Transgene integration into the rice genome, and integrity was examined by PCR analysis of genomic DNA obtained from young leaves (T0 plants) as the template. PCR primers were designed to amplify either the *AMP* gene or the complete cassette for expression of each *AMP* gene (that is the *promoter::AMP::Nos-T* DNA fragment). Agarose gel electrophoresis of products revealed amplification of the expected DNA fragment in each case (compared to the pCAMBIA-derived plasmid). Representative results are presented in **Fig. 4.13**. Almost all the regenerated rice lines recovered from the various transformation events had the complete cassette for *AMP* expression integrated into their genomes (as revealed by PCR analysis using primers located at the promoter and terminator regions, blue arrows in **Fig. 4.13 A**).

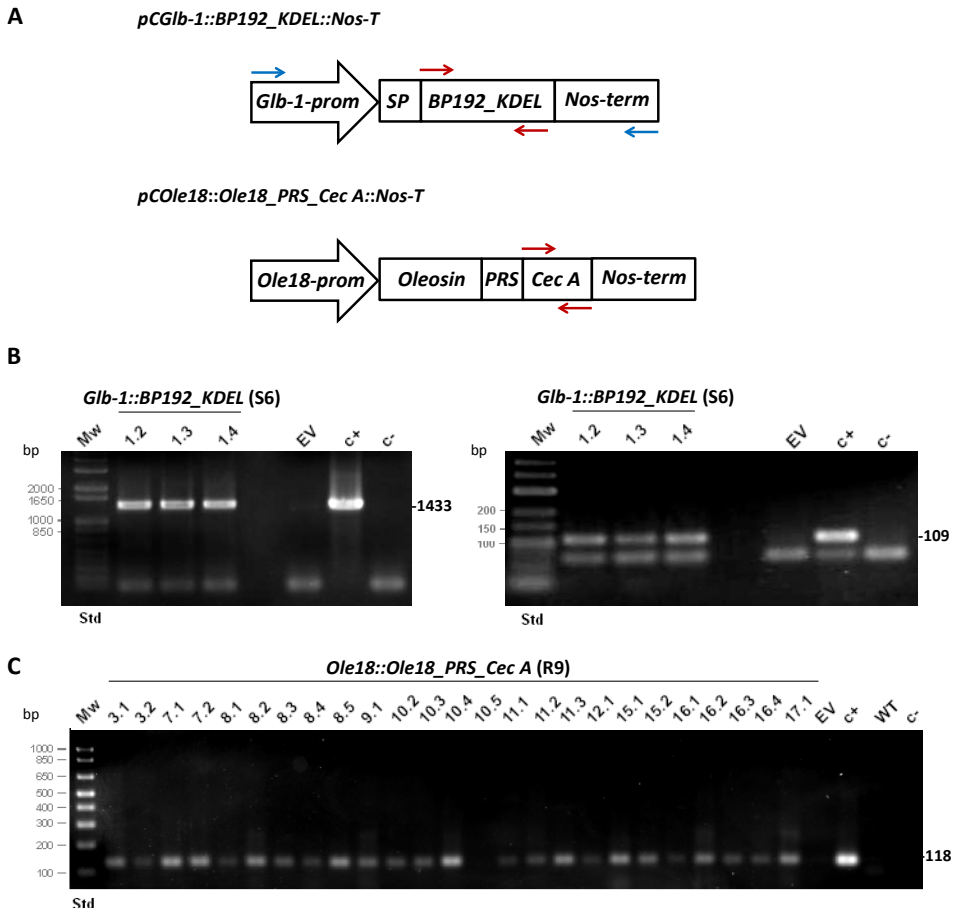


Fig. 4.13 Transgene integration in the rice genome. PCR analysis of genomic DNA. Locations of primers used for PCR amplification are denoted in (A). The PCR amplified DNA fragments were analysed by agarose gel stained with ethidium bromide. (B) Lines *pCGlb-1::BP192_KDEL* (S6 lines). Detection of the complete cassette (1433 bp, left panel; location of the PCR primers is show in A, upper panel, blue arrows), or the *AMP* gene (109 pb, right panel; primers used for PCR are shown in A, upper panel, red arrows). (C) Lines *pCOle18::Ole18_PRS_Cec A* (R9 lines). Detection of the cecropin A sequence (location of the primers is shown in A, lower panel, red arrows). c+, the pCAMBIA-derived plasmid used for the rice transformation served as a positive control of the PCR reaction; c- negative control of the PCR reaction; EV, rice plants transformed with the empty vector (pCAMBIA1300); WT, non transformed seeds. Std, molecular standard 1 Kb plus. Numbers in the left side indicate the size of the DNA fragments of the molecular weight DNA markers. Bp, base pairs.

Several studies have been shown that the insertion of multiple copies of the transgene in the plant genome may result in a loss of expression of some, if not all, integrated copies (a phenomenon referred to as transgene silencing). Moreover, when introducing DNA sequences homologous to endogenous

genes, the expression of the transgene and the endogenous gene can be silenced (co-suppression) (Ingelbrecht et al., 1994; Flavell, 1994, Stam et al., 1997; Vaucheret and Fagard, 2001). In addition to this, Vain and co-workers (2002) established a relationship between high transgene copy number and reduced seed viability.

Therefore, in order to achieve a rice seed-based platform producing foreign peptides, the estimation of transgene copy number was crucial. Southern hybridization analysis of restriction enzyme-digested genomic DNA has been the standard method to estimate transgene integrity and copy number, as well as to determine the integration patterns of the transgene in the various transformants (Southern, 1975). However, this method is laborious and time-consuming and requires large amounts of tissue for genomic DNA extraction (Ingham et al., 2001; Mason et al., 2002). Also, detection of restriction enzyme-digested DNA fragments by gel-transfer hybridization presents important limitations: (1) incomplete DNA digestion, (2) underestimation of transgene copies when more than one copy is inserted into one locus; and (3) variable blotting efficiency of the DNA fragment. To overcome these limitations an alternative method was developed to determine the transgene copy number which is based on quantitative PCR (qPCR) analysis. This method requires minimal amounts of genomic DNA and it is performed in a few hours, and it allows to estimate transgene copy number (except when rearrangements occur in the primer binding site). It also avoids post-PCR manipulation, reducing contamination risks (Ingham et al., 2001; Song et al., 2002; De Preter et al., 2002; Weng et al., 2004; Yi et al., 2008; Jiang et al., 2009). A correlation of 90-95% between Southern blot and qPCR for transgene copy number estimation has been reported (Li et al., 2004; Chu et al., 2013). Other studies indicated that qPCR is more effective and accurate than Southern blot in determining transgene copy number, being especially useful to estimate low or single copy transformation events (Yang et al., 2005; Chen et al., 2010). However, qPCR also presents some limitations as the method detects DNA fragments usually below 200 bp.

In the present study, the transgene copy number was estimated by qPCR in all T0 regenerated plants. This allowed us the identification of single copy events which were next used as the parental lines for the obtention of

homozygous lines (at the T2 generation). In this way, the high copy number transgenic plants were discarded to save efforts and space in the greenhouse.

For the identification of the low copy number events, genomic DNA was extracted from young leaves of T0 plants and analysed by qPCR targeting to either the *AMP* or the *hptII* genes. The single copy gene, *β-actin,* was used for normalization (previous studies in our research group showed that *β-actin* was a suitable endogenous control gene). qPCR was performed based on SYBR-Green technology and PCR primers were designed using the Beacon designer 7.0 software as described by Nadal et al., 2012. The *AMP* and *hptII* gene copy number was estimated by comparing quantitative PCR data of the *AMP* and the *hptII* transgene with those of *β-actin* gene, obtained from standard calibration curves. Each genomic DNA sample (including genomic DNA from Senia and EV controls) was analysed in three independent amplification reactions, performed simultaneously. Since T0 plants are hemizygous for the transgenes (*hptII* and *AMPs*) and homozygous for the *β-actin*, ratios close to 0.5 suggested that these T0 regenerated plants integrated a single copy for the *AMP* transgene (Mason et al., 2002; Weng et al., 2004; Chen et al., 2010). The percentage of the qPCR-tested T0 lines showing a single copy of both transgenes ranged from 25 to 81.4% depending on the line **(Table 4.17)**. Interestingly, some events presented a different copy number for the *hptII* and the *AMPder* transgenes, suggesting T-DNA rearrangements during the transformation process (Hiei et al., 1994; Yang et al., 2005).

It was not possible to optimize standard curves from genomic DNA of T0 plants transformed with *pCGluB-1::Cec A*, *pCGluB-4::Cec A_KDEL* and *pCGlb-1::BP192_KDEL* due to a low primer hybridization level, so transgene copy number in this plants could not be estimated.

Table 4.17 Estimation of transgene copy number (and possible rearrangements) in the T0 transgenic rice lines as determined by qPCR. The percentage (%) of the total number of lines tested showing single copy or multiple copies inserted is shown.

			Rice lines (%)		
Promoter	**Peptide/Fusion protein**	**Code**	**Single copy**	**Multiple copy**	**Rearrangements**
GluB-1	BP178_KDEL	S11, A11	50.0	50.0	0.0
	BP188_KDEL	R7, A12	25.0	75.0	0.0
	BP192_KDEL	R4, MC2, S8	50.0	10.0	40.0
	Cec A_KDEL	MC3	n.d	n.d	n.d
	Cec A	A5	33.3	66.6	0.0
GluB-4	BP178_KDEL	A14	100.0*	0.0	0.0
	BP188_KDEL	R8	59.2	25.9	14.9
	BP192_KDEL	A7	0.0	100.0	0.0
	Cec A_KDEL	S2	n.d	n.d	n.d
	Cec A	R10, A1	41.6	45.9	12.5
Glb-1	BP178_KDEL	R11	52.0	24.0	24.0
	BP192_KDEL	S6	n.d	n.d	n.d
Ole18	18 kDa oleosin_PRS_Cec A	R9	81.4	14.8	3.8

n.d, not determined; *, only one transgenic event was obtained that showed a single *AMP* and *hptII* integration.

Most of the transgenic plants showing high copy numbers, or possible rearrangements, were discarded to prevent transgene loss or silencing. Those transgenic lines containing a single copy of the transgene were selected for further characterization.

During the course of this work, qPCR analysis was also performed on selected T3 homozygous lines. In most cases, the number of copies of the transgene correlated well between T0 and T3 plants, indicating stable transgene inheritance through successive generations of the transformed rice plants (at least up to the T3 generation).

4.3.3 Transgene expression in immature seeds

Several studies demonstrated no accumulation of the transgene product in plants expressing transgenes encoding small peptides (<50 aa) (Matoba et al., 2001; Yasuda et al., 2005; Takagi et al., 2008, 2010). Failures to detect accumulation of small peptides in plants with integrated transgenes could

be, however, attributable to problems that might occur at one or another level of transgene expression (transcriptional, post-transcriptional, translational and post-translational mechanisms, including peptide degradation by host proteases).

To verify that the *AMPs* genes were properly expressed in the transgenic rice lines here obtained, a random set of seeds were subjected to RT-PCR analysis using primers specific to the coding region of each transgene, that is the *SP-AMPder* (signal peptide-*AMPder*), in the PCR reaction. For this purpose, total RNA was extracted from developing seeds (17-21 days after pollination, DAP) of T1 and T2-T3 homozygous plants, and used for RT-PCR analysis. The *actin* gene was used as the endogenous control of gene expression and the PCR product was analysed by DNA gel electrophoresis **(Fig. 4.14)**.

Most of the randomly selected seeds gave the expected DNA amplification product for the *AMP* transcript under study. Only *AMP* transcripts could not be observed in *Ole18::Ole18_PRS_Cec A::Nos-T* (R9, line 7.2) seeds (results are not shown because the **Fig. 4.14** is referred to T3 homozygous plants, and, as described below **(Section 4.3.7, Results)**, no T3 homozygous plants from the event R9 7.2 could be obtained). Integration of multiple copies of both transgenes *AMP* and *hptII* was observed by qPCR. As expected, AMP transcripts were not detected either in wild-type or empty vector pCAMBIA seeds. In this way, RT-PCR analyses confirmed transgene expression in developing seeds from the different transformation events generated in this work.

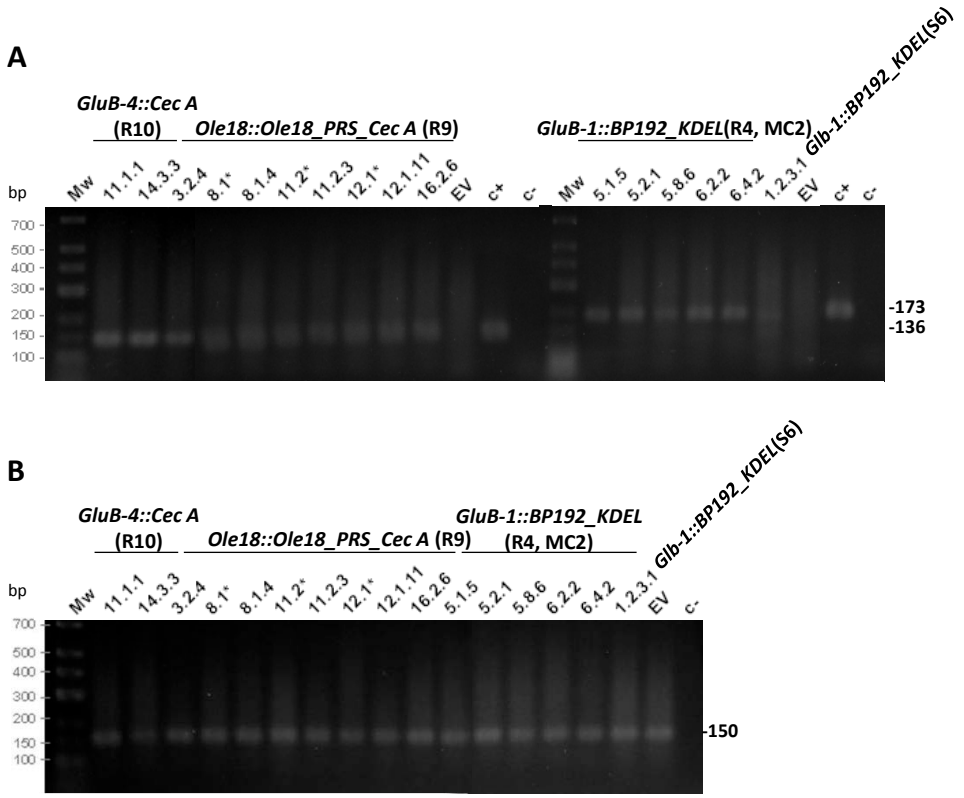


Fig. 4.14 Detection of transgene transcripts in immature rice seeds. *AMP* gene expression was assessed by RT-PCR using 2 µg of total RNA from immature T2-T3 rice seeds. (A) Representative results for transgenic lines harbouring the following transgenes are presented: *pCGluB-4::Cec A* (R10, 136 bp); *pCOle::Ole18_PRS_Cec* A (R9, 136 bp), *GluB-1::BP192_KDEL* (R4, MC2, 173 pb) and *Glb-1::BP192_KDEL* (S6, 173 pb). (B) The rice *actin* gene was used as endogenous control (150 bp). EV; empty vector line used as a negative control for transgene expression. c-, negative control of the PCR reaction; c+, DNA from pCAMBIA1300 control. Numbers in the left side indicate the size of the DNA fragments of the molecular weight DNA markers. Bp, base pairs.

4.3.4 Detection of the AMPs in rice seeds

The aim of the present study was to produce recombinant antimicrobial peptides in rice seeds by targeting the AMP of interest to a particular subcellular compartment such as protein bodies (seed endosperm) and oil bodies (seed aleurone and embryo). The detection of the *in planta* produced AMPs have proven to be a difficult task in this study. The main problem encountered relates to the biochemical nature of such peptides. In addition to being small sized peptides, they are amphyphatic and sticky in nature. Seeds from representative transformation events were analysed, namely

seeds from *Ole18::Ole18_PRS_Cec A* (R9), *GluB-1::Cec A_KDEL* (MC3), *GluB-1::Cec A* (A5), *GluB-1::BP178_KDEL* (A11), *GluB-4::Cec A* (R10), *GluB-4::BP178_KDEL* (A14) and *Glb-1::BP178_KDEL* (R11) lines (T1 generation). For these studies, seeds from plants harbouring single or multiple insertions were used.

We approached AMP detection by subcellular fractionation followed by Western blot analysis. The advantage of using this experimental approach is that it allows not only detection of the AMP and variations in AMP accumulation among tested lines, but also confirmation of expected subcellular compartments (oil bodies, protein bodies).

4.3.4.1 Detection of the cecropin A peptide accumulating in oil bodies of seeds

Rice oil bodies (and associated proteins) can be easily isolated and separated from other seed components by floating centrifugation where the oil body fraction forms an immiscible "fat pad" onto the upper surface of the supernatant. Thus, oil bodies were purified from seeds of non transgenic and transgenic lines bearing the *pCOle::Ole18_PRS_Cec A* construct (R9) by two consecutive cycles of the two layer flotation method as described in **Section 3.16.1.2, Methods**. Further, protein extracts from the oil body fractions were subjected by SDS-PAGE gel electrophoresis and transferred onto PVDF membrane. Cec A was immunodetected using an anti-Cec A antibody (Coca et al., 2006). This analysis demonstrated that a recombinant oleosin_PRS_Cec A fusion protein was properly produced and accumulated at different levels in oil bodies of all the transgenic lines assayed **(Fig. 4.15 and 4.16)**. No immunological reaction occurred with protein extracts from oil bodies isolated from control seeds (WT and empty vector). We noticed that the apparent molecular weight of the protein band corresponding to the 18 kDa oleosin_Cec A fusion protein of ~18 kDa was slightly lower than expected **(Fig. 4.15).** Thus, the addition of the cecropin A peptide and the processing site for the TEV protease to the C-terminus of the oleosin protein accounts for 47 extra amino acids. In this way the theoretical molecular mass

of the 18 kDa oleosin_PRS_Cec A fusion protein is 23.2 kDa (18 kDa oleosin (272 aa) + 1.2 kDa PRS (10 aa) + 4 kDa Cec A (37 aa), whereas the apparent molecular weight of the fusion protein in our electrophoretic system was ~18 kDa **(Fig. 4.15)**. In our electrophoretic system, however, the rice oleosins (18 kDa and 16 kDa isoforms) also exhibited an apparent molecular mass slightly lower that the expected one **(Fig. 4.15 A).** The identity of the oleosin isoforms was confirmed by mass spectrometry. For this, proteins obtained from the oil bodies of *japonica* rice were electrophoretically separated. Gel slice containing proteins in the range of 13-15 kDa and subjected to LC MS/MS analysis (Proteomic Service of Barcelona Science Park, Barcelona, Spain). MS data was analysed by MASCOT SEARCH. The true molecular weights of 17211 Da (corresponding to the 18 kDa rice oleosin), and 15190 Da (corresponding to 16 kDa oleosin) were obtained (p value <0.05) further confirming the anomalous electrophoretic mobility of these proteins in SDS-PAGE gels.

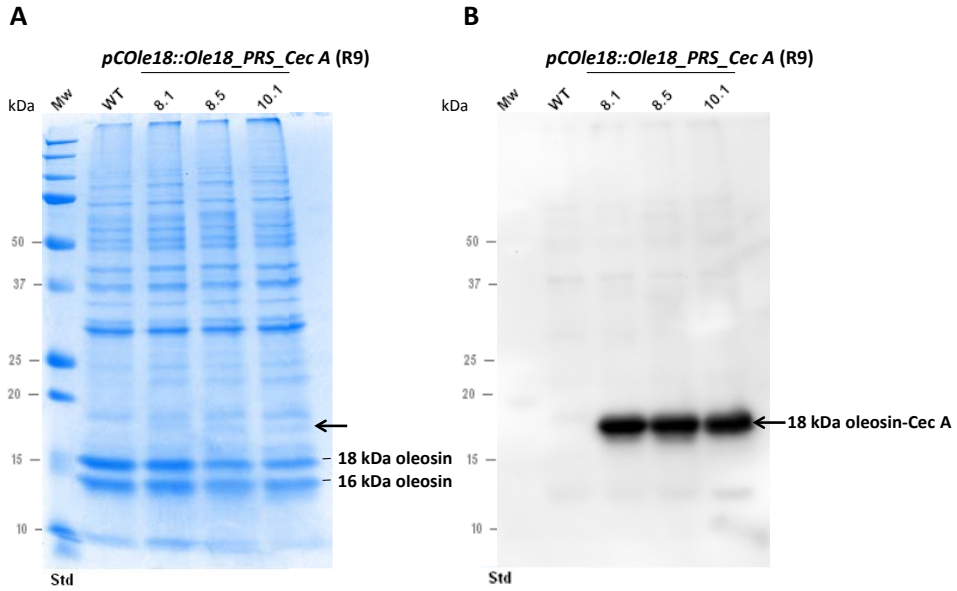


Fig. 4.15 Protein blot analyses of oil bodies isolated from seeds of rice plants transformed with the *pCOle::Ole18_PRS_Cec A* gene (R9 lines, T1 seeds). Oil bodies were purified from 5 seeds from each line. OB extracts (50 µg) were separated by SDS-PAGE gel and transferred onto PVDF membrane. Membranes were probed with an anti-Cec A polyclonal antibody. Immunological reactions were detected with an anti-rabbit IgG secondary antibody linked to horseradish peroxidase. (A) Coomassie-blue staining of rice proteins extracted from isolated oil bodies. The two rice oleosins isoforms (18 and 16 kDa) are indicated according to Chuang et al., 1996. (B)

Western blot analysis for detection of the oleosin_Cec A fusion protein in samples shown in (A). WT, non-transgenic rice. The oleosin_Cec A fusion protein is indicated by arrow. Numbers in the left indicate the molecular weight of protein markers (Precision Plus Protein™ Dual Color, Bio-Rad).

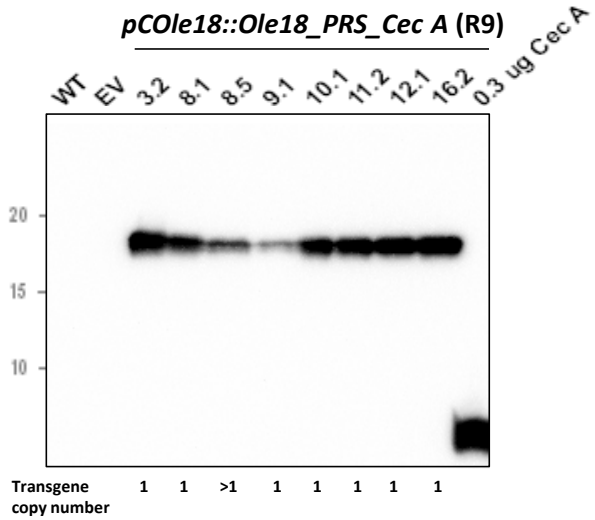


Fig. 4.16 Western blot analysis of oil body protein extracts (50 µg) from T1 transgenic rice seeds expressing the chimeric gene *oleosin_PRS_Cec A* using the anti-Cec A antibody. Depending on the line, different levels of oleosin_PRS_Cec A accumulation were observed. The same protein extracts were also probed with the anti-oleosin A antibody produced in this work **(see Fig. 4.17).** WT, untransformed seed; EV, seeds from transgenic event transformed with the empty pCAMBIA 1300 vector. 0.3 µg of the synthetic cecropin A peptide, used as a positive control. The transgene copy number of each line is indicated.

Detection of the oleosin_Cec A fusion protein was also approached using a polyclonal antibody raised against the 18 kDa rice oleosin. For this, the 18 kDa oleosin protein was purified from mature rice seeds and used to raise a polyclonal antibody. Oil body proteins from transgenic rice seeds (*Ole18::Ole18_PRS_Cec,* R9) and control seeds (Senia, EV) were probed with the anti-18 kDa oleosin antibody. Replicate membranes were prepared. One membrane was probed with anti-oleosin antibody, while the other membrane was probed with anti-Cec A antibody. As expected the oleosin_PRS_Cec A fusion protein was recognised by anti-Cec A antibody in *Ole18::Ole18_PRS_Cec A* seeds, whereas no immunological reaction occurred with protein extracts from, either wild-type or empty vector seeds **(Fig. 4.17 A)**. When the same protein samples were probed with anti-oleosin antibodies, the endogenous 18 kDa oleosin was detected in oil body proteins from both transgenic and wild type seeds, whereas the oleosin_Cec A fusion

protein was exclusively detected in the oil body protein fractions from *Ole18::Ole18_PRS_Cec A* transgenic seeds **(Fig. 4.17 B)**. The observed immunological reaction of the ~23 kDa protein with the anti-Cec A and the oleosin antibodies, confirmed the presence of the recombinant oleosin_Cec A fusion protein in the oil bodies of *Ole18::Ole18_PRS_Cec A* homozygous seeds. However, variable levels of recombinant protein accumulation were observed between independently transformed events.

Finally, when using Criterion technology (Precast gels, Bio-Rad) for electrophoretic separation of oil body extracts and immunodetection of the 18 kDa oleosin and 18 kDa oleosin_PRS_Cec A fusion protein, the "true" molecular mass (18 and ~23 kDa, respectively) was observed for these proteins **(Fig. 4.17)**. The apparent lower molecular mass that is observed for the aforementioned proteins in handmade SDS-PAGE gels could be attributed to conformational features of these proteins.

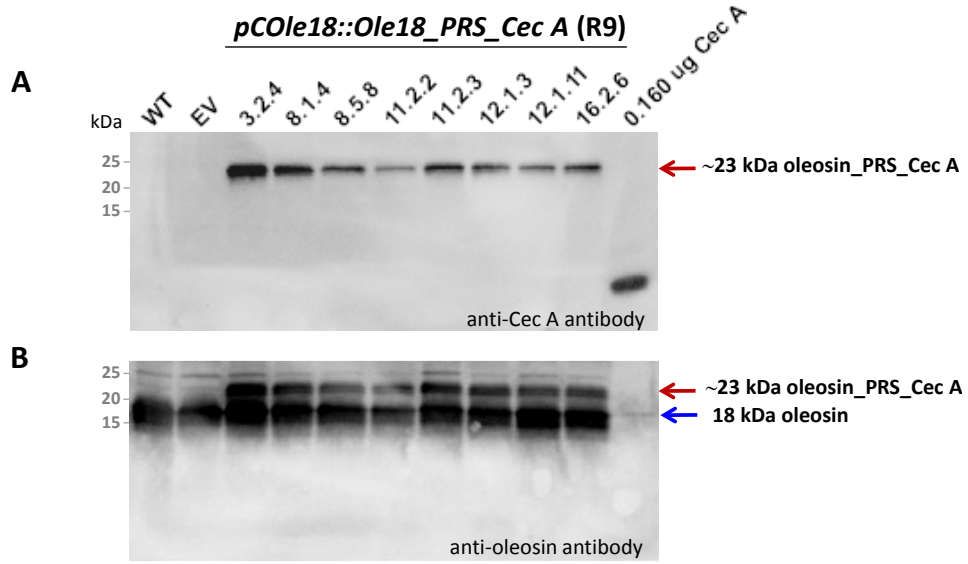


Fig. 4.17 Oleosin_Cec A fusion protein accumulates in the rice oil bodies of *pOle18::Ole18_PRS_Cec A* (R9) lines. Oil body protein extracts were prepared from mature rice seeds (T3 homozygous lines) and subjected to SDS-PAGE (35 µg each). Western blot analysis was performed using the anti-Cec A (A) or anti-oleosin (B) antibodies. Immunological reactions were revealed using an anti-rabbit IgG secondary antibody linked to horseradish peroxidase. WT, untransformed seed; EV, seeds from transgenic event transformed with the empty pCAMBIA 1300 vector. 0.16 µg of the synthetic cecropin A peptide was used as a positive control. (←) a 23 kDa protein band was recognized by both the anti-Cec A and anti-18 kDa oleosin antibodies.

4.3.4.2 Detection of AMP peptides in protein bodies of seeds

Traditionally, the obtention of protein bodies from rice seed endosperm has been performed by sucrose density gradient centrifugation (SDG) (Tanaka et al., 1980; Yamagata et al., 1982). Specifically, the endosperm homogenates are layered on top of 52% to 62% (w/w) linear SDG formed on a bed of 65% (w/w) of sucrose. After ultracentrifugation, PB are collected at about 56%-59% (w/w) sucrose. However, this technique is laborious and time-consuming. Clearly, this methodology was not convenient for the screening of the large number of independent transformation events generated in this work (≥ 125 independent transgenic lines).

In an attempt to simplify the methodology and to speed up the process of screening of seeds from rice transformants with reduced cost, a simplified method for preparation of partially purified protein bodies was set up. For this, seeds from the CecAder-lines (fifteen T1 seeds from each *GluB-4::Cec A* (R10), *GluB-1::Cec A_KDEL* (MC3) and *GluB-1::Cec A* (A5) lines) were randomly selected and their endosperms excised were manually excised. Endosperm tissues were then homogenized in buffer (0.6 M Sucrose, 10 mM Na_2HPO_4, pH 7.5) and centrifuged at low speed to remove most of the starch present in the samples. After a second centrifugation step (speed increased from 200 to 2000 rpm), the enriched protein body fraction was recovered from the supernatant. When necessary, an additional centrifugation step was performed. During protocol optimization, all steps were monitored by phase-contrast optical microscopy, to ensure starch removing and proper PB enrichment **(Fig. 4.18)**. By using this protocol, partially purified protein bodies can be prepared in about 15 min.

Once the enriched PB fractions were obtained, the proteins from PBs were analysed on a Tris-Tricine SDS-PAGE gel, blotted onto PVDF membrane and probed with antibodies raised against either the BP178 (or BP192) peptide or the cecropin A peptides. Results are presented in **Fig. 4.19**.

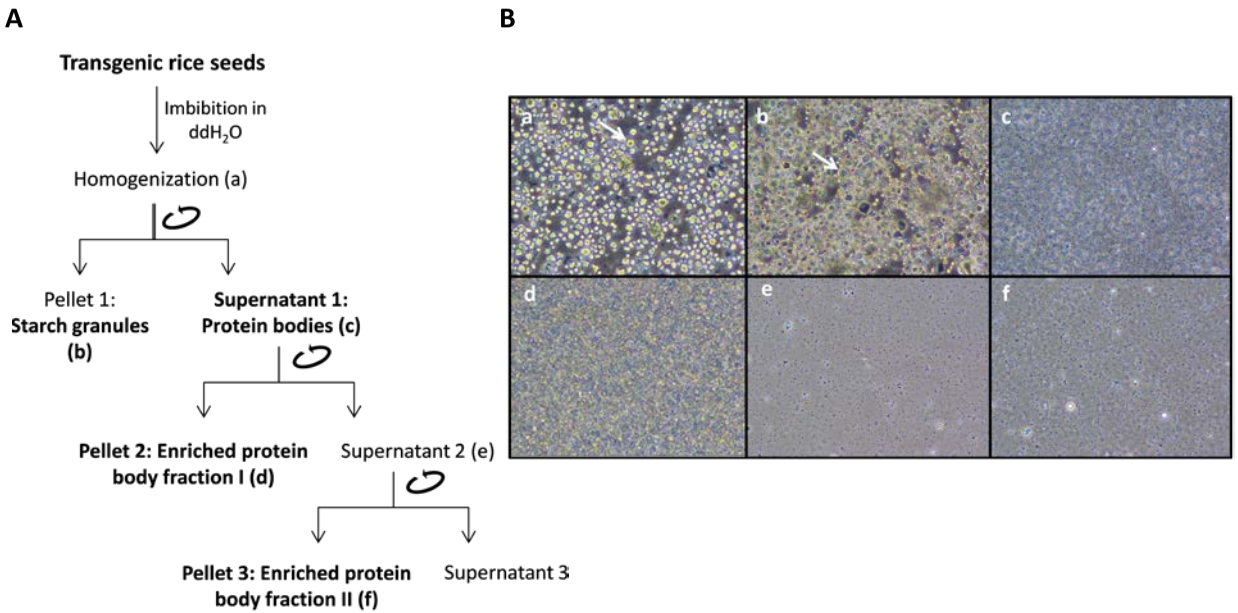


Fig. 4.18 Preparation of partially purified PBs from rice seeds of *pGluB-4::Cec A* transgenic lines (R10, line 9.1). (A) Procedure followed for the preparation of the protein body enriched fraction from mature seeds. (B) Phase-contrast microscope observations of protein body fractions from the rice endosperm of *GluB-4::Cec A* mature seeds. Protein body fractions were obtained following the protocol shown in A. (B) After homogenization of the endosperms in grinding buffer (0.6 M Sucrose, 10 mM Na_2HPO_4, pH 7.5), successive centrifugations were carried out to obtain an enriched protein body fraction. (a) Endosperm homogenates showing numerous starch granules. The homogenate was centrifuged recovering a pellet 1 (b) consisting of most starch granules and supernatant (c). Supernatant 1 was centrifuged, obtaining a pellet 2 composed of the enriched protein body fraction I (d). Supernatant 2 (e) was centrifuged obtaining a pellet 3 (f), which was added to the protein body fraction I obtaining a final enriched protein body fraction (d + f). Starch granules are indicated with an arrow.

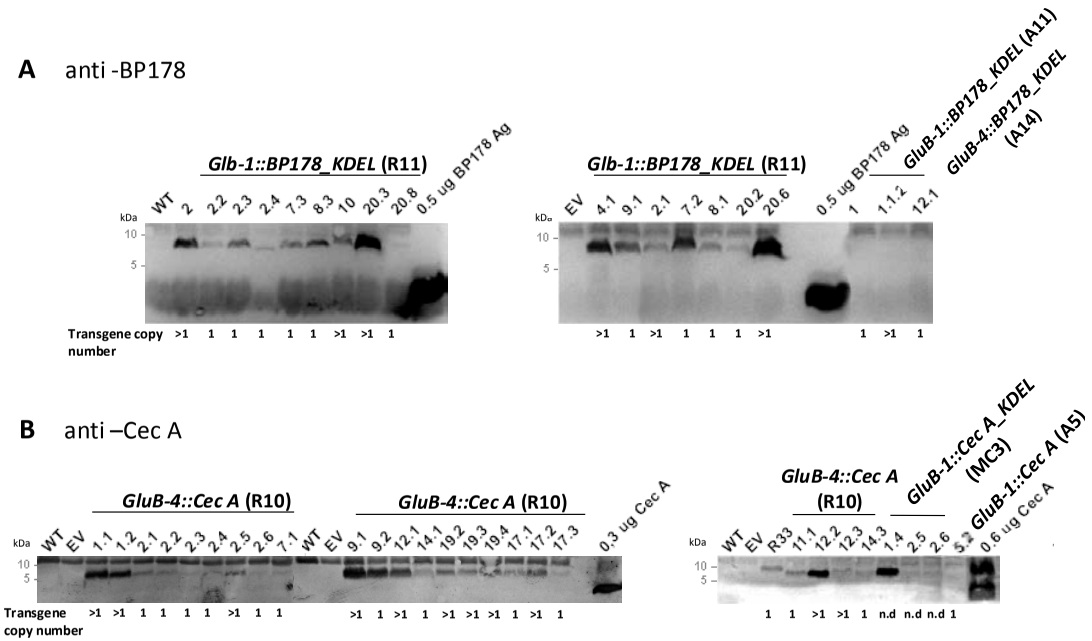


Fig. 4.19 Accumulation of BP178, Cec A_KDEL and Cec A recombinant peptides in seed protein bodies of transgenic rice (T1 generation). Protein extracts (50 µg) were separated on Tris-tricine-SDS-PAGE gel and transferred onto PVDF membrane. Membranes were incubated with the anti-BP178 antibody (A) or anti-Cec A antibody (B) followed by anti-rabbit IgG (Fc) secondary antibody linked to alkaline phosphatase. (A) Detection of BP178 peptide in *pCGlb-1::BP178_KDEL* (R11), *GluB1::BP178_KDEL* (A11) and *GluB-4::BP178_KDEL* lines. (B) Detection of Cec Ader in transgenic rice expressing the *GluB-4::Cec A* (R10), GluB-1::Cec A (A5) and *GluB-1::Cec A_KDEL* (MC3) transgene. WT, wild-type seeds used as a negative control; EV, seeds transformed with the empty pCAMBIA 1300 vector, used as a negative control. 0.3 µg of Cec A synthetic reference used as a positive control. T3 seeds bearing the *pCubi::ER-CecA::No-T* (Coca et al., 2006) were used as a positive control. Std, molecular standard Precision Plus Protein™ Dual Color. Numbers in the left indicate the size of the proteins of the length standard molecular weight. The trangene copy number in each line is indicated below.

The BP178 peptide was detected in all the lines harbouring the *pCGlb::BP178_KDEL* (R11) construct, but not in seeds carrying either the *pCGluB1::BP178_KDEL* (A11) or the *pCGluB4::BP178_KDEL* (A14) construct **(Fig. 4.19 A),** but the size of the peptide product was higher than the expected one (see comments below). Further analysis demonstrated that BP178 accumulates, but at very low levels in T1 seeds from *pCGluB4::BP178_KDEL* and *pCGluB1::BP178_KDEL* (A14 and A11, respectively) **(Section 4.4, Results)**. Finally, Cec A accumulated in the PBs of all *pCGluB-4::Cec A* (R10) and *pCGluB1::Cec A_KDEL* (MC3) events (T1 seeds) **(Fig. 4.19 B)**. As expected, no immunological reaction occurred in protein extracts from control seeds (WT, EV plants) **(Fig. 4.19 A, B).**

A large variation in accumulation levels was observed among different lines generated in a given transformation event, although using the same promoter to drive transgene expression. Variations observed in lines harbouring the same construct, could be attributed to: (1) different transgene copy number; (2) different integration site of the transgene into rice genome; (3) silencing mechanisms; or (4) physiological conditions in the various plants. Also, when the protein profiles of the transgenic and WT seeds samples were compared by SDS-PAGE, no major differences in the accumulation of native storage proteins (glutelins, prolamins) were observed (data not shown).

Based on their mobility on Tricine-SDS PAGE, the molecular weights of AMPders were estimated to be larger than the molecular weight of the synthetic peptide (samples subjected to Tris-Tricine SDS-PAGE in parallel in which the peptide was mixed with protein extracts from wild type seeds). This discrepancy may be explained by improper processing of the signal peptide, post-translational modifications in the peptide sequence, or to the presence of self-assembled peptide multimers.

The size of the synthetic Cec A (monomer) is 4.0 kDa. However, protein bands of approximately 4 kDa and 8 kDa were observed when synthetic Cec A was analysed by Western-blot. When large amounts of the synthetic BP178 are subjected to immunoblot analysis, we noticed that at least three immunoreactive bands are detected **(see Fig. 4.36, Section 4.4, Results)** that could be associated to peptide multimerization. To this point, it should be

mentioned that dimer formation was reported to occur in protein extracts from leaves of rice plants constitutively expressing the *Cec A* gene (Coca et al., 2006). Similarly, when a large amount of Cec A peptide is subjected to Western blot, a pattern of immunological reactions occur which can be correlated with multimers formation (Coca et al., 2006). The "abnormal" electrophoretic mobility of the in planta synthesized AMPs in the Tris-Tricine SDS-PAGE system will be further discussed in the following section **(Section 4.4, Results)**.

Overall, the method here developed for subcellular fractionation in combination with Western-blot analysis confirmed accumulation of BP178_KDEL, BP192_KDEL, BP188_KDEL, Cec A_KDEL and Cec A in transgenic rice seeds, and in the expected subcellular compartments, protein bodies or oil bodies depending on the strategy.

The next step was to confirm that the recombinant peptides accumulated in seed tissues and not in vegetative tissues of the transgenic rice. For this purpose, protein extracts were prepared from leaves and roots of 10 day-old transgenic and non transgenic rice plants, as well as from mature seeds. Protein samples were subjected to Tris-Tricine SDS-PAGE gel electrophoresis, followed by Western-blot analysis. As controls, transgenic rice lines expressing *Cec A* or *Cec A_KDEL* transgenes under the control of the constitutive maize *ubiquitin-1* (*Ubi*) promoter were assayed (Coca et al., 2006).

As expected, the Cec A and Cec A_KDEL peptides were detected in all the tissues, vegetative and seed tissues, of the *pCUbi::Cec A::Nos-T* and *pCUbi::Cec A_KDEL::Nos-T* transgenic plants **(Fig. 4.20)**. By contrast, no accumulation of Cec A occurred in leaves or roots from plants that express the AMP gene under the control of a seed specific promoter. As expected, when protein extracts from empty vector and WT rice tissues were analysed, the Cec A peptide was not detected in any of the tested tissues **(Fig. 4.20)**.

It is worthwhile to emphasise that the localization site of the CecAder peptides in the rice seed was not influenced by the presence or absence of the KDEL retention signal. This observation suggests that the signal peptide sequence of the seed storage protein (either GluB-1 or GluB-4) contains the

targeting signals that allow compartmentalization of the recombinant peptides into seed protein bodies. Similar results to those presented for the Cec A lines were obtained when BP178-lines were analysed (data not shown).

These results confirmed that AMPs accumulate in seeds, but not in leaves and roots, of the transgenic lines produced in this work, further supporting the already described seed-specific activity of the *GluB-1* and *GluB-4* promoters. In addition, the *in planta* produced AMP peptides were found to accumulate in a stable manner in the rice seeds as they could be detected in the transgenic seeds even after long periods of storage (at least up to 2-3 years), even when stored at room temperature.

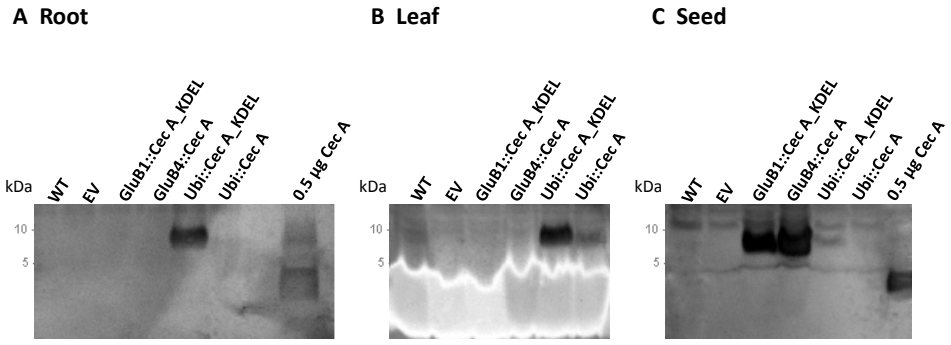


Fig. 4.20 Western blot analysis of protein extracts from rice plants expressing the *Cec A* or *Cec A_KDEL* gene under the control of either a seed-specific rice promoter or the constitutive *ubiquitin* (*Ubi*) promoter. Total protein extracts (60 µg) from mature seeds, leaves or roots were separated by Tris-Tricine SDS-PAGE and transferred onto PVDF membrane. Membranes were incubated with the anti-Cec A antibody followed by anti-rabbit IgG (Fc) secondary antibody linked to alkaline phosphatase. WT, non transformed event; EV, rice plants with the empty pCAMBIA 1300 vector, used as a negative control. Plants expressing the *pCGluB-1::Cec A_KDEL* gene (MC3 1.4); plants expressing the *pCGluB-4::Cec A* gene (R10 9.2). Events carrying the *pCUbi::Cec A_KDEL* or *pCUbi::Cec A_KDEL* genes, were examined (R33 and R26, respectively). 0.5 µg of Cec A was used as a positive control of western blot analysis.

4.3.5 Resistance to pathogen infection in rice seeds

Additional evidences that the *in planta* produced AMP peptides were properly synthesized and biologically active came from the evaluation of disease resistance of the transgenic rice plants. A germination test was developed to assess the susceptibility of transgenic rice seedlings to

infection by bacterial (*Dickeya* sp. 1552.10.1) and fungal (*Fusarium verticillioides*) pathogens. *Dickeya* sp. is a phytopathogenic bacteria that causes soft rot disease, being the causal agent of the dark brown sheath rot of rice (Li et al., 2012). *Fusarium verticillioides* (anamorph stage of *Gibberella fujikuroi*, mating population A) is the causal agent of the bakanae disease of rice and infects seed and root tissues (Wulff et al., 2010). By using this experimental approach, we determined the incidence and severity of infection as well as differences in susceptibility within and between transgenic events and control plants. The results obtained through infection bioassays together with those obtained on transgene copy number **(Section 4.3.3, Results)** and AMPder peptide detection **(Section 4.3.4, Results)** were the basis for selection of the transgenic events to be next used for massive seed production and purification of AMPs from rice seeds **(Section 4.3.6, Results)**.

As a preliminary set-up of the infection assay, five plant pathogenic bacterial strains were tested using the rice cv. Senia (non-transgenic seeds). This study allowed us to choose the most suitable strain to perform bioassays on transgenic seeds. For this purpose Senia wild-type seeds were surface sterilized and inoculated with increasing doses of the pathogenic strains here assayed (*Dickeya* sp. 1552.10.1, *Dickeya* sp. 1428 1a, *P. carotovorum* 578.3, *P. carotovorum* 3902.3 and *P. carotovorum* 558-6-b). As control, Senia seeds were inoculated with double distilled H_2O. Under our experimental conditions, seven days after inoculation, the *Pectobacterium carotovorum* 578.3, 3902.3 and 558-6-b strains-infected seeds germinated normally. These seedlings presented the same morphology as the control seedlings (water-inoculated). By contrast, the Senia seeds infected with the *Dickeya* sp. strains showed a clear decrease on seed germination capability, and in most cases the infected seeds were unable to germinate. **Fig. 4.21** shows representative results of infection bioassays of Senia seeds with each of the five pathogenic bacterial strains assayed.

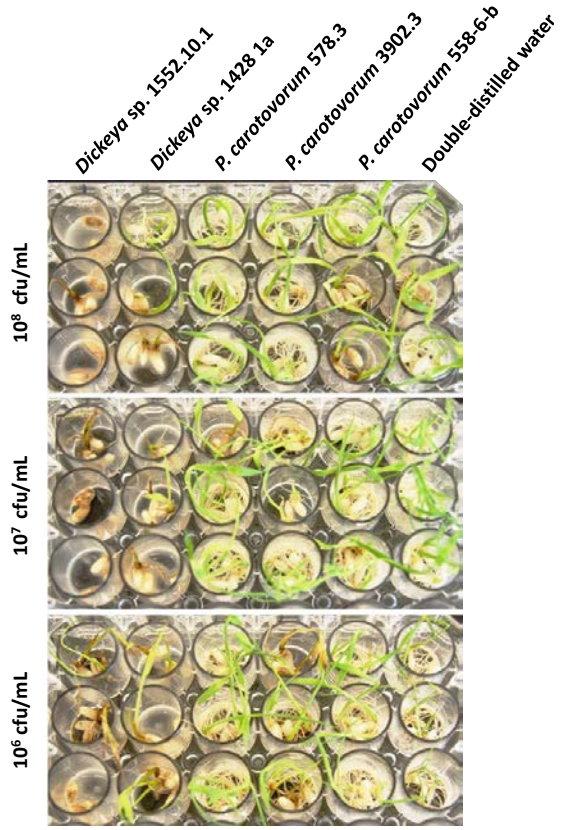


Fig. 4.21 Susceptibility of Senia seeds to infection by plant pathogenic bacterial strains. Seeds of non-transformed (WT) plants were surface-sterilized and vacuum infiltrated with the bacterial suspension at the indicated concentrations. Bacterial pathogens (and strains) were *Dickeya* sp. 1552.10.1; *Dickeya* sp. 1428 1a; *P. carotovorum* 578.3; *P. carotovorum* 3902.3; *P. carotovorum* 558-6-b. Pictures show the appearance of the infected seedlings 7 days after infiltration with the bacterial suspension.

The *Dickeya* sp. 1552.10.1 was selected to perform infection bioassays on rice seeds from the different transgenic lines: *pCOle18::Ole18_PRS_Cec A* (R9), *pCGluB-1::BP178_KDEL* (A11, S11), *pCGluB-1::BP188_KDEL* (A12, R7), *pCGluB-1::BP192_KDEL* (MC2), *pCGluB-1::Cec A_KDEL (*MC3), *pCGluB-1::Cec A* (A5), *pCGluB-4::BP188_KDEL* (R8), *pCGluB-4::Cec A* (R10), *pCGluB-4::BP192_KDEL* (A7), *pCGlb-1::BP192_KDEL* (S6) and *pCGlb-1::BP178_KDEL* (R11). Representative results are presented in **Fig. 4.22**.

Even though differences in susceptibility occurred among different lines from the same transformation event, the *Dickeya*-infected transgenic seeds showed in many cases, no disease or reduced disease symptoms compared to control seeds (WT and empty vector). In our experimental conditions, transgenic lines expressing *pCGluB-4::BP188_KDEL, pCOle18::Ole18_PRS_Cec A*, *pCGluB-4::Cec A, pCGluB-1::BP192_KDEL*, *pCGluB-1::Cec A_KDEL*, *pCGluB-1::Cec A*, *pCGlb-1::BP178_KDEL*, *pCGluB-4::BP178_KDEL* and *pCGluB-4::Cec A_KDEL*, showed a clear and consistent phenotype of resistance to infection by the bacterial pathogen *Dickeya* sp. 1552.10.1, compared to control plants (WT, empty vector) **(Fig. 4.22)**.

Surprisingly, disease symptoms became much more pronounced in transgenic events expressing *BP178_KDEL* or *BP188_KDEL* under the control of the *GluB-1* promoter **(Fig. 4.22, panel B)**. Infected seedlings from these transformation events, exhibited shorter shoots and brownish roots. Finally, rice seeds harbouring the *BP192_KDEL* gene under the control of *Glb-1* promoter as well as with lines expressing either the *BP178_KDEL* or *BP188_KDEL* under the control of *GluB-1* promoter were found to be the most sensitive lines, these seeds being in most cases unable to germinate in the presence of the bacterial phytopathogen.

An index was used to estimate *Dickeya* sp. effects on germinating WT and transgenic seeds using a semi-quantitative scale **(Table 4.18)**.

Table 4.18 Scale used to asses resistance to bacterial infection

Scale	Germination	Description of seedling growth
0	-	-
0.5	+	The seedling length was less than 25% of control seedling
1.0	+	The seedling length was less than 50% of control seedling and presented an altered root or shoot development
1.5	+	The seedling length was less than 75% of control seedling and presented an altered root or shoot development
2.0	+	The seedling presented the same morphology as the control seedling

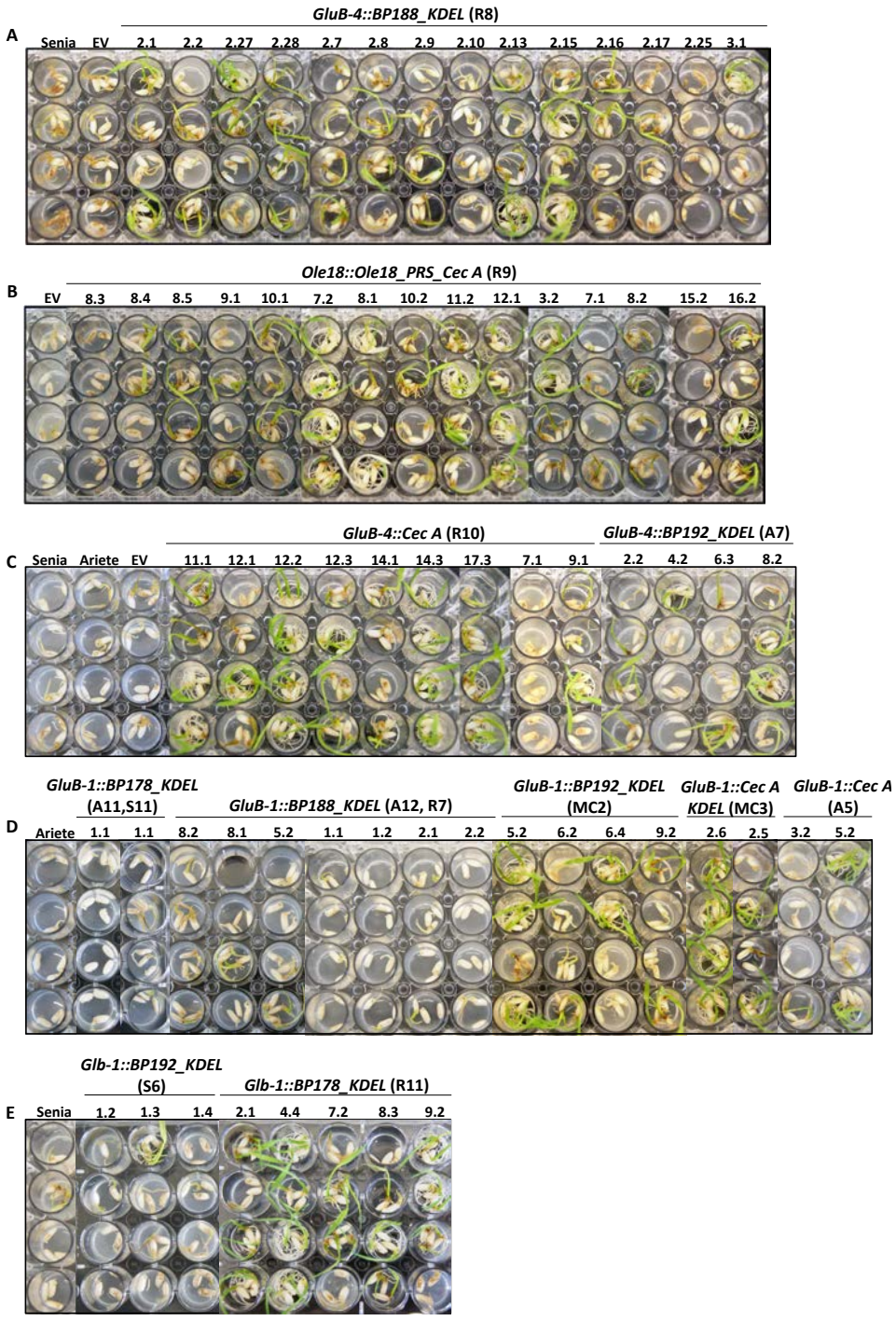


Fig. 4.22 Disease symptoms caused by *Dickeya* sp. 1552.10.1 infection on germinating wild-type and transgenic rice seeds (T1 generation). Representative results are presented for the following lines: (A) *pCGluB-4::BP188_KDEL* (R10), (B) *pCOle18::Ole18_PRS_Cec A* (R9), (C) *pCGluB-4::Cec A*

(R10), *pCGluB-4::BP192_KDEL* (A7), (D) *pCGluB1::BP178_KDEL* (A11, S11), *pCGluB-1::BP188_KDEL* (A12, R7), *pCGluB-1::BP192_KDEL* (MC2), *pCGluB-1::Cec A_KDEL (*MC3), *pCGluB-1::Cec A* (A5)and (E) *pCGlb-1::BP192_KDEL* (S6) and *pCGlb-1::BP178_KDEL* (R11). Seeds were inoculated with a bacterial suspension at 10^7 cfu/mL, vacuum infiltrated and allowed to germinate. Pictures were taken 7 days after inoculation. Senia and Ariete, non-transformed seeds; EV, transformed event carrying the pCAMBIA 1300 vector. At least two independent experiments using 12 seeds for each line, were performed.

Disease severity was determined as in the previous bioassay and the relative resistance was calculated. Control seeds incubated with water served to check seed germination capability. The relative germination index (transgenic events *vs.* to controls) is presented in **Fig. 4.23** and row data on *Dickeya*-inoculated seeds are presented in **Annex 8.2**.

Generally, the expression of *BP134der* and *CecAder* transgenes in rice seeds under the control of seed specific promoters resulted in a high level of protection to infection by the bacterial phytopathogen *Dickey*a sp. 1552.10.1.

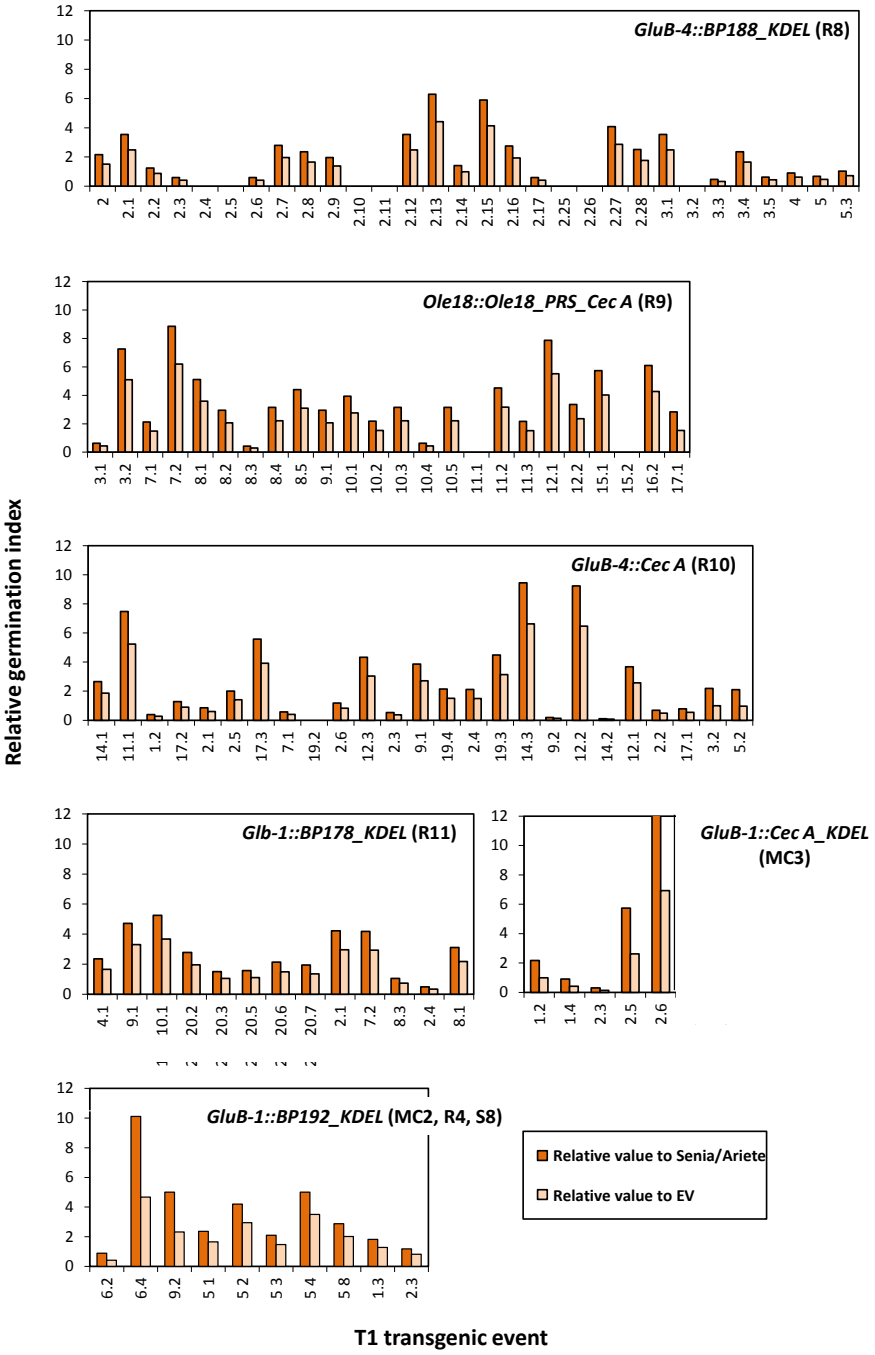


Fig. 4.23 Germination capability of *Dickeya*-infected T1 rice seeds expressing *AMPder*. Seeds were inoculated with a bacterial suspension (*Dickeya* sp. 1552.10.1, 10^7 cfu/mL), vacuum infiltrated and allowed to continue germination. Means of relative germination index were determined as relative values of transgenic events in relation to WT (Senia or Ariete) and empty vector (EV, R1II 7 line) controls. Germination index was estimated at 7 days after inoculation. A value of 0 indicate no germination while values > 1 indicate the number of times in which the transgenic germination index was higher than that of controls (Senia or Ariete, non-transformed seeds); EV, seeds transformed with the empty pCAMBIA 1300 vector.

Similar experiments were carried out using T3 seed (homozygous plants). As an example, **Fig. 4.24** represents the susceptibility of representative lines, from different transformation events, to the infection by the bacterial pathogen *Dickeya* sp. 1552.10.1.

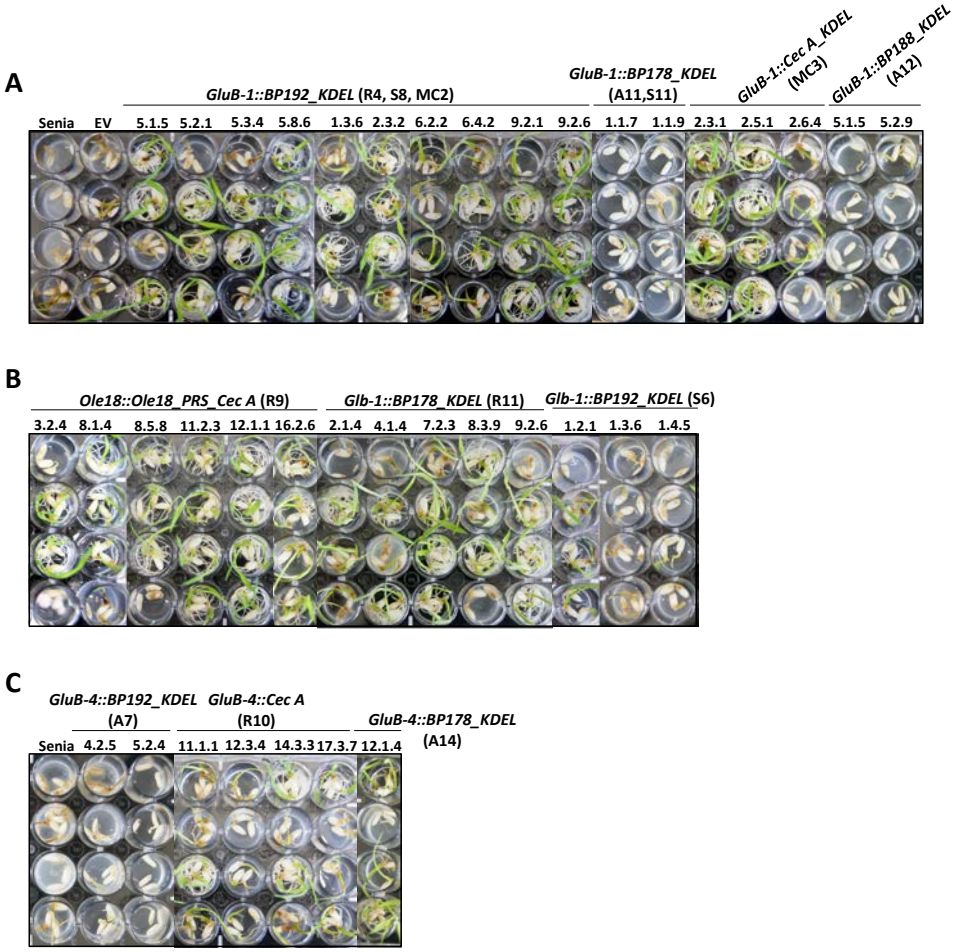


Fig. 4.24 Disease symptoms caused by *Dickeya* sp. 1552.10.1 infection on germinating wild-type and transgenic T3 rice seeds. (A) *pCGluB-1::BP192_KDEL* (R4, MC2 and S8), *pCGluB-1::BP178_KDEL* (A11, S11), *pCGluB-1::Cec A_KDEL (*MC3) and *pCGluB-1::BP188_KDEL* (A12). (B) *pCOle18::Ole18_PRS_Cec A* (R9), *pCGlb-1::BP178_KDEL* (R11) and *pCGlb-1::BP192_KDEL* (S6). (C) *pCGluB-4::BP192_KDEL* (A7), *pCGluB-4::Cec A* (R10), *pCGluB-4::BP178_KDEL* (A14). Seeds were inoculated with a bacterial suspension at 10^7 cfu/mL, vacuum infiltrated and allowed to germinate. Pictures were taken 7 days after inoculation. Senia and Ariete, non-transformed seeds; EV, seeds carrying the pCAMBIA 1300 empty vector. At least two independent experiments containing 12 seeds for each line were performed.

When comparing the results obtained with T1 and T3 seeds, the T3 lines expressing the *AMPder* transgene in homozygosis frequently showed higher level of protection to infection by *Dickeya* sp. 1552.10.1 relative to that in the corresponding T1 hemizygous transformants **(Fig. 4.25**. Row data on germinating capability of *Dickeya*-inoculated T1 *vs* T3 seeds are presented in **Annex 8.3)**. This fact could be associated to the transgene dosage and/or transgene expression levels in transgenic plants (homozygous *vs*. hemizygous). However, results obtained in our infection assays (T3 seeds) did not revealed a clear relationship between phenotype (resistance/susceptibility) and transgene copy number or level of transgene expression. The possibility that the observed phenotype of resistance/susceptibility of the transgenic plants is associated not only to transgene expression levels, but also to pleiotropic effects of transgenesis that are not related to the antimicrobial activity of the AMP should be considered (Bhat and Srinivasan, 2002). Further studies are needed to clarify this aspect.

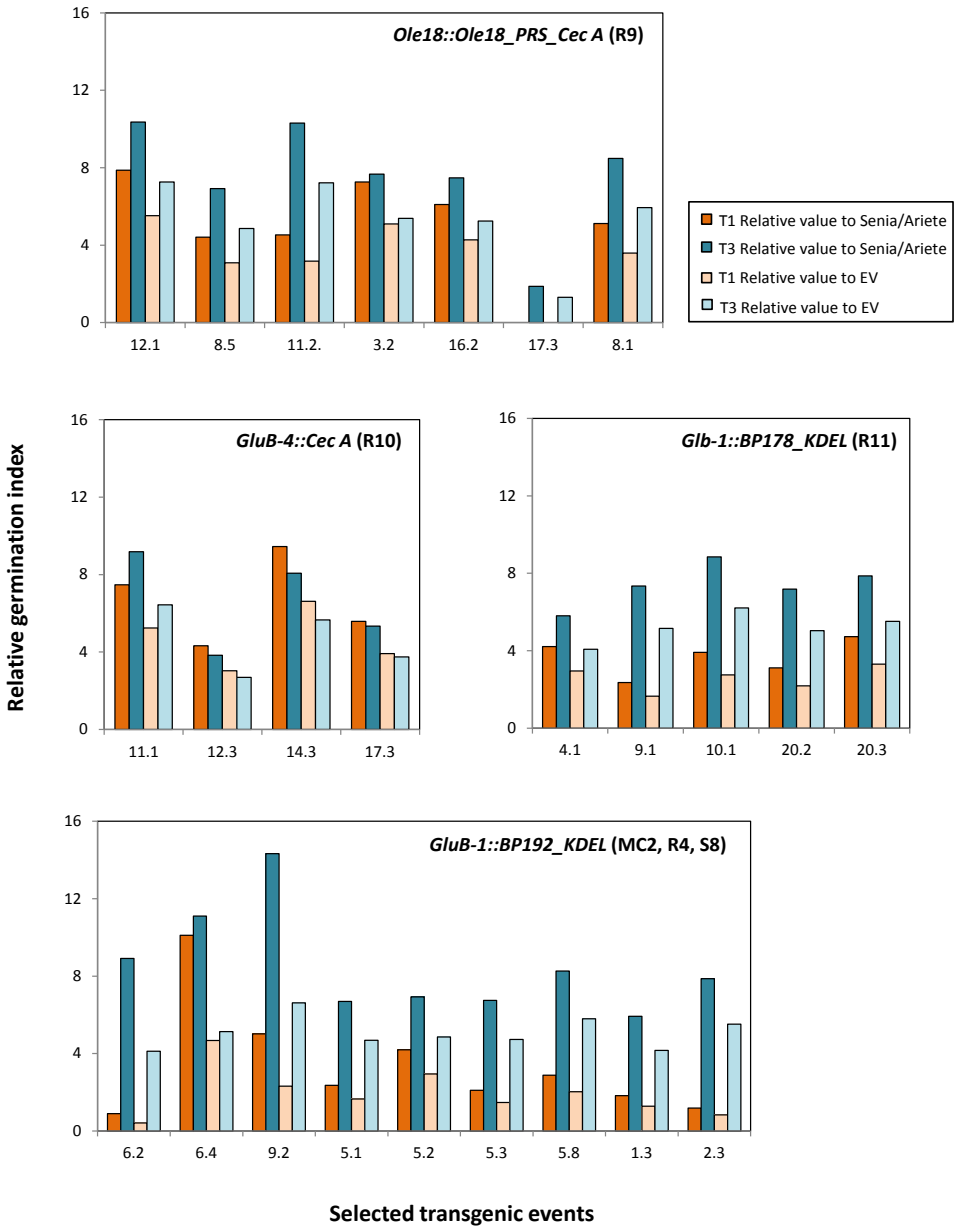


Fig. 4.25 Germination capability of *Dickeya*-infected T1 and T3 rice seeds expressing an *AMP* gene. Seeds were inoculated with a bacterial suspension at a concentration of 10^7 cfu/mL, vacuum infiltrated and allowed to germinate. Means of relative germination index were determined as a relative values of transgenic events in relation to WT (Senia or Ariete) and EV controls. The germination index was estimated at 7 days after inoculation. A value of 0 indicates no germination while values of >1 indicate the number of times that transgenic germination index is higher than that of controls. WT and Ariete, non-transformed seeds; EV, transformed event carrying the empty pCAMBIA 1300 vector.

Next, we examined susceptibility of the transgenic rice seeds to infection by the fungal pathogen *Fusarium verticillioides*. *Fusarium* infection bioassays were carried out by M. Coca (CRAG Centre, Barcelona, Spain) within the framework of the European Trilateral Plant KBBE project SEPSAPE (Safe and Efficient Plant Systems for Antimicrobial PEptide production).

Essentially, this study revealed a better germination capability of the fungal-infected transgenic seeds harbouring a *Cec A* gene (with or without the KDEL extension) under the control of a glutelin promoter (*GluB-1, GluB-4*), compared to fungal-infected control seeds (WT, empty vector). Seeds harbouring the *oleosin_Cec A* fusion gene showed no differences in susceptibility to fungal infection. As for the rice seeds harbouring a BP134der, all the transgenic events expressing BP192 exhibited susceptibility to *F. verticillioides* infection and at similar levels than *F. verticillioides*-infected control seeds. By contrast, rice seeds expressing *BP188_KDEL* (BP134_Cec A hybrid peptide) or *BP178_KDEL* (BP134_magainin hybrid) showed enhanced resistance to *F. verticillioides* infection. Here, it should be mentioned that, among other requisites, BPders were previously selected based on their antibacterial activity **(Section 4.1, Results)** and their antifungal activity was not assayed by *in vitro* infection assays.

Overall, the observed accumulation of the AMPders in rice seeds together with the results obtained in disease resistance (protection to pathogen infection) indicated that the *in planta* produced AMPders were biologically active.

4.3.6 Selection of transgenic rice lines

The criteria for selection of the parental lines to be next used for the obtention of homozygous transgenic lines included: (1) verification of transgene insertion and copy number, that is lines harbouring single copy insertion of transgenes; eventually more than one, e.g. two or three insertions; (2) resistance to infection by *Dickeya* sp. and *F. verticillioides*; and (3) high level of AMP accumulation in the corresponding subcellular compartment of rice seeds. Based on these criteria, at least 6 lines for each transformation event were selected to obtain the corresponding

homozygous lines and subsequent seed amplification at CIDSAV and CRAG. The selected lines are listed in **Table 4.19**.

Table 4.19 Molecular and phenotypic characterization of selected transgenic lines

		Transgene copy number			Seed infection bioassay	
Construct	**Code**	***AMPder***	***hptII***	**AMP detection**	***Dickeya***	***Fusarium***
pCGluB-1::BP178_KDEL	S11 1.1	n.d	n.d	+	S	R
	A11 1.1	1	1	-	S	R
pCGluB-1::BP188_KDEL	A12 5.1	n.d	n.d	n.c.	R	S
	A12 5.2	1	1	n.c.	R	S
	A12 8.1	n.d	n.d	n.c.	R	S
	A12 8.2	1	1	n.c.	R	S
pCGluB-1::BP192_KDEL	R4 5.1	>1	1	n.c.	R	S
	R4 5.2	1	0	n.c.	R	S
	R4 5.3	1	1	n.c.	R	S
	R4 5.4	>1	0	n.c.	R	S
	R4 5.8	n.d	n.d	n.c.	R	S
	S8 1.3	>1	1	n.c.	R	S
	S8 2.3	>1	1	n.c.	R	S
	MC2 5.1	n.d	n.d	n.c.	R	S
	MC2 5.2	>1	>1	n.c.	R	S
	MC2 5.3	n.d	n.d	n.c.	R	S
	MC2 6.2	1	1	n.c.	R	S
	MC2 6.3	n.d	n.d	n.c.	R	S
	MC2 6.4	1	1	n.c.	R	S
	MC2 6.5	n.d	n.d	n.c.	R	S
	MC2 9.2	1	1	n.c.	R	S
pCGluB-1::Cec A_KDEL	S3 2.2	n.d	n.d	+	R	R
	MC3/A3	n.d	n.d	+	R	R
	MC3 1.3	n.d	n.d	+	R	R
	MC3 2.3	n.d	n.d	+	R	R
	MC3 2.5	n.d	n.d	+	R	R
	MC3 2.6	n.d	n.d	+	R	R
pCGluB-1::Cec A	A5 1.2	>1	>1	+	R	R
	A5 3.2	>1	>1	+	R	R
	A5 5.1	n.d	n.d	+	R	R
pCGluB-4::BP178_KDEL	A14 11.1	n.d	n.d	-	R	R
	A14 12.1	1	1	-	R	R
	A14 13.2	n.d	n.d	-	R	R

Table 4.19 (continued)

Construct	Code	Transgene copy number		AMP detection	Seed infection bioassay	
		AMPder	*hptII*		*Dickeya*	*Fusarium*
pCGluB-4::BP188_KDEL	R8 2.7	1	1	n.c.	R	R
	R8 2.13	>1	1	n.c.	R	R
	R8 2.15	1	1	n.c.	R	R
	R8 3.1	1	1	n.c.	R	R
	R8 2.27	>1	>1	n.c.	R	R
pCGluB-4::BP192_KDEL	A7 4.2	>1	>1	n.c.	R	S
	A7 5.2	>1	>1	n.c.	S	S
	A7 6.2	>1	>1	n.c.	R	S
pCGluB-4::Cec A_KDEL	S2 1.1	n.d	n.d	+	R	R
	S2 2.2	n.d	n.d	+	R	R
	S2 4.8	n.d	n.d	+	R	R
pCGluB-4::Cec A	R10 9.1	>1	>1	+	R	R
	R10 11.1	1	1	+	R	R
	R10 12.1	>1	>1	+	R	R
	R10 12.2	>1	>1	+	R	R
	R10 12.3	>1	>1	+	R	R
	R10 14.3	1	1	+	R	R
	R10 17.3	1	1	+	R	R
	R10 19.3	>1	>1	+	R	R
	A1 2.1	n.d	n.d	+	R	R
	A1 3.1	n.d	n.d	+	R	R
	A1 3.2	n.d	n.d	+	R	R
	A1 5.2	n.d	n.d	+	R	R
pCGlb-1::BP178_KDEL	R11 2.1	>1	1	+	R	n.d
	R11 4.1	>1	1	+	R	n.d
	R11 7.2	1	1	+	R	n.d
	R11 8.1	1	>1	+	R	n.d
	R11 9.1	1	1	+	R	n.d
	R11 10	>1	>1	+	R	n.d
	R11 20.3	>1	>1	+	R	n.d
	R11 20.6	>1	>1	+	R	n.d
	R11 20.7	1	1	+	R	n.d
pCGlb-1::BP192_KDEL	S6 1.2	n.d	n.d	-	R	S
	S6 1.3	n.d	n.d	-	S	S
	S6 1.4	n.d	n.d	-	S	S

Table 4.19 (continued)

Construct	Code	Transgene copy number		AMP detection	Seed infection bioassay	
		AMPder	*hptII*		*Dickeya*	*Fusarium*
pCOle18::Ole18_PRS_Cec A	R9 3.2	1	1	+	R	S
	R9 7.2	>1	>1	+	R	S
	R9 8.1	1	1	+	R	S
	R9 8.5	>1	>1	+	R	S
	R9 10.1	1	1	+	R	S
	R9 11.2	1	1	+	R	S
	R9 12.1	1	1	+	R	S
	R9 16.2	1	1	+	R	S
	A16 2.1	n.d	n.d	+	n.d	S
	A16 3.2	n.d	n.d	+	n.d	S
	A16 4.1	n.d	n.d	+	n.d	S
	A16 6.2	n.d	n.d	+	n.d	S
	A16 9.2	n.d	n.d	+	n.d	S

Rice genotype used for rice transformation: Ariete (A, MC); Senia (R, S). n.d; not determined; R and S, resistance and sensibility to infection by *Dickeya* or *Fusarium*. Western blot analysis: +, positive reaction with AMPder antibodies; -, absence of positive reaction with AMPder antibodies; n.c., not clear immunoreactive band.

4.3.7 Production of homozygous transgenic rice lines

As described previously, the T-DNA inserted into plant genome contains the *hptII* selectable marker gene that confers resistance to hygromycin. Accordingly, homozygous lines were selected for their ability to germinate in the presence of hygromycin. Because segregation ratios depend on the number of inserts integrated into the T0 rice genome, the frequency of hygromycin-resistant plants among the seed progeny allowed us to estimate the presumptive number of transgene insertions. Mendelian ratio of 3:1 for the transgene (presence:absence) was expected for a single integration locus while ratios of 15:1 and 63:1 were expected in those T0 events presenting two or three copies of the insert integrated into independent loci, respectively (Christou et al., 1989; Srivatava et al., 1996; Chen et al., 1998; Sallaud et al., 2003; Zhang et al., 2006, Tizaoui and Kchouk, 2012).

To obtain homozygous seeds and analyse the *hptII* inheritance, 24 randomly chosen T1 seeds of the selected T0 transgenic lines were germinated in the presence of hygromycin. Based on theoretical considerations, a 3:1 ratio is

expected for segregation of T0 events harbouring a single copy of the transgene (as determined byqPCR assays). To examine whether experimental segregation values were consistent with the theoretical values corresponding to a 3:1 Mendelian segregation, the χ^2 test was performed **(Table 4.20 and 4.21)**.

Table 4.20 Representative data of segregation analysis for the hygromycin selectable marker in the T1 progeny. The segregation frequencies correspond to a total of 24 seeds.

	T1 seed code	**Estimated *hptII* copy number by qPCR**	**Segregation of Hyg^R**	**Probability X^2 <0.05 (ratio 3:1[a])**
pCAMBIA1300	EV	1	19	0.637 (n.s)
pCGluB-1::BP192_KDEL	MC2 5.2	>1	0	0.000 (s)
	MC2 6.2	1	20	0.346 (n.s)
	MC2 6.4	1	19	0.637 (n.s)
	MC2 9.2	1	17	0.637 (n.s)
	S8 1.3	1	16	0.346 (n.s)
	S8 2.3	1	20	0.346 (n.s)
	R4 5.1	1	18	1.000 (ns)
	R4 5.2	0	17	0.637 (n.s)
	R4 5.3	1	18	1.000 (n.s)
	R4 5.4	0	22	0.059 (n.s*)
	R4 5.8	n.d	17	0.637 (n.s)
pCGluB-1::Cec A_KDEL	MC3 1.2	n.d	0	0.000 (s)
	MC3 1.4	n.d	0	0.000 (s)
	MC3 2.3	n.d	22	0.059 (n.s*)
	MC3 2.5	n.d	17	0.637 (n.s)
	MC3 2.6	n.d	17	0.637 (n.s)
pCGlb-1::BP192_KDEL	S6 1.2	n.d	20	0.346 (n.s)
	S6 1.3	n.d	24	0.000 (s)
	S6 1.4	n.d	20	0.346 (n.s)
pCOle18::Ole18_PRS_Cec A	R9 3.2	1	21	0.157 (n.s)
	R9 7.2	>1	21	0.157 (n.s)
	R9 8.1	1	19	0.637 (n.s)
	R9 8.5	>1	17	0.637 (n.s)
	R9 10.1	1	18	1.000 (n.s)
	R9 11.2	1	15	0.157 (n.s)
	R9 12.1	1	18	1.000 (n.s)
	R9 16.2	1	19	0.637 (n.s)

Table 4.20 (continued)

	T1 seed code	Estimated *hptII* copy number by qPCR	Segregation of Hyg^R	Probability X^2 <0.05 (ratio 3:1[a])
pCGluB-4::Cec A_KDEL	R10 9.1	>1	0	0.000 (s)
	R10 11.1	1	18	1.000 (n.s)
	R10 12.1	>1	0	0.000 (s)
	R10 12.2	>1	0	0.000 (s)
	R10 12.3	>1	19	0.637 (n.s)
	R10 14.3	1	20	0.346 (n.s)
	R10 17.3	1	18	1.000 (n.s)
	R10 19.3	>1	17	0.637 (n.s)
pCGlb-1::BP178_KDEL	R11 2.1	1	20	0.346 (n.s)
	R11 4	1	20	0.346 (n.s)
	R11 7.2	1	19	0.637 (n.s)
	R11 8.1	>1	23	0.018 (s)
	R11 9	1	19	0.637 (n.s)
	R11 10	>1	0	0.000 (s)
	R11 20.3	>1	0	0.000 (s)
	R11 20.6	>1	0	0.000 (s)
	R11 20.7	1	19	0.637 (n.s)

Hyg^R , hygromycin-resistant phenotype; (a), theoretical segregation ratio for a single independently integrated insert; n.s, statistically not significant; s, statistically significant; n.s*, Although χ^2 test was not significant for a 3:1 Mendelian segregation, significant differences were assumed since probability value was very closed to 0.05. EV, empty vector line.

Table 4.21 Relationship between the copy number of the selectable marker gene (*hptII*) in T0 regenerants and segregation analysis in T1 seed progeny (24 seeds were assayed)

qPCR *hptII*		hptII segregation in T1 seeds (3:1)	
		χ^2 test	
Number of copies	**Frequency**	**n.s**	**s**
<1	2	1	1
1	22	22	0
>1	12	4	8
Total	36	27	9

n.s, not statistically significant; s, statistically significant

The chi-square analysis of *hptII* segregation patterns was non-significant in 27 out of the 36 tested lines, indicating no differences between the expected and observed frequencies for a segregation ratio of 3:1 for *hptII* transgene in T1 seeds. In fact, 100% of the T0 tested plants containing a single copy *hptII* transgene (as determined by qPCR) presented a *hptII* segregation

corresponding to 3:1 Mendelian ratio, suggesting that the results obtained by qPCR were in concordance to those obtained by segregation analysis. However, in the multiple copies group, two segregation patterns were observed. Specifically, 33% of total group presented a segregation pattern of 3:1, indicating that *hptII* inserts were integrated into the same locus (either in tandem or split) since they were segregated as a single insert. By contrast, 66% of events contained multiple copies of the *hptII* transgene presented a deviation regarding to typical Mendelian segregation for a single dominant gene. So, multiple copies were inserted into independent loci of the rice genome. If more than one *hptII* gene has been integrated into the rice genome, then truncated *hptII* gene might be among them. Also, experimental errors during transgene copy number estimation should not be discarded.

A correlation analysis was performed in order to establish if there was a significant dependency between results obtained by qPCR and segregation analysis. As it shown in shown in **Table 4.22**, a significant correlation was observed ($p<0.05$).

Table 4.22 Comparison between assessment of *hptII* copy number in T0 plants by qPCR and segregation analysis

Segregation[a]	**qPCR**		
	<1	**1**	**>1**
<1	1	0	0
1	1	22	4
>1	0	0	8

(a) A segregation ratio of 3:1 was expected when a single transgene (1 copy) or tandem transgene (>1 copy) was integrated into the same locus of the rice genome

Homozygous T2 seeds obtained by self-pollinization of T1 plants were selected based on *hptII* gene expression, and, as expected, the T2 progeny of the single copy group was composed of segregants that expressed the selectable gene in 100% of the progeny and segregants that showed the 3:1 Mendelian ratio. Thus, homozygous lines for the *hptII* transgene were identified in the T2 lines **(Fig. 4.26)**.

Then, T3 seeds were obtained by self-pollinating of T2 homozygous plants and the expression of *hptII* was confirmed in 100% of the progeny **(Fig. 4.27)**.

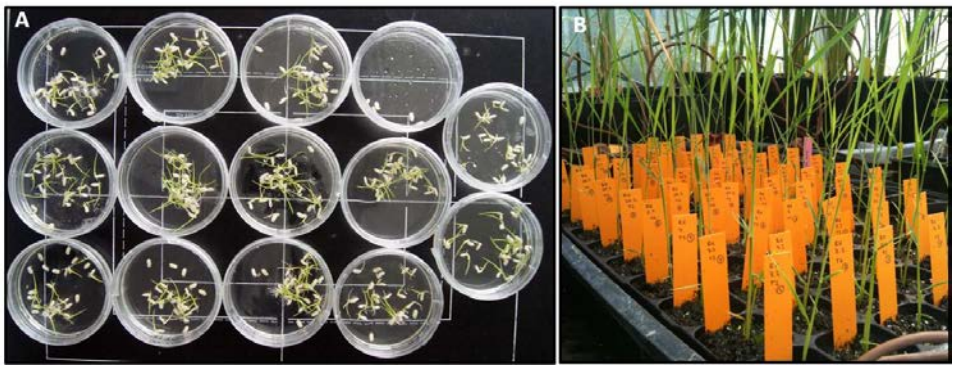


Fig. 4.26 Progeny segregation analysis. (A) Segregation analysis of the self-pollinated T2 progeny of transgenic rice plants bearing single copy for the *hptII* and *Ole18_PRS_Cec A* transgenes (R9, line 12.1). 24 seeds of each 13 segregants of T2 transgenic progeny were pregerminated and grown in the presence of hygromycin (50 mg/mL). The wild-type seeds were grown under the same conditions than the transgenic seeds. Seeds were grown at 28 °C for seven to ten days using a 12 h/12 h (day/night) photoperiod under a photon flux of 110-150 µmol. m^{-2}. s^{-1}. Depending on the segregation pattern, two groups of segregants could be distinguished. One group expressed *hptII* in 100% of the progeny, while the other group showed a segregation of 3:1. Wild-type seeds were unable to germinate. (B) Nine T2 seedlings being part of one of the group expressing the *hptII* transgene in 100% of the progeny were planted in substrate to obtain T3 homozygous lines for the gene marker *hptII* and for the transgene.

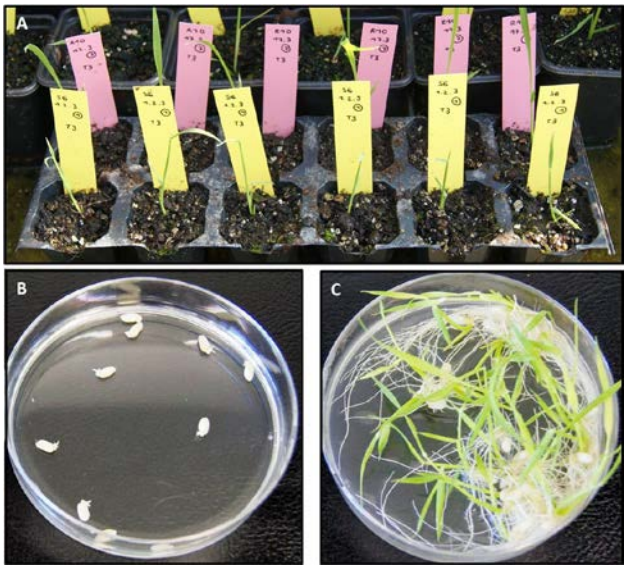


Fig. 4.27 Confirmation of *hptII* expression in 100% of the homozygous T3 progeny of *pCGlb::BP192_KDEL* lines (S6, line 1.2.3). (A) T2 transgenic plants growing in the greenhouse to obtain the T3 progeny (self-pollination). Wild-type and transgenic (S6 1.2.3 T3) seeds germinated in the presence of hygromycin (50 mg/mL) (B and C, respectively). After incubation at 28 °C for seven to ten days, at a 12 h/12 h (day/night) photoperiod and a photon flux of 110-150 µmol. m^{-2}. s^{-1}. 100% of germination was observed for transgenic T3 line. Wild-type seeds were unable to germinate.

As expected, Mendelian segregation of 3:1 was not observed in those events presenting more than 1 copy for the transgene inserted into independent loci. Moreover, T1 plants were not obtained from most of the multiple copy lines such as the lines harbouring the *GluB-1::Cec A_KDEL* gene (MC3 1.2 and MC3 1.4), the *GluB-4::Cec A* gene *(*R10 9.1, R10 12.1 and R10 12.2) or the *Glb-1::BP178_KDEL* gene (R11 10, R11 20.3 and R11 20). Finally, T1 plants were obtained from the multiple copy insertion line R9 7.2, but, T2 homozygous plants for this line were not identified in the course of this work.

In summary, at least three homozygous lines from each transformation event corresponding to fourteen different constructs were obtained. In most cases, these lines carried a single copy of both the selectable marker and the *AMPder* transgene. **Table 4.23** shows details of the transgenic lines that were selected for subsequent analyses.

Table 4.23 Homozygous lines finally selected for subsequent analysis

Construct	Cultivar	Code	Selected lines
pC1300	Senia	EV	6, 7
pCGluB-1::BP178_KDEL	Senia	S11	1.1
	Ariete	A11	1.3, 1.7, 1.9
pCGluB-1::BP188_KDEL	Ariete	A12	5.1, 5.2, 8.1, 8.2
pCGluB-1::BP192_KDEL	Senia	R4	5.1, 5.2, 5.3, 5.8
		S8	1.3, 2.3
	Ariete	MC2	5.1, 5.3, 6.2, 6.3, 6.4, 6.5, 8.1, 9.2
pCGluB-1::Cec A_KDEL	Senia	S3	2.2
	Ariete	MC3	1.3, 2.3, 2.5, 2.6
pCGluB-1::Cec A	Ariete	A5	1.2, 3.2, 5.1
pCGluB-4::BP178_KDEL	Ariete	A14	11.1, 12.1, 13.2
pCGluB-4::BP188_KDEL	Senia	R8	2.7, 2.13, 2.15, 3.1, 2.27
pCGluB-4::BP192_KDEL	Ariete	A7	4.2, 5.2, 6.2
pCGluB-4::Cec A_KDEL	Senia	S2	1.1, 2.2, 4.8
pCGluB-4::Cec A	Senia	R10	11.1, 12.3, 14.3, 17.3
	Ariete	A1	2.1, 3.1, 3.2, 5.2
pCGlb-1::BP178_KDEL	Senia	R11	2.1, 4.1, 7.2, 8.1, 9.1
pCGlb-1::BP192_KDEL	Senia	S6	1.2, 1.3, 1.4
pCOle18::Ole18_PRS_Cec A	Senia	R9	3.2, 8.1, 8.5, 11.2, 12.1, 16.2
	Ariete	A16	2.1, 3.2, 4.1, 6.2, 9.2
pCOle18::Ole18	Ariete	MCX	u.d.

EV, empty vector line; u.d., under development

4.4 Characterization and yield estimation of the recombinant peptides produced in seeds of transgenic rice plants

In the present study, seeds of the selected transgenic plants were used as a system for the development of a plant based-production of synthetic antimicrobial peptides. The results presented in the previous sections indicated that the rice expression systems developed in this work allow proper expression of the *AMP* genes under study in rice seeds. Furthermore, the transgenic rice seeds accumulated the peptides at distinct subcellular compartments, namely protein bodies or oil bodies. Additionally, the transgenic seeds showed enhanced resistance to infection by bacterial (*Dickeya* sp. 1552.10.1) and fungal pathogens (*Fusarium verticillioides*), suggesting that the *in planta* produced AMPs are biologically active.

However, when transgenic plants are used as a production platform, biomass and downstream processing steps to recovery the recombinant peptides are key factors. Thus, the challenge is to achieve high valuable products in a cost-effective manner while avoiding the need of time-consuming methods. Because losses of at least 80% of recombinant product are currently reported during purification steps (Kusnadi et al., 1997; Protein purification Handbook, Amersham Biosciences), the number of processing steps should be considered. Moreover, the final product has to be purified in a suitable condition for further applications as plant pesticide, pharma or food preservative. Taking into account the aforementioned considerations, in this work different types of chromatographic techniques were evaluated as purification methods for AMPs accumulating in rice seeds.

4.4.1 Analysis of the AMPders accumulating in the rice seed endosperm

Based on the selection criteria previously described **(Section 4.3, Results)** up to six homozygous lines derived from each transformation event were selected for further characterization. They were:

(1) plants accumulating an AMP in the rice seed endosperm (*GluB-1::Cec A_KDEL*, *GluB-4::Cec A_KDEL*, *GluB-1::Cec A, GluB-4::Cec A*, *GluB-1::BP178_KDEL*, *GluB-4::BP178_KDEL*, *Glb-1::BP178_KDEL*, *GluB-*

1::BP188_KDEL, *GluB-4::BP188_KDEL, GluB-1::BP192_KDEL*, *GluB-4::BP192_KDEL*, *Glb-1::BP192_KDEL* transgenic lines),
(2) plants accumulating an oleosin_AMP fusion protein in the rice oil body (*Ole18::Ole18_PRS_Cec A* transgenic lines).

4.4.1.1 LC-MS/MS identification of the recombinant AMPder peptides accumulating in the rice seed endosperm

In addition to immunodetection of AMPs in rice seed tissues, in this work a more detailed characterization of the AMPders produced in the rice seed endosperm was carried out by using LC-MS/MS proteomic analyses. Briefly, protein extracts from PB-enriched fractions from AMPder-lines, EV control as well as increasing amounts of AMPder synthetic peptides, were loaded onto 16.5% Tris-Tricine SDS-PAGE gels. Gel slices were excised (from the dye front to ~13 kDa), digested with trypsin, and subsequently analysed by mass spectrometry. These analyses allowed the detection of tryptic-derived peptides from BP178_KDEL, BP192_KDEL, Cec A_KDEL and Cec A **(Table 4.24)** that were unequivocally identified in all the AMPder-enriched PB fractions, but not in EV control. Specifically, these peptides were identified in those bands corresponding to ~4 and ~8 kDa polypeptides. Cec A was also detected in a gel slice at the estimated 35-45 kDa size, suggesting that Cec A might interact with some of the storage rice proteins. By contrast, no BP188_KDEL peptide was detected in either synthetic BP188_KDEL samples, or BP188_KDEL-PB fraction.

LC-MS/MS data also confirmed that the ~8 kDa immunodetected polypeptide **(Figs. 4.19 and 4.20)** corresponded to the expected recombinant peptides in seeds of the transgenic plants, namely *GluB-1::BP192_KDEL* (MC2, S8, R4), *GluB-1::Cec A_KDEL* (MC3, S3), *GluB-1::BP178_KDEL* (A11), *GluB-1::Cec A* (A5), *GluB-4::BP192_KDEL* (A7), *GluB-4::Cec A_KDEL* (S2), *GluB-4::Cec A* (R10, A1), *Glb-1::BP178_KDEL* (R11) and *Glb-1::BP192_KDEL* (S6) rice plants.

Table 4.24 Tryptic-peptide sequence tags for AMPder identification by mass spectrometry

		Precursor $[M+H]^+$		
Peptide	**Amino acid sequence**	**Sequence tag**	**$[M+H]^{2+}$ (Da)**	**Rice lines[a]**
BP192_KDEL	KKLFKKILKYLAGPAKKLFKKILKY LKDEL	YLKDEL	390.71 ± 0.02	R4, S6, S8, A7, MC2
BP178_KDEL	KKLFKKILKYLAGPAGIGKFLHSA KKDEL	YLAGPAGIGK	473.77 ± 0.02	A11, R11
Cec A_KDEL	KWKLFKKIEKVGQNIRDGIIKAG PAVAVVGQATQIAK**KDEL**	AGPAVAVVGQATQIAK	740.92 ± 0.02	S2, S3, MC3
Cec A	KWKLFKKIEKVGQNIRDGIIKAG PAVAVVGQATQIAK	AGPAVAVVGQATQIAK	740.92 ± 0.02	A1,R10, A5,

(a) Transgenic rice lines in which the specific sequence tag was detected. Seeds from the following transgenic events were analyzed: *GluB-1::BP192_KDEL* (MC2, S8, R4), *GluB-1::Cec A_KDEL* (MC3, S3), *GluB-1::BP178_KDEL* (A11), *GluB-1::Cec A* (A5), *GluB-4::BP192_KDEL* (A7), *GluB-4::Cec A_KDEL* (S2), *GluB-4::Cec A* (R10, A1), *Glb-1::BP178_KDEL* (R11) and *Glb-1::BP192_KDEL* (S6)).

The tryptic peptides from the GluB-1, GluB-4 and Glb signal peptides **(Table 4.25)** were not detected in the PB samples, suggesting a correct processing of the N-terminal signal peptide during translocation into the ER system. Thus, the higher molecular weight (compared to standard peptides) observed for the *in planta* produced AMPs in the Tris-Tricine SDS-PAGE can not be explained by the presence of the signal peptide due to a failure of N-terminal processing site of the recombinant peptides.

Table 4.25 Tryptic-peptide sequence tags containing the N-terminal signal sequence of each recombinant peptide

Signal peptide	
Code	**Sequence tag**
SPα-Glb-BP178	VVFFAAALMAAMVAISGAK
SPGLUB-1-Cec A	FSIYFCVLLLCHGSMAK
SPGLUB-1-Cec A_KDEL	FSIYFCVLLLCHGSMAK
SPGLUB-4-Cec A	LSIYFCVLLLCHGSMAK
SPGLUB-4-Cec A_KDEL	LSIYFCVLLLCHGSMAK

In bold indicate the first amino acid from AMPder peptide

In this way, LC-MS/MS confirmed the identity and correct N-terminal processing in the recombinant AMPder peptides accumulating in the PBs of the transgenic rice seeds.

LC-MS/MS identification of the recombinant AMPder peptides was carried out by E. Izquierdo and M. Rossignol (Mass Spectrometry Proteomics Platform-MSPP, Laboratoire de Protéomique Fonctionnelle, INRA, Montpellier, France) within the framework of the European Trilateral Plant KBBE project SEPSAPE (Safe and Efficient Plant Systems for Antimicrobial PEptide production.

4.4.1.2 Stability of the cecropin A peptide in germinating rice seeds

We investigated the stability of the Cec A peptide accumulating in the rice seeds during germination. Seeds of *GluB-4::Cec A* (R10) rice plants were germinated for 4 and 7 days. Before preparation of total protein extracts, the germinating seeds were dissected into embryo (including coleoptile) and endosperm **(Fig. 4.28 A)**. For comparison, protein extracts from endosperm tissues of mature seeds were also analysed (embryo of mature seeds was not analysed due to the difficulty to be excised from the mature seeds free from endosperm contamination).

The Cec A peptide in R10 seeds was detected in the endosperm of 4- and 7-day germinated seeds, suggesting that this peptide is stable in the endosperm tissue, at least during the first stages of seed germination **(Fig. 4.28 B)**. No cecropin-related degradation products were observed in protein extracts from germinating seeds**.** The estimated Cec A yield is shown in **Table 4.26**.

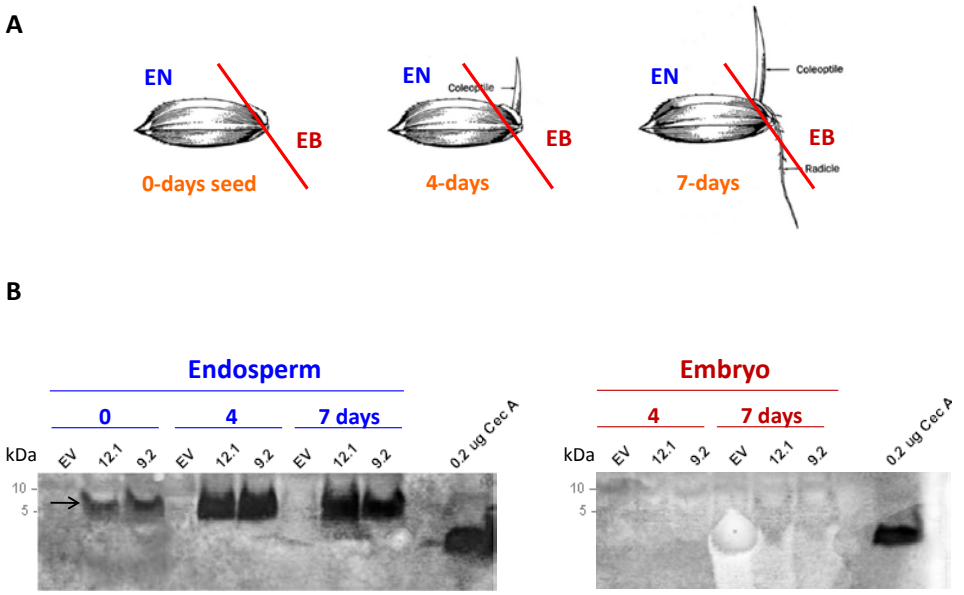


Fig. 4.28 Stability of the CecA peptide accumulating in the endosperm of germinating rice seeds (*GluB-4::Cec A*, R10, lines 9.2). The embryos (EB) and endosperm (EN) tissues were manually dissected from germinating rice seeds (4 and 7 days of germination). Total protein extracts were then obtained from embryos and endosperms as well as from mature seeds (0-days). Protein extracts (60 µg) were separated by Tris-tricine SDS-PAGE and transferred onto PVDF membrane. Membranes were incubated with the anti-Cec A antibody followed by anti-rabbit IgG (Fc) secondary antibody linked to alkaline phosphatase. EV, lines transformed with the empty pCAMBIA 1300 vector, used as a negative control. Cec A synthetic peptide was used as a positive control. Arrows, indicate Cec A-related bands in rice extracts. Numbers in the left side indicate the size of the molecular markers (kDa).

Table 4.26 Yield of the Cec A peptide accumulated in *GluB-4::Cec A* rice seeds (R10) during germination

		Yield	
Time of germination (days)	**Line**	**ng/grain**	**pmols/mg**
0	9.2	45	0.6
	12.1	65	0.8
4	9.2	110	1.4
	12.1	113	1.4
7	9.2	121	1.5
	12.1	104	1.3

The higher level of accumulation that is observed in the endosperm of germinating *GluB-4::Cec A* seeds relative to mature seeds can be explained by the difficulty in extracting the Cec A peptide from dry seeds (as compared

to germinating seeds). In addition to this, the relative contribution of storage proteins in total protein extracts (e.g. glutelins, globulins) is expected to decrease during germination (during seed germination, seed storage proteins are mobilized to provide nutrients for seedling growth). Therefore, there is the possibility that these changes in overall protein composition of germinating seeds might favour Cec A recovery from the rice seeds.

4.4.1.3 Extraction and purification of the recombinant peptides from rice seed endosperm (analytical scale)

As described in **Section 4.3 (Results)**, partial purification of recombinant peptides can be easily achieved by subcellular fractionation and protein body isolation. The purity obtained from the PB fraction isolated in this way was sufficient for protein analyses, such as Western-blot and LC-MS/MS analysis. However, in order to further characterize the recombinant product and to evaluate whether the *in planta* produced recombinant peptide can be produced at important levels for subsequent applications, (transgenic rice seeds as a biofactory), a procedure for the obtention of purified products is required.

Based on previous findings **(Section 4.3.4, Results)**, removing starch from seed proteins was the first limitation to recover high yields of the *in planta* synthesised peptides. Once most of the starch in the sample was removed and the PB fraction was obtained, the next challenge we had to overcome was the poor solubility of the protein extract after protein body disruption **(Section 3.16.2.1, Methods)**. Due to the physicochemical nature of the recombinant peptides, both structural integrity and biological activity of the recombinant products should not to be affected by the solvent used for extraction. However, the presence of the cationic recombinant peptides together with the acidic nature of storage proteins (physicochemical properties are described in the **Introduction, Table 1.5**) present in protein extracts from PBs has proven to interfere in peptide purification. In fact, purification of AMPs free of host proteins has been a major difficulty in the present study and a great effort was made to set up methods suitable for

the solubilization and purification of the recombinant peptides. In order to develop an efficient method for peptide purification, the physicochemical properties of AMPs (BP134der and Cec A peptides) were carefully examined. Here, it should be mentioned that the AMPs produced in rice seeds were initially designed considering the net charge and the hydrophobicity/amphipathicity required to maintain the antimicrobial activity of the leader peptide BP134 **(Section 4.2, Results)**. Thereby, the resulting peptides carried strongly cationic net charge and their amphipathic character was evident when representing the ideal α-helix of these peptides by means of and Edmunson wheel plot (Badosa et al., 2007).

The physicochemical properties of AMPder, such as hydrophobicity, hydrophilicity, total net charge, isoelectric point and estimated solubility, are compiled in **Table 4.27**. In this study, the online bioinformatic tools provided by Innovagen (http://www.innovagen.es/), GenScript (http://www.genscript.com/) and APD: the Antimicrobial Peptide Database (http://aps.unmc.edu/AP/main.html, Wang and Wang, 2004) were used. Additional support was provided by LIPPSO (University of Girona, Girona, Spain). Minor differences in peptide properties were observed when bioinformatic analysis of peptide sequences was carried out using different peptide property calculators.

Table 4.27 Theoretical physicochemical properties of the AMPders accumulating in rice seeds according to databases

	Peptide properties							
		Charged residues						
Peptide	Aa	# acidic	# basic	Total net charge	Hydrophilic (%)	Hydrophobic uncharged residues (%)	pI[a]	Attribute
BP178	29	2	9	7+	38	45	10.84	Basic
BP188	28	2	7	5+	32	50	10.72	Basic
BP192	30	2	11	9+	43	47	10.92	Basic
Cec A_KDEL	41	4	9	5+	32	46	10.80	Basic
Cec A	37	2	8	6+	27	49	11.18	Basic

BPders contained the KDEL retention signal in their amino acid sequence.
(a) Isoelectric point

Clearly, improper solubilisation of recombinant peptides can potentially result in a loss of the peptide during processing. Accordingly, a search for the

identification of an appropriate solvent yielding a biologically active AMP was carried out. Taking into account the peptide properties, the first step was to determine the solvent in which AMPder peptides, theoretically, are better solubilized. Theoretical considerations detailed in **Table 4.28** were assessed.

Table 4.28 Recommended guidelines for peptide solubilization according to databases

Peptide characteristics	**Peptide solubilization**
Less than 5 residues	Aqueous solution (except when all residues are hydrophobic).
Hydrophilic (>25% charged residues and <25% uncharged hydrophobic residues)	Aqueous solution when charged residues are equally distributed throughout the sequence.
Hydrophobic (50-75% uncharged hydrophobic residues)	Insoluble or partially insoluble in aqueous solution. Before aqueous solubilisation, pre-dissolved in minimal amount of stronger solvents such as acetonitrile, isopropyl alcohol, acetic acid, 4-8 M urea or DMSO.
Very hydrophobic (>75% uncharged hydrophobic residues)	Initial solubilisation in a small volume of very strong acids such as TFA or formic acid before applying the aqueous solution. Often, final solution may require high content of organic solvents or denaturing agents.
>75% of S, T, E, D, K, R, H, N, Q or Y residues	As these peptides have the tendency to aggregate and they have to be treated as hydrophobic peptides. Variations in pH values might be considered.

Wang, Z. and Wang, G. (2004) APD: the Antimicrobial Peptide Database +, Innovagen (http://www.innovagen.es/), GenScript (http://www.genscript.com/)

As a proof of concept, peptide solubility was initially tested using known amount of synthetic peptides. Peptides under study are strongly basic and contain >25% of positive charged residues and >25% of uncharged hydrophobic residues. Thus, based on the physicochemical properties described in **Table 4.27**, the AMPder could be included within the second or third group of peptides compiled in **Table 4.28**. Accordingly, the synthetic lyophilized peptides were completely solubilized in double distilled H_2O.

The next step consisted in the development of an efficient chromatographic system for AMP purification. Due to the strong cationic properties of the

AMPs, initial attempts to set-up the purification method focused on ion-exchange chromatography (followed by HPLC to check the purity of the recovered peptide). For this purpose, strong cation exchange chromatography (HyperSepSCX Strong Cation Exchanger SPE Columns) of synthetic peptides was performed following the Amersham Bioscience Protein Purification Handbook. Although a small portion of synthetic peptide was recovered (60-80% of elution buffer, 20 mM Na_2HPO_4 + 1 M NaCl, pH 6.8), neither the recovery yield nor the high salt content was good enough to carry on with the next step.

The next approach used for synthetic peptide purification was based on the isoelectric point of these peptides. Once the peptide was retained into the column sorbent, the strategy for peptide elution consisted in decreasing the positive net charge of the retained peptides by increasing the pH of the buffer. Unfortunately, no peptide was detected in any of the recovered fractions. Thereby, the synthetic peptides either remained tightly bound to the sorbent or aggregated, thus, becoming too large to enter into the column sorbent.

Knowing the difficulty to elute strong cationic peptides from a strong cation exchange sorbent, the next strategy was based on the following principles: (1) the use of Weak Cation Exchange-Reversed Phase chromatography (this methodology is suitable only for ionic species), (2) solubilisation of the synthetic peptides in saline buffer at pH 6.0 (to increase both the solubility and the positive charge of peptides, thus providing stronger interaction with the negative charges of the column sorbent), and (3) recovering of retained peptides by elution with organic solvents (better choice as they can be removed by lyophilisation). Nanodrop was used to confirm that peptides were completely solubilized in the selected solvent.

After several trials to evaluate the aforementioned factors, the synthetic peptides were found to better solubilised in aqueous 25 mM ammonium acetate, pH 6.0 buffer. Furthermore, 50% acetic acid was selected to elute the synthetic peptides from the chromatography sorbent. After optimizing the loading, washing and elution steps, the synthetic peptides were successfully purified using a Weak Cation Exchange-Reverse Phase system **(Section 3.20.1, Methods)**, adjusting it to strong cationic peptide properties.

Next, the recovered peptides were analysed and semiquantified by HPLC chromatography. The yield of peptide recovery after purification was ~60% relative to the initial amount of peptide. No further peptide was eluted when the cation exchange resin was washed with 100% of acetic acid, formic acid, acetonitrile or NaCl 2 M.

Once the procedure for peptide purification was established, the transgenic rice seeds were used as the starting material for peptide purification. Total protein extracts from protein body fractions were obtained by precipitation using acetone+TCA+DTT. Surprisingly, while CecAder was able to precipitate together with seed storage proteins, the BP178_KDEL was found to soluble in these conditions. Taking advantage of BP178_KDEL solubility in acetone+TCA+DTT, the precipitated seed storage proteins were easily removed from the protein extracts by centrifugation, and the acetone fraction was lyophilized **(Fig. 4.29)**. In this way, a single purification step was required to obtain a high quality BP178 peptide from PB enriched fraction.

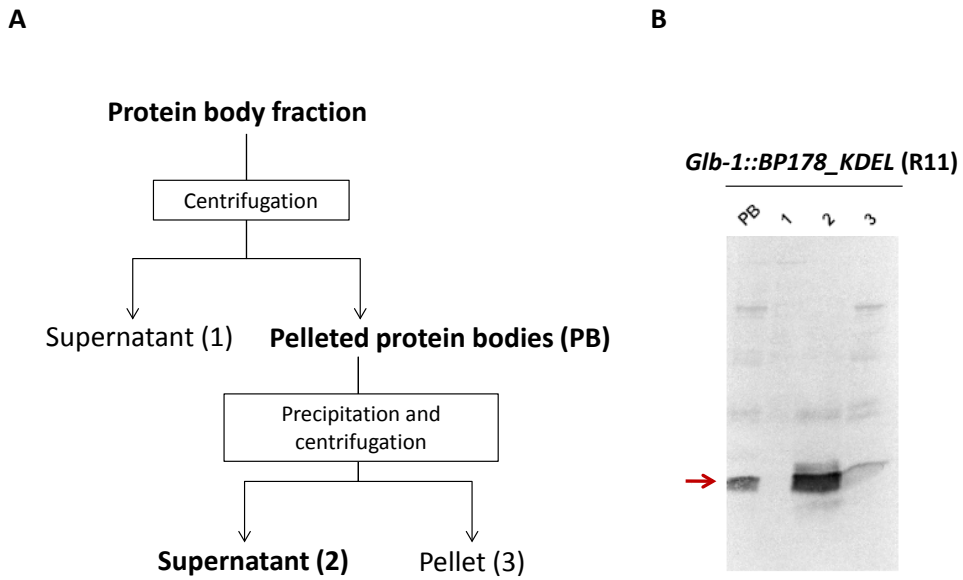


Fig. 4.29 Purification of BP178_KDEL peptide from protein bodies by acetone solubilization. Protein body fraction was obtained from *pCGlb::BP178_KDEL* (R11, line 4.1.4) homozygous plants. (A) Procedure followed for purification of BP178 peptide. Protein body fraction (50 µg) was centrifuged to recover the pelleted PB (PB). Supernatant was also analysed (1) to confirm the absence of BP178 peptide in this fraction. Precipitation mixture consisting of acetone+TCA+DTT was added to pelleted PB, followed by one-step centrifugation. Acetone supernatant (2) contained the purified BP178 peptide, while seed storage proteins were recovered in the pellet fraction (3). (B) Tris-Tricine SDS-PAGE and Western blot analysis of protein body fractions recovered during the process of purification of BP178 shown in A. PVDF membrane was incubated with the anti–BP178 antibody followed by anti-rabbit IgG (Fc) secondary antibody linked to alkaline phosphatase. Arrow indicates the BP178-related band in the rice sample.

In the present study, efforts were also made to obtain a partially purified Cec A peptide from protein body fractions of transgenic seeds. To remove the acetone remaining in the protein precipitate, CecAder-fraction extract was lyophilized and washed by dissolving it in double distilled H_2O, followed by another freeze-drying step.

Knowing that the synthetic Cec A peptides solubilize in 25 mM ammonium acetate, pH 6.0 and that Cec A peptides were found to be successfully recovered from cation exchange chromatography (50% of acetic acid as an elution solvent), the procedure for purification of Cec Ader from rice seeds was as follows. Initially, the protein mixture (seed storage proteins and

CecAder products) was resuspended in 25 mM ammonium acetate, pH 6.0. Unfortunately, although synthetic peptides were completely soluble in the aqueous buffer, the protein extracts obtained from protein bodies were insoluble. Based on the theoretical considerations to solubilise protein aggregates, the use of acetonitrile, isopropyl alcohol, acetic acid, 8 M urea, SDS and Triton X-100 (among others) was the next choice. To this point, it should be taken into account that these agents must then be removed if the obtention of biologically active recombinant peptides is pursued (e.g. if antimicrobial activity of the recombinant product). Generally, these agents are removed by chromatography methods but these methods are not suitable for our recombinant peptides due to their sticky nature (strongly cationic and amphyphatic). Equally, denaturing agents are not recommended.

Several saline buffers and organic solvents were tested for their ability to solubilize rice seed protein aggregates to be suitable for subsequent liquid chromatography purification. When the protein extract was solubilized using acetic acid, the AMP peptide precipitated together with storage proteins (probably by interacting with glutelins) and no CecAder peptide remained in the supernatant fraction. Similar results were obtained when using acetonitrile or saline buffers at different pH ranges. Finally, considering that the protein extract was highly insoluble, the lyophilized product was initially resuspended in a small volume of TFA before applying double distilled H_2O (0.1% TFA final solution). Although complete solubilization of the protein mixture was not achieved, the protein sample was found to be suitable for subsequent liquid chromatography purification.

The different steps followed for the obtention of Cec A from rice protein bodies are summarized in **Fig. 4.30 A**. The CecAder was partially purified from seed protein extract by using the Weak Cation Exchange-Reverse Phase column following the protocol designed for synthetic AMPders. Fractions recovered were lyophilized, washed (by dissolving it in double distilled H_2O) and freeze-dried. After resuspending this material in TT1X buffer, all fractions were checked by Western blot analysis **(Fig. 4.30 B)**. Previous results have shown the complete elution of synthetic AMPder in the acetic acid fraction, so, CecAder was expected to be eluted in the same fraction.

Surprisingly, CecAder from transgenic seeds eluted in the wash fraction consisting of 100% methanol. Thus, Cec A produced in transgenic seeds showed an unexpected high solubility in methanol, compared to that of the corresponding synthetic peptide, indicating differential properties from those of the synthetic peptide.

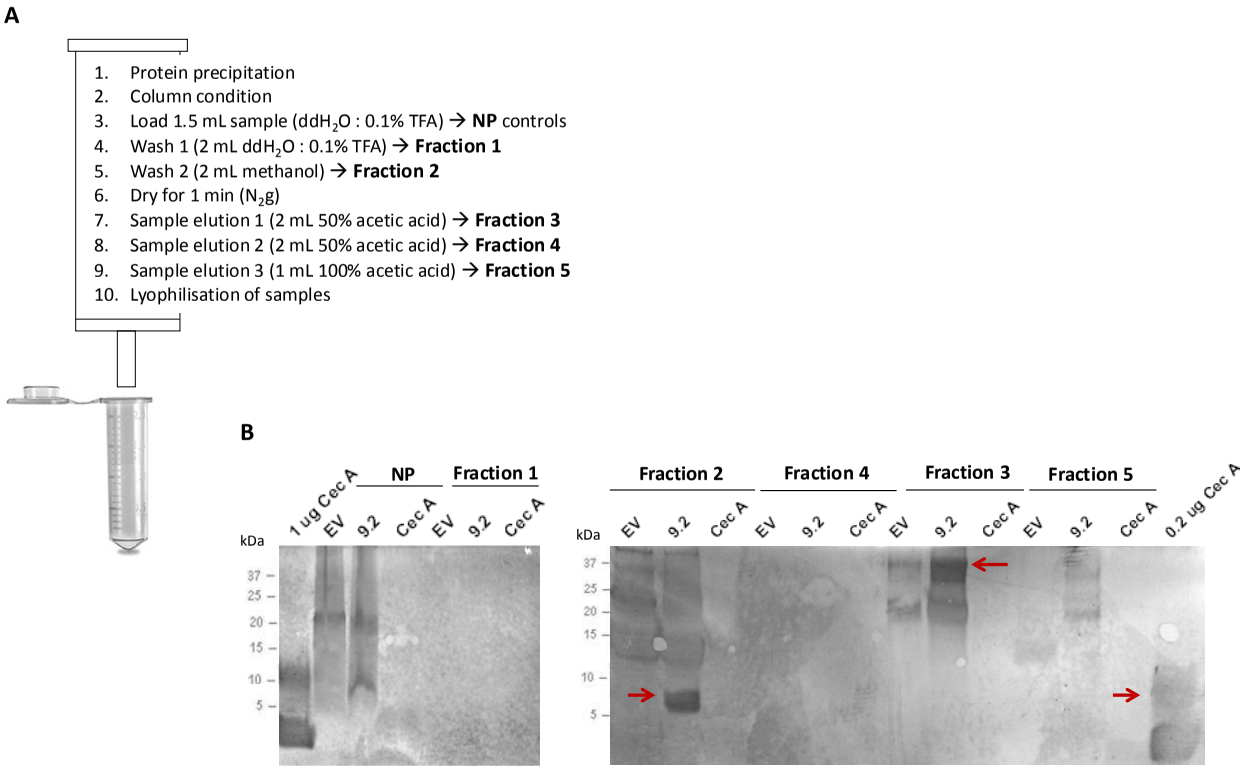


Fig. 4.30 Purification of the Cec A peptide from seeds of *GluB-4::Cec A* rice plants (R10, line 9.2). Proteins obtained from protein body preparations (50 µg) and synthetic Cec A (15 µg) were dissolved in 0.1% TFA : H_2O and separated on a Weak Cation Exchange-Reverse Phase column. (A) Protocol designed to elute the Cec A peptide from *GluB-4::Cec A* enriched-protein body fraction. (B) Each recovered fraction was analyzed by western blot using anti-Cec A antibodies. EV, PB extracts from empty pCAMBIA 1300 vector line, used as a negative control; samples NP, 0.1% TFA:H_2O - non soluble proteins in the PB sample. 1 µg of synthetic Cec A was used as a positive control. Arrows indicate Cec A-related bands. Numbers in the left side indicate the molecular weight protein markers (kDa).

The observed high solubility of the Cec A product in methanol suggests changes in peptide polarity, based on the functional groups of the amino acids contained in the peptide sequence. Moreover, as described in **Section 4.3 (Results)**, the CecAder eluted with methanol had an apparent molecular weight close to 8 kDa. Interestingly, when 50% of acetic acid was used after methanol elution, an immunoreactive polypeptide of ~37 kDa was detected with the anti-Cec A antibody **(Fig. 4.30 B, fraction 3)**. This result was in agreement with that obtained by LC/LC Mass spectrometry, where Cec A was identified in a gel slice band at 35-45 kDa **(Section 4.4.1.1, Results)**. Then, the hypothesis that Cec A was interacting with a host protein gained more strength. If so, the denaturing electrophoretic conditions did not dissociate the possible protein-peptide interaction. It is also noteworthy that both, glutelins and AMPder were soluble in acetic acid. There is then the possibility that Cec A and glutelins might interact. However, further studies are needed to clarify this aspect.

To study in more detail the observed affinity changes, the washing step with methanol was eliminated from the purification protocol **(Fig. 4.31 A)**. Interestingly, CecAder was eluted in the 50% acetic acid fraction (Fraction 1) and no peptide was eluted in subsequent washes. In addition, three Cec A-related bands were immunodetected with apparent weights of 12, 8 and 4 kDa (the expected monomer molecular weight is ~4 kDa) **(Fig. 4.31 B, fraction 1)**. Moreover, when 50 % of acetic acid was used as elution solvent, a ~20 kDa polypeptide detected with the anti-Cec A antibody **(Fig. 4.31 B, fraction 1)**. Once again, these results support the hypothesis that Cec A might interact with acid-soluble host proteins.

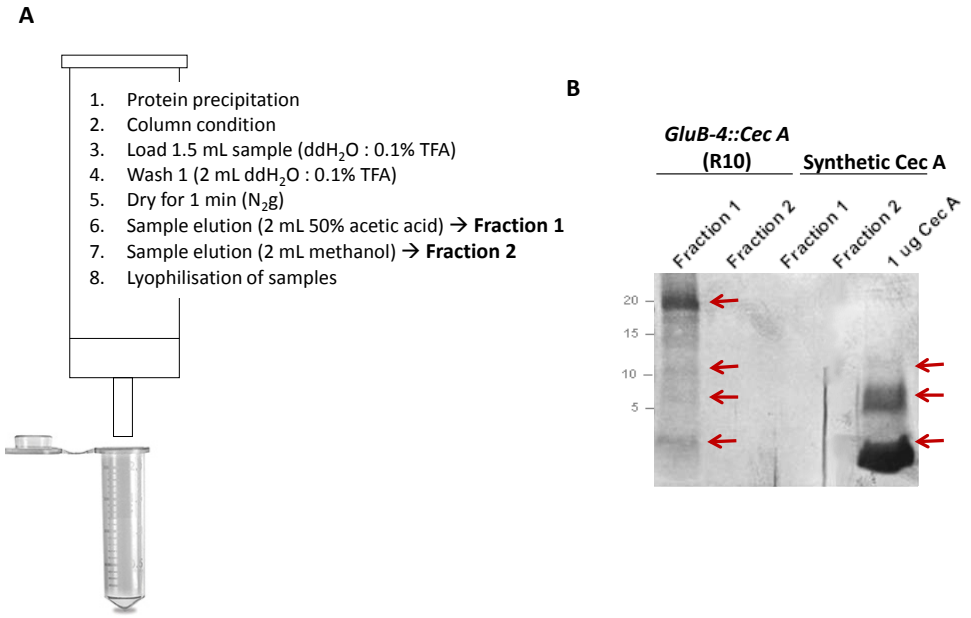


Fig. 4.31 Modified method for purification of Cec A peptide from seeds of *GluB-4::Cec A* rice plants. Protein body extracts (50 µg) were dissolved in 0.1% TFA:H_2O and separated on a Weak Cation Exchange-Reverse Phase column. Each recovered fraction was analysed by western blot using anti-Cec A antibodies. (A) Modified protocol designed to elute Cec A peptide from *GluB-4::Cec A* (R10, line 9.2) enriched-protein body fraction. B) PB protein extracts from seeds transformed with *pCGluB-4::Cec A* (R10, line 9.2) eluted with 50% acetic (Fraction 2). Three Cec A-related bands were estimated at ~4, 8 and 12 kDa size in fraction 2. Synthetic Cec A (15 µg) appears to strongly interact with the negative charges of the resin. 1 µg of synthetic Cec A used as a positive control. Arrows indicate Cec A-related bands. Numbers in the left side indicate the molecular weight of the protein markers (kDa).

The observation that no synthetic peptide was recovered from the column when eluting with either methanol or 50% acetic (up to 15 µg of Cec A was loaded onto the column), supports that the synthetic Cec A strongly interacts with the negative charges of the sorbent (Cec A was found to be soluble in 0.1% TFA solvent, the mobile phase) **(Figs. 4.30 and 4.31, Lanes Cec A)**.

During the course of this work, the possibility of posttranslational modifications in the rice-produced Cec A was also considered. Thus, the apparent high molecular mass of AMPs (as revealed by SDS-PAGE) might be associated to glycosylation of the in planta- produced peptide (and not to n-

mers formation). Clearly, the attachment of carbohydrates to the recombinant AMP might interfere not only with its solubility and stability but also with its antimicrobial activity.

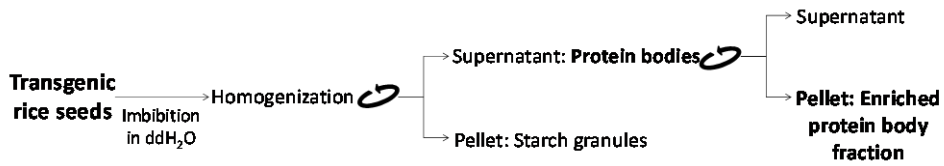
It is well known that the glycosylation of proteins has a great impact both on their physicochemical properties and on their biological function. Carbohydrates attached to proteins can be classified into two main categories: N-glycans (linked to the amide group of asparagine residues), and O-glycans (linked to the hydroxyl group of serine, threonine, hydroxylysine or hydroxyproline (Hyp) residues) in the polypeptide chain. With the exception of O-GlcNAcylation, protein N- and O-glycosylations typically occur in the secretory pathway. Thus, protein N-glycosylation starts in the endoplasmic reticulum, with the cotranslational or post-translational transfer of preformed lipid-linked oligosaccharide onto the nascent polypeptide. In contrast, O-glycans are synthesized by a stepwise transfer of monosaccharides on the folded protein in the ER and the Golgi apparatus (Gomord et al., 2010). Evidence in the literature supports O-glycosylation of rice seed prolamines (Kilcoyne et al., 2009) and glutelins (Kishimoto et al., 1999). Protein glycation (non-enzymatic reaction of proteins with glucose and other reducing sugars, predominantly at lysine residues according to the NetGlycate 1.0 Server) might also lead to structural and functional modifications of plant proteins. Accordingly, the possibility of glycosylation and glycation posttranslational modifications of AMPders was explored.

The O-glycosylation and glycation potential sites of AMPs were predicted by using the YinOYang 1.2 Server and the NetGlycate 1.0 Server, respectively. The AMPder amino acid sequences (BPs and Cec A) contain threonine and proline residues that are potential O-glycosylation sites as well as several potential glycation sites (lysine). No potential N-linked sites (Asn-X-Ser/Thr) were identified in the synthetic AMPder amino acid sequence. For this, a protein Mix Deglycosylation System (combination of enzymes for deglycosylation of both, N and O-deglycosylated residues) was used. The results obtained in this study were, however, not conclusive. Thus, in our hands, changes on BP178_KDEL, Cec A or Cec A_KDEL electrophoretic mobility were not observed after endoglycosidase H or Peptide-N-Glycosidase F (PNGase F) treatment compared to the untreated controls.

Another glycoprotein detection system based on the biotin-labelling of a carbohydrate, followed by its immunodetection (streptavidin-horseradish peroxidase) was also used in the present work, and again unclear results were obtained. Although not completely proven, these observations indicate that N-linked glycosylation, most probably, does not occur in the seed-produced AMP peptides, although there is still the possibility that the enzymes used in this study were not able to deglycosylate glycans linked to the AMPders. Based on the above results, no evidences on posttranslational modifications (glycosylation, glycation) of the *in planta* produced peptide was obtained in this work. Then, further studies are needed to elucidate if the higher molecular weight mass of AMPs in Tris-Tricine SDS-PAGE gels reflects peptide multimerization (e.g. dimerization), abnormal behaviour of the plant-made AMPder in this electrophoretic system, or post-transcriptional modifications of the recombinant peptide products.

Finally, once the procedure for isolation of AMPders from endosperm tissues was established at the analytical level, the next step was to evaluate the feasibility of using this method at a preparative scale and to assess whether the recovered AMPs are biologically active. Towards this end, seeds from transgenic plants harbouring the *Glb-1::BP178_KDEL* construct (R11, homozygous line 7.2) and seeds from empty vector control plants (EV, line R1II 7) were selected and used for seed amplification (T4 homozygous seeds).
For this purpose, 120 T3 seeds from each line were grown in the greenhouse until maturity. PB fractions were obtained from 1.500 seeds. Protein extracts were then prepared from protein bodies for each transgenic event (*Glb-1::BP178_KDEL* and control plants (~245 mg of proteins were obtained in each case) **(Fig. 4.32)**.

A



B

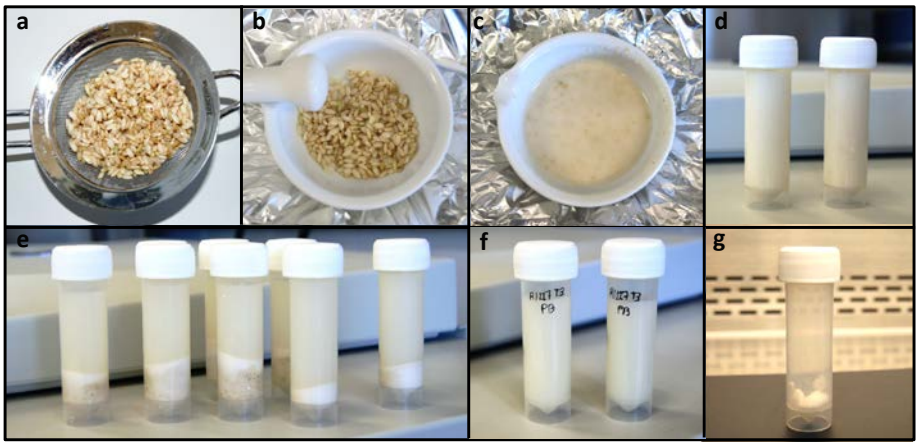


Fig. 4.32 Preparation of protein extracts from PB (*pCGlb::BP178_KDEL::Nos-T*, R11) and EV control lines at preparative scale. (A) Procedure followed for the preparation of protein extracts from protein bodies of transgenic rice seeds. (B) Preparation of protein extracts from PB from transgenic rice seeds following the procedure shown in A. (a-c) 1.500 seeds were embedded in water and were homogenized in grinding buffer (0.6 M Sucrose, 10 mM Na_2HPO_4, pH 7.5). (d) The homogenate was centrifuged (e-f). The pellet consisted of most starch granules and the supernatant was enriched with protein bodies (f). The supernatant was centrifuged, obtaining a pellet composed of enriched protein body fraction. (g) After precipitation (acetone+TCA+DTT), the protein extract was resuspended in double distilled water and lyophilized. Lyophilized product represents the starting material for further applications.

This material represents the departing biotechnologically produced material for subsequent evaluation of the AMP product. Current work at the CIDSAV for macropurification of endosperm-seed derived AMP peptides is based on the use of an ÄKTA purifier system and evaluation of the antibacterial activity of AMP products against phytopathogens.

4.4.2 Analysis of the oleosin_Cec A fusion protein accumulating in the rice seed embryo

As previously mentioned, an *oleosin_Cec A* fusion gene was expressed in rice under the control of the *18 kDa oleosin* promoter. The fusion protein was correctly synthesized and accumulated in oil bodies of the seed embryo. The identity of the Cec A peptide, being part of the fusion protein, was demonstrated (western blot analysis using an anti-Cec A antibody). A more detailed characterization of the oleosin_Cec A fusion protein and a protocol for large scale production of Cec A accumulating in rice oil bodies are presented in this section.

4.4.2.1 LC-MS/MS identification of the Cec A peptide produced in rice as an oleosin_Cec A fusion protein

The identity and correctness of the oleosin_Cec A fusion protein was analysed by mass spectrometry. Briefly, samples of *pCOle18::Ole18_PRS_Cec A* (R9) lines, EV control (as well as increasing amounts of Cec A synthetic peptide) were digested with trypsin and subsequently analysed by LC-MS/MS. The peptide sequence AGPAVAVVGQATQIAK (monoisotopic mass 740.92), this sequence being part of the Cec A polypeptide, KWKLFKKIEKVGQNIRDGIIKAGPAVAVVGQATQIAK) was unequivocally identified in all the protein samples from Ole18_Cec A seeds tested as well as in Cec A synthetic controls, but not in protein samples from empty vector control seeds. Thus, Western-blot analysis in combination with LC-MS/MS confirmed the identity and the correct synthesis of the Cec A peptide produced using fusion technology.

LC-MS/MS identification of the recombinant AMPder peptides was carried out by Esther Izquierdo (Mass Spectrometry Proteomics Platform-MSPP, Laboratoire de Protéomique Fonctionnelle, INRA, Montpellier, France) within the framework of the European Trilateral Plant KBBE project SEPSAPE (Safe and Efficient Plant Systems for Antimicrobial PEptide production.

4.4.2.2 Purification of Cec A from the oleosin_Cec A fusion protein

In order to optimize the proteolytic digestion of the oleosin_Cec A fusion protein, a synthetic peptide (named as BP251) was prepared (LIPPSO, Universitat of Girona, Spain). The BP251 amino acid sequence included the processing site for the TEV protease (KKLFKKILKYL**PTTENLYFQS**AVAVVGQATQIAK). Three different commercial preparations of TEV protease were used in these studies. Although the proteolytic cleavage efficiency greatly varied, the best results were obtained by using the ProTEV plus purchased from Promega.

Next, oil body proteins from transgenic and non transgenic seeds were treated with the TEV protease. A polypeptide of ~4 kDa was immunodetected in TEV-digested protein samples from transgenic lines, suggesting that Cec A was released from the fusion protein (data not shown). As it was observed with the standard peptide BP251, results were variable in terms of the efficiency of the proteolytic digestion. Moreover, there were cases in which the Cec A was cleaved from the fusion protein (as determined by Western blot analysis using the anti-oleosin antibody), but free Cec A was not always immunodetected. In this case, digestion mixtures were centrifuged and Cec A was detected in the infranatant fraction (as revealed by LC-MS/MS analysis of these samples). This observation supports cleavage of the Cec A peptide from the fusion protein.

In a further step to optimize conditions for the proteolytic cleavage of the oleosin_Cec A protein, the fusion protein was purified by SDS-PAGE and then incubated with the TEV protease **(Fig. 4.33)**. Results obtained using the purified fusion protein were essentially the same than those obtained using the total proteins from oil bodies for TEV protease digestion.

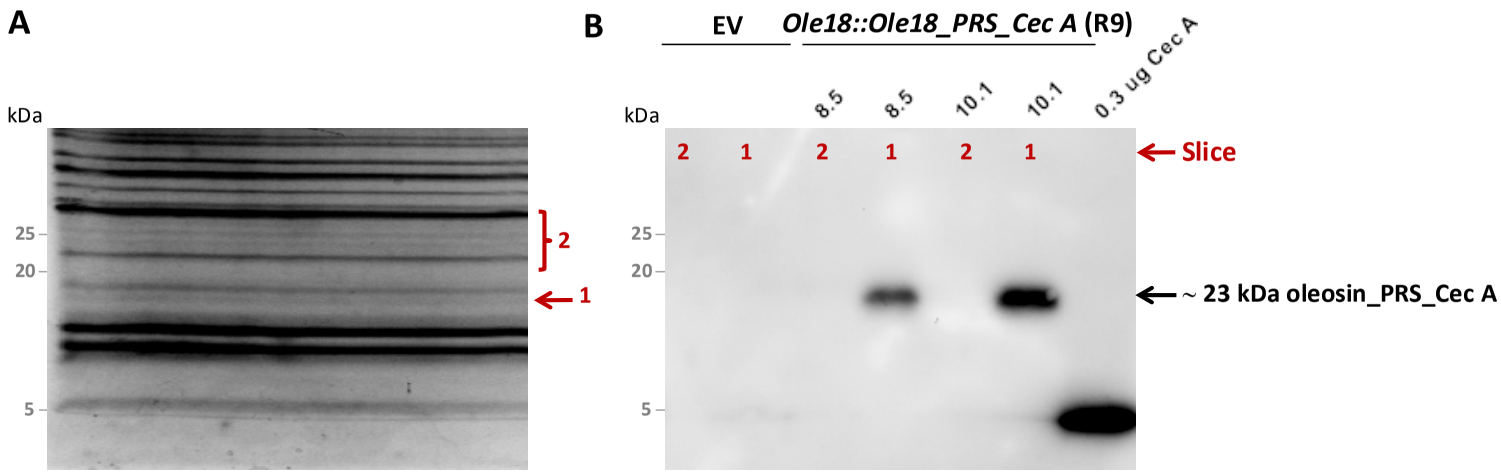


Fig. 4.33 Purification of oleosin_Cec A fusion protein from *pOle18::Ole18_PRS_Cec A* (R9) oil bodies. SDS-PAGE of OB proteins from *pCOle::Ole18_PRS_Cec A* seeds (R9, lines 8.5 and 10.1). (A) Gels were stained (colloidal Coomassie staining), and proteins at the indicated positions (1 and 2) recovered. (B) The purified proteins were subjected to immunoblot analysis using the anti-Cec A antibody. A reactive band against anti-Cec A antibody was detected in section 1 of the oil body proteins from both transgenic lines, while it was not detected in gel section 2. No immunoreactive band was detected in any gel sections from EV seeds.

The next step focused on purification of the Cec A peptide after TEV digestion. It should be here noticed that TEV digestion of the oleosin_Cec A fusion protein releases the Cec A polypeptide in which an extra serine residue is attached to its N-terminal sequence (S-Cec A). The S-Cec A peptide was purified using the Weak Cation Exchange-Reversed phase system, previously used for the purification of the Cec Ader produced in the rice seed endosperm **(Section 4.4.1.3, Results)**. Two starting materials were used for the oleosin_Cec A fusion: (1) OB fraction and (2) OB protein extracts **(Sections 3.16.1.2 and 3.16.2.2, Methods)**, and from both *Ole18::Ole18_PRS_Cec A* transgenic seeds and control seeds (empty vector). Every sample was diluted in 0.1% TFA before loading into the column. All recovered fractions were lyophilized, washed by adding double distilled water and again freeze-dried. Lyophilized samples were resuspended in TT1X loading buffer and subjected to Western blot analysis using anti-Cec A antibodies **(Fig. 4.34 B, C)**. The oleosin_Cec A fusion protein eluted in the acetic fractions **(Fig. 4.34 B, fractions F3 (50% acetic) and F5 (100% acetic))**. After TEV digestion, the S-Cec A product was exclusively recovered in the fraction eluted with 100% of acetic acid **(Fig. 4.34 C, fraction F5)**. In fact, S-Cec A was eluted together with the fraction of undigested fusion protein, indicating either the partial proteolysis of the fusion protein or interaction of the S-Cec A with the oleosin_Cec A fusion protein.

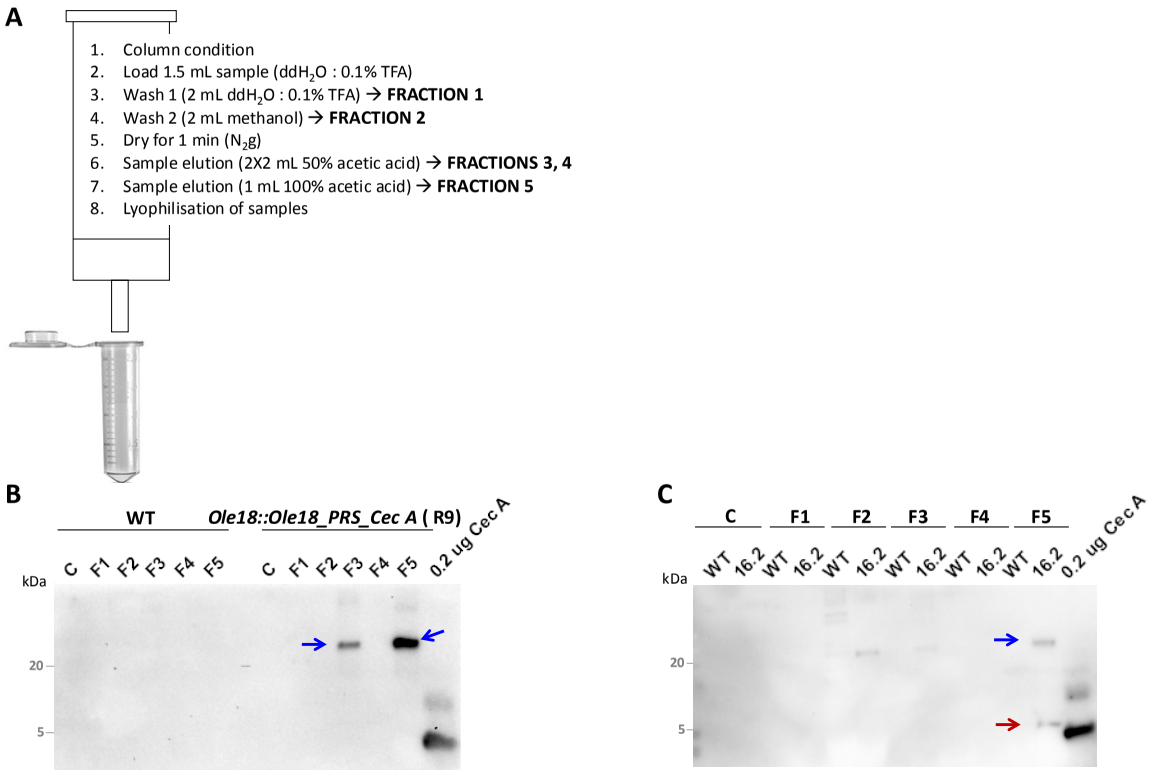
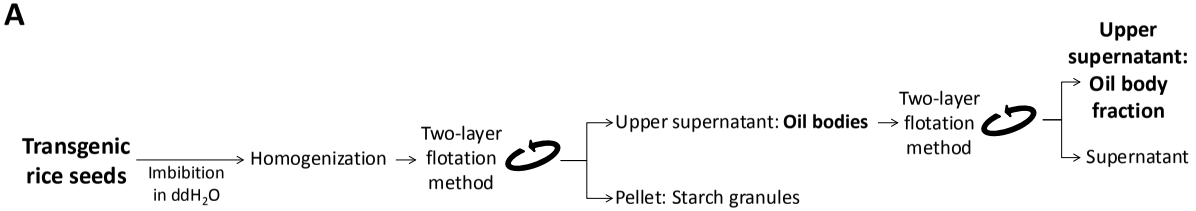


Fig. 4.34 Purification of the oleosin_Cec A fusion protein from *Ole18::Ole18_PRS_Cec A* seeds (R9, line 16.2) and isolation of S-Cec A by TEV digestion of oleosin_PRS_Cec A. Oil body proteins (35 µg) were dissolved in 0.1% TFA:H_2O and separated on a Weak Cation Exchange-Reverse Phase column. Each recovered fraction was analysed by western blot using anti-Cec A antibodies. (A) Protocol designed to elute the fusion protein and the digested S-Cec A peptide from *Ole18::Ole_PRS_Cec A* (R9, line 16.2) oil body fraction. B) Analysis of protein extracts from oil bodies (after OB disruption and protein precipitation). (C) Elution of the S-Cec A peptide after TEV digestion. TEV digestion was done using crude proteins from oil bodies. WT, from non transformed seeds. Synthetic Cec A used as a positive control (0.2 µg).

4.4.2.3 Purification of the S-Cec A at a preparative scale and bioactivity against bacterial pathogens

Plants expressing the *oleosin_Cec A* fusion gene (*pCOle18::Ole18_PRS_Cec A*, R9, line 12.1), as well as empty vector control plants, were grown in the greenhouse until maturity (120 T3 plants each). The procedure followed for purification of the oleosin_Cec A produced in rice OBs is illustrated in **Fig. 4.35**. Briefly, oil bodies were purified transgenic seeds (1.500 seeds) as described in **Sections 3.16.1.2 and 2.16.2.2 (Methods)**. Proteins from the oil body fraction were then obtained (~245 mg of proteins for each transgenic event).

A



B

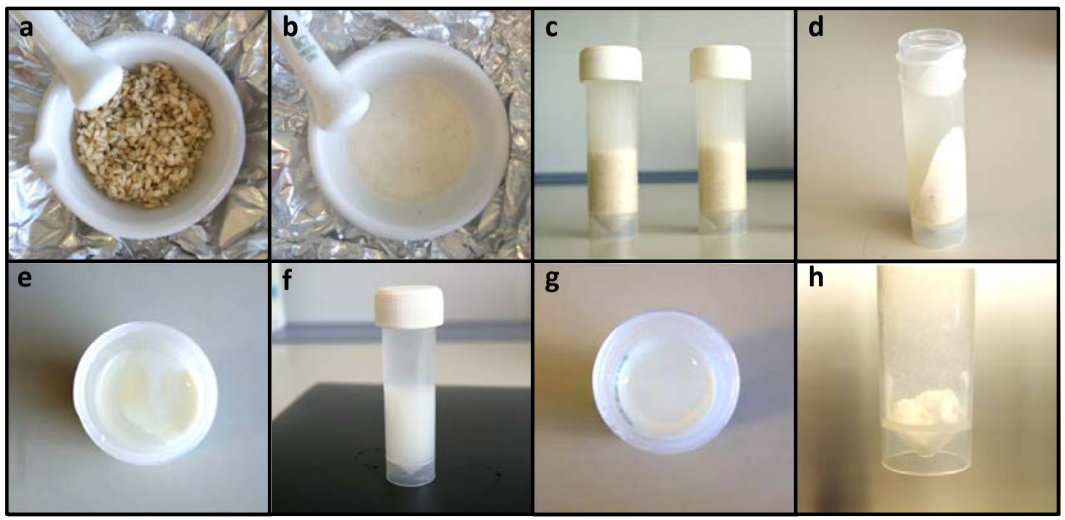


Fig. 4.35 Preparation of protein extracts from OB (*Ole18::Ole18_PRS_Cec A*, R9) and EV control lines at preparative scale. (a-b) 1.500 seeds were embedded in water and were homogenized in grinding buffer (0.6 M Sucrose, 10 mM Na_2HPO_4, pH 7.5). (c-d) The oil bodies were isolated from homogenate components by floating centrifugation using sucrose gradient density. (e) The immiscible "fat pad" in the upper supernatant was recovered and dissolved. (f-g) A second density gradient was performed and the OB fraction was obtained. (h) After precipitation, the oil body extract was resuspended in double distilled water lyophilized. Lyophilized product represents the starting material for further applications.

Cleavage of oil body proteins with the TEV protease was carried out (10 U TEV/20 µg of oil body protein extract, in ProTEV buffer, 2 h at 30 °C) and confirmed by Western-Blot analyses.

To estimate the amount of the recombinant peptide present in the protein extracts from oil bodies, increasing concentrations of synthetic Cec A (from 0.097 µM to 1.56 µM) were prepared to obtain the calibration curve.

Next, products from the proteolytic digestion (as well as synthetic Cec A peptide) were tested for antibacterial activity. The antimicrobial activity of the synthetic Cec A peptide (at known concentrations) was compared to that of the released S-Cec A, under similar assay conditions. A contact exposure test was performed with the plant-pathogenic bacteria *Dickeya* sp. by incubating the bacterial suspension (10^4 cfu/100 µL of total volume) with 69 or 138 µg of OB proteins obtained from either *pCOle::Ole_PRS_CecA* seeds (R9) or empty vector seeds (EV). Bacterial survival was determined at 2 h after exposure to the Cec A products. Bactericidal activity of the digested oil body protein extracts was determined after lyophilisation as well as with extracts in which the lyophilisation step was omitted (oil body proteins were prepared as shown in **Fig. 4.35,** and after proteolytic digestion, the samples were subjected or not to a lyophilisation step). As it is shown in **Fig. 4.36**, TEV cleavage of the recombinant fusion protein yielded a S-Cec A peptide that was biologically active. This observation suggests that the S-Cec A released from the oleosin_Cec A fusion protein is properly folded and adopts an active conformation that allows the peptide to interact with the bacterial membrane. The serine residue attached at the NH_2-terminal amino acid sequence of the recombinant Cec A would not interfere with Cec A antimicrobial activity. Similarly, when the serine residue was attached to the NH_2-terminal amino acid sequence of BP134der, no significant effect on bactericidal activity was observed **(Table 4.2, Section 4.1, Results,** peptides BP193, BP195, BP205, BP206, BP207, BP208, BP211, BP212, BP209, BP210**)**.

To note, when the bacterial suspension was incubated with either OB protein extract controls or EV digested-OB, bacterial colonies increased in

size, indicating that compound(s) in the protein preparation might stimulate bacterial growth.

As shown in **Fig. 4.36 A**, after exposure to 1 µM synthetic Cec A for 2 h, *Dickeya* sp. did not survive. The concentration of the S-Cec A produced *in planta* was estimated according to standard curves obtained in antibacterial activity assays of TEV-digested oil body proteins, with and without lyophilisation **(Fig. 4.36 B and C,** respectively**)**. When using non lyophilised protein extracts, no survivals were detected **(Fig. 4.36 C, arrows)**, suggesting that the assay mixture contained a concentration of S-Cec A of at least 1 µM (total volume, 100 µL). Taking into account the volumes and dilutions made during preparation of protein samples (twenty-five µL of OB protein extract in a microtiter plate with 75 µL of the bacterial suspension, **Section 3.22.1, Methods**), the concentration of the S-Cec A present in antibacterial activity assay using oil body proteins (69 µg of oil body proteins in 25 µL) was estimated to be around 4 µM in the total volume solution.

A slightly lower bactericidal activity of the S-Cec A was detected in lyophilized TEV-digested material relative to that of the corresponding non lyophilized protein samples **(Fig. 4.36 B and C)** (this samples, however, had a gelled appearance).

In summary, the S-Cec A peptide produced and accumulated in oil bodies of *Ole18::Ole18_PRS_Cec A* rice seeds showed a clear bactericidal activity against *Dickeya* sp. No activity was observed in non digested proteins from the same oil body preparations, supporting that S-Cec A peptide displayed bactericidal activity only when released from the oleosin fusion protein.

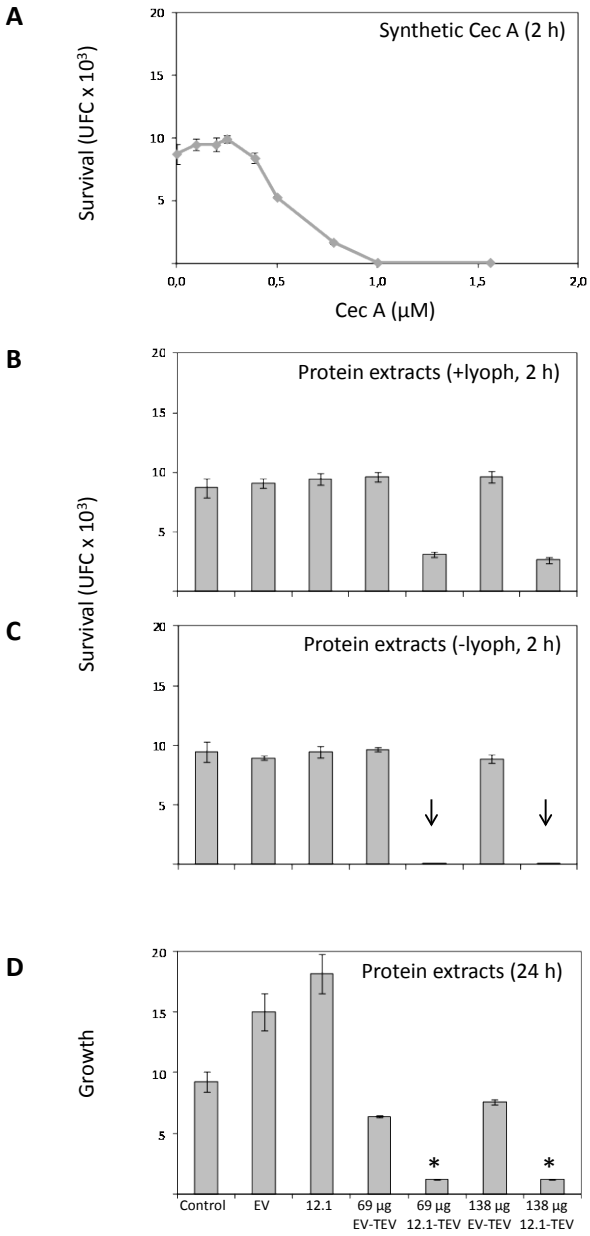


Fig. 4.36 *In vitro* bactericidal and bacteriostatic activities of the Cec A peptide produced in the oil bodies of *Ole18::Ole18_PRS_Cec A* (R9, line 12.1) rice seeds. Bactericidal activity of the synthetic Cec A (A) and the *in planta* produced peptide (B-C) was evaluated against *Dickeya* sp. (A) Bacterial survival was determined in the absence (Control) or in the presence of increasing concentrations of synthetic Cec A. (B and C) Bactericidal activity of the digested oil body protein extract upon lyophilisation (B) or without the lyophilisation step(C). Cell viability was measured after 2 h exposure period. (D) Effect of the S-Cec A peptide on the growth of *Dickeya* sp. Bacterial suspensions were incubated for 24 h. Asterisks indicate the limit of detection.

Additional experiments were conducted to further investigate the antimicrobial activity of the *in planta* produced S-Cec A peptide by assessing *in vitro* growth inhibition of *Dickeya* sp. **(Fig. 4.36 D)**. Contrary to the bactericidal activity test (contact assay), in growth inhibition test, target cells are continuously exposed to the antimicrobial peptide. For this purpose bacterial suspensions were incubated with TEV-digested oil body protein extracts from *pCOle::Ole_PRS_Cec A* (R9, line 12.1) for 24 h. OB protein extracts from R9 12.1 clearly inhibited bacterial growth (more than 90% inhibition compared to control **(Fig. 4.36 D)**. Noteworthy, bacterial growth was detected when oil body protein controls (non-digested samples) were used **(Fig. 4.36 D, EV and R9)**. This fact could be associated to a stimulatory effect on *Dickeya* growth, which results in an increased absorbance value compared to that of control extracts. This would be in agreement with the observed higher size of colonies that was observed when performing the contact exposure test (see above). While the presence of TEV protease mixture in the bactericidal assay did not affect the bacterial survival in TEV digested-OB form EV samples, digestion mixture appears to somehow stimulate the bacterial growth.

4.4.3 Yield of recombinant AMPs produced in the rice seeds

The accumulation of BP134 derivatives and Cec A peptides in transgenic seeds was estimated in T3 homozygous seeds. Results obtained for the *Glb::BP178_KDEL* (R11) plants are shown in **Fig. 4.37**. The amount of the recombinant peptides was determined by western blot analysis of protein extracts followed by quantification of immunoreactive polypeptides using the Quantity Tools (Image Lab™ Software, Version 4.1, Bio-Rad). The intensity of those immunoreactive bands was compared to that of serial dilutions of Cec A or BP134 derivatives at known concentrations. Since synthetic Cec A and BP134der show generally n-mer bands in Western-blot analysis, the amount of the standard peptide was calculated taking into account the sum of all individual immunodetected bands. This analysis was carried out with seeds accumulating an AMP in either protein or oil bodies. Representative results obtained on the analysis of seeds from *Ole18::Ole18_PRS_Cec A* (R9), *GluB-4::Cec A* (R10), *GluB-1::Cec A_KDEL* (MC3) and *Glb-1::BP178_KDEL* (R11) are presented in **Table 4.29**.

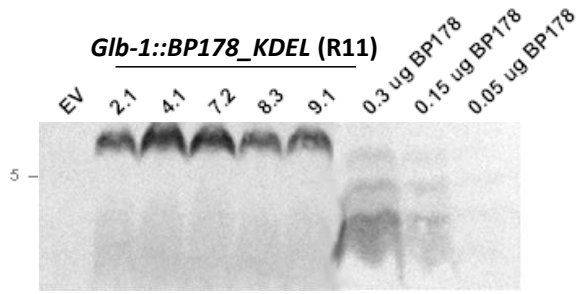


Fig. 4.37 BP178 peptide accumulates in rice seed protein bodies of *Glb-1::BP178_KDEL* (R11) transgenic seeds (T3 seeds). Protein body extracts (50 µg) were separated by Tris-Tricine SDS-PAGE gel and transferred onto PVDF membrane. Membranes were incubated with the anti-BP178 antibody followed by anti-rabbit IgG (Fc) secondary antibody linked to alkaline phosphatase. EV, rice plants with the empty pCAMBIA 1300 vector, used as a negative control. 0.3, 0.15 and 0.05 µg of BP178 synthetic peptide used as a positive control and to compare their band intensity to that of transgenic lines.

Table 4.29 Estimated yields of recombinant peptides accumulating in transgenic rice seeds

							Yield	
Peptide	**Mw**	**Localization**	**Promoter**	**Strategy**	**Event**	**Line**	**ng/grain**	**pmols/ mg of seed**
Cec A	4004.8	Embryo	*Ole18*	Oleosin fusion	R9	3.2	863	10.8
						8.1	479	6.0
						8.5	286	3.6
						11.2	317	4.0
						12.1	194	2.4
						16.2	253	3.2
						Mean	**398±197**	**5.0±2.5**
Cec A	4004.8	Endosperm	*GluB-4*	-	R10	1.1	64	0.8
						1.2	35	0.4
						9.1	103	1.3
						9.2	81	1.0
						11.1	37	0.5
						12.1	60	0.7
						12.2	92	1.2
						14.3	56	0.7
						Mean	**66±17**	**0.8±0.2**
Cec A_KDEL	4490.4	Endosperm	*GluB-1*	KDEL ER signal	MC3	1.4	129	1.4
						2.3	73	0.8
						2.6	25	0.3
						Mean	**75±59**	**0.8±0.65**
BP178	3243.0	Endosperm	*Glb-1*	KDEL ER signal	R11	2.1	85	1.3
						4.1	117	1.8
						7.2	108	1.7
						8.3	71	1.1
						9.1	76	1.2
						20.6	94	1.5
						Mean	**92±14**	**1.4±0.2**

Peptide yields were calculated by comparing the signal intensity of standard (synthetic peptides) and the *in planta*-produced peptides in western blot analysis of rice protein extracts obtained from either PBs or OBs. The confidence interval for $\alpha = 0.05$ is indicated.

The highest amount of the BP178_KDEL (3.24 kDa polypeptide), Cec A (4.00 kDa polypeptide) and Cec A_KDEL (4.49 kDa polypeptide) was found in rice seeds of the transgenic lines *Ole18::Ole18_PRS_Cec A* (R9), *GluB-4::Cec A* (R10), *GluB-1::Cec A_KDEL* (MC3) and *Glb-1::BP178_KDEL* (R11). These lines

accumulated the corresponding AMP peptide at the following levels: 10.8 pmols/mg of seed, for the *Ole18::Ole18_PRS_Cec A* lines (R9), 1.3 pmols/mg of seed, for the *GluB-4::Cec A* lines (R10), 1.4 pmols/mg of seed for the *GluB1::Cec A_KDEL* lines (MC3), and 1.8 pmols/mg of seed for the *Glb-1::BP178_KDEL* lines (R11).

Summarising the results presented in **Chapter 4**, AMPs were found to be properly synthesized and accumulated in a stable manner in protein bodies of the rice seed endosperm of transgenic rice plants. The identity of the AMPs accumulated in the endosperm was demonstrated by Western blot analysis and the correct processing was confirmed by LC-MS/MS analysis. The Cec A peptide accumulating into protein bodies of the starchy endosperm was partially purified by chromatographic separation methods. A highly purified BP178 peptide was obtained by using a single purification step. The attachment of the KDEL retention signal to the Cec A peptide does not influence its accumulation in the rice seeds.

As for the *oleosin_PRS_Cec A* transgenic plants, the fusion protein was found to accumulate in the oil bodies of the rice embryo. Both, the fusion protein and the S-Cec A (obtained by TEV digestion of the oleosin_Cec A fusion protein) were partially purified by chromatographic separation methods. Most importantly, the S-Cec A peptide displayed In vitro bacteriostatic and bactericidal activity against *Dickeya* sp. 1552.10.1.

Among the different strategies assayed in this work for production of Cec A in rice seeds, the highest levels of AMP accumulation was achieved by expressing the *Cec A* gene as a translational fusion to the 18 kDa oleosin gene.

5. DISCUSSION

The future of the control of plant diseases is linked to the emergence of new compounds that must meet the strict regulations about toxicity, environmental impact and biodegradability of pesticides in many countries. Antimicrobial peptides represent a promising alternative to conventional antibiotics for plant disease protection against bacterial and fungal diseases, while reducing the emergence of resistance in target pathogens. Biotechnological applications of AMPs also include their use as therapeutic agents to treat fungal and antibiotic-resistant bacterial infections in animals and humans; as biopreservatives in cosmetics, materials and food; and substitutes of antibiotics in animal feed.

Naturally existing antimicrobial peptides can be optimized by synthesis and combinatorial chemistry approaches to increase their antimicrobial activity and stability, and to reduce the possible toxic effect against non-targeted organisms. However, the chemical synthesis of antimicrobial peptides entails high production costs.

In the present study, we evaluated the feasibility of using rice as biofactory for the production of biologically active AMPs, by designing strategies for AMP accumulation in seed tissues, and in particular subcellular compartments. Towards this end, we generated transgenic rice plants accumulating *de novo* designed antimicrobial peptides in seeds. The strategy followed involved the use of various seed-specific promoters and codon-optimized AMP DNA sequences, as well as, the preparation of chimeric genes in which the nucleotide sequence coding for specific targeting and stabilization amino acid sequences were fused to the antimicrobial peptide. Different procedures for extraction and purification of the recombinant AMPs from seed materials have been explored, and the identity and bioactivity of the in planta produced AMPs determined.

5.1 Synthetic BP100 derivatives for plant expression

The heterologous expression of AMP genes requires overcoming several challenges to allow its expressability in plants. Accordingly, in this work, synthetic peptides have been designed and synthesized by introducing modifications directed 1) to increase the peptide size (elongation, n-merization), and 2) to include sequences for stabilization (AGPA), internalization and/or retention into the endoplasmatic reticulum of the plant cell (N-terminal signal peptide, C-terminal KDEL extension). All these approaches aimed to achieve high expression levels of the AMPs under study while avoiding the risk of peptide degradation by host proteases and/or loss of antimicrobial activity.

To identify optimum peptide candidates for plant expression, the aforementioned modifications have been made on the leader peptide BP134 (H-KKLFKKILKYL-OH). The development of BP134 derivatives started with the application of a combinatorial approach on the sequence of the Pep 3 (WKLFKKILKVL- NH2) peptide, a cecropin A_melittin hybrid exhibiting both antibacterial and antifungal activity (Cavallarin et al., 1998; Ferré et al., 2006). Based on the amino acid sequence of this peptide, a library of linear undecapeptides displaying increased antimicrobial activity against the plant pathogenic bacteria (*Xanthomonas axonopodis* pv. vesicatoria, *Pseudomonas syringae* pv. syringae and *Erwinia amylovora*), with minimized cytotoxicity and moderate protease susceptibility compared to Pep 3, was generated (Ferré et al., 2006). Among them, the BP76 peptide (KKLFKKILKFL-NH2) was chosen as the best lead compound and used as a template for the design and synthesis of a peptide library comprising 125 linear undecapeptides (the CECMEL11 library) exhibiting improved physicochemical properties compared to the parental BP76 peptide (Badosa et al., 2007). Among the CECMEL11 peptide library, BP100 was the best performing compound (amidated form, KKLFKKILKYL-NH2), and was further modified for its production in plant systems, thus, resulting in the BP134 peptide (H-KKLFKKILKYL-OH).

Modifications on the BP134 sequence were made on the basis of the specific requirements for their expression in rice plants, while maintaining the ideal α-helical wheel of the original BP134 peptide (accordingly to the hydrophobicity/amphipathicity required to maintain or increase the antimicrobial activity). In this way a peptide library was constructed with the *de novo* designed BP134 derivatives which comprised 52 peptides (Badosa et al., 2013).

When analysing the antimicrobial activity of peptides in the library of BP134 derivatives, we observed that changes or additions on peptide sequences can have an important effect on their activity and toxicity (Badosa et al., 2013). These results are in agreement with those previously reported in the CECMEL11 library members (Badosa et al., 2007). Thus, the BP178_KDEL, BP188_KDEL and BP192_KDEL peptides showed improved antibacterial activity profiles, decreased cytoxicity and increased length. These three peptides contain the ER retention signal and the AGPA stabilization sequence (BP178 and BP192) and were the AMPs chosen to be produced heterologously in transgenic plants.

The selected peptides exhibit a strong cationic net charge (pI in the range of 10.72 to 11.18) and their amphipathic character was evident when the ideal α-helix was represented by means of an Edmunson wheel plot. Most important, these peptides display strong antibacterial and bactericidal effect against the economically important plant-pathogens *Xanthomonas axonopodis* pv. vesicatoria, *Pseudomonas syringae* pv. syringae, *Erwinia amylovora* and *Dickeya* sp., while having reduced hemolytic and phytotoxic activities (Badosa et al., 2013). Indeed, the antibacterial activity displayed by the selected peptides BP178_KDEL, BP188_KDEL, BP192_KDEL, was comparable to that of antibiotics routinely used for bacterial disease control. These peptides presented MIC values in the range of 0.6 to 10 μM, with hemolytic and phytotoxic activities only significant at concentrations 50-100 times higher than the corresponding MIC values.

In addition, the Cec A and Cec A_KDEL peptides were chosen as references because previous reports indicated that the Cec A peptide displays potent antimicrobial activity against economically important plant pathogens (Cavallarin et al., 1998). Furthermore, the constitutive expression of the cecropin A gene was successfully accomplished in rice plants, conferring a resistance phenotype to infection by fungal pathogens (Coca et al., 2006). To note, targeting of the Cec A to the apoplastic space (without KDEL retention signal) resulted in non fertile plants (Coca et al., 2006). Along with this, when the cecropin A hybrid undecapeptide (namely Pep 3) without ER retention signal was intended to be produced in rice plants, the peptide was found to be toxic towards plant cells (San Segundo, unpublished results). The problems encountered in the above mentioned strategies may be due to the expression under the control of a constitutive promoter (*ubiquitin-1* promoter). The strategies for tissue-specific expression of *Cec A* in our study allowed us to determine whether the unintended negative effects that were previously observed in rice plants constitutively expressing *Cec A* were abolished in transgenic rice accumulating Cec A in seed tissues.

5.2 Transgenic rice plants producing antimicrobial peptides in seeds

Plant systems offer several advantages for the large-scale production of AMPs in comparison to other systems such as microbial or mammalian cell cultures. Recombinant proteins expressed in plants are produced at lower production costs, with low risk of contamination by human or animal pathogens, the proteins are correctly folded, easy of scaling up or down and using relatively simple purification technology (Hood et al., 2002). Moreover, the use of seeds for the production of recombinant proteins, instead of other plant tissues, has been shown to have several advantages (Delaney, 2002).

Seeds represent an attractive system for production of important amounts of recombinant proteins because of their ability to naturally store large amounts of proteins in the endoplasmatic reticulum and Golgi-derived cell compartments, the so called protein bodies (PB) types I and II, respectively.

In addition, recombinant products accumulated in seeds may remain stable for long periods of time at room temperature without loss of activity (Sharma and Sharma, 2009). Additionally, the oil bodies in the embryo and aleurone tissues offer another strategy to target the foreign protein in this organelle.

The intracellular deposition site of the foreign proteins greatly determines the production levels, which in turn are related to the own physical properties of the foreign protein. Moreover, when the recombinant protein bears the KDEL signal, its interaction with other ER resident proteins, or the assembly state of the protein itself, might be critical in determining its accumulation level (Stoger et al., 2005; Wang et al., 2013).

The accumulation of recombinant proteins in PBs might protect the protein from the action of cytosolic proteases and facilitates their purification by a simple sucrose density centrifugation, thus, increasing the efficiency of the production system. In addition, peptides can be accumulated in such organelles without subjecting the ER to an excessive level of stress (Vitale and Cerotti, 2004).

Previous studies in our group failed in the obtention of stable transgenic rice plants expressing certain *BP134der* genes under the control of constitutive promoters (including ER targeting), suggesting toxicity of these AMPs to the host plant (Nadal et al., 2012; Company et al., 2013). In other cases, the *BP134der* genes were constitutively expressed in rice, but their expression resulted in abnormal growth and development of the transgenic rice plants (Company et al., 2013). Accordingly, aspects related to the possible cytotoxicity to the host cells were considered in the present work. Results here presented illustrate the usefulness of the rice seed system for the production of AMPs, with no adverse effects on growth and development of the plants.

Particularly, four seed-specific promoters which are known to be active in seed tissues, either in the endosperm (*glutelin B-1*, *glutelin B-4*, and *26 kDa α-globulin*), or in the embryo (*18 kDa oleosin*) were assayed. The *glutelin B-1*, *glutelin B-4*, and *26 kDa α-globulin* promoters were isolated from *Oryza*

sativa ssp. *japonica* cv. Senia genomic DNA. In addition, the signal peptide sequence of the corresponding storage protein (glutelin B-1, glutelin B-4, and 26 kDa α-globulin) was used to direct internalization of the AMP into the ER. Whenever necessary, the KDEL signal was added to the C-terminal end of the peptide for its retention in the ER. As an additional advantage, in some cases the addition of the C-terminal KDEL tetrapeptide has been shown to increase the accumulation of fusion proteins while decreasing their toxicity to the plant cell (Ko et al., 2002; Yasuda et al., 2006). As for the *18 kDa oleosin* (*Ole18*) promoter, it was isolated from *Oryza sativa* ssp. *indica* cv. IR36 and used to drive *AMP* gene expression in the rice embryo (the activity of this promoter in the aleurone layer has not been explored in the present work). Finally, the *18 kDa oleosin_AMP* fusion gene was expressed under the control of the *Ole18* promoter.

Before their use for rice transformation, we verified the nucleotide sequences of the isolated endosperm-specific promoters (including the N-terminal signal peptide sequences) which were compared to those annotated in the GenBank database (GenBank accession numbers, AY427569 (*GluB-1*), AY427571 (*GluB-4*) and AY427575 (*Glb-1*)). This analysis revealed major nucleotide changes in the sequence of the *18 kDa oleosin* promoter in comparison to that reported in the database (AY427563). Presumably, these changes in the *18 kDa oleosin* promoter can be associated to different rice subspecies and cultivars used (*indica* rice, present work; *japonica* rice in the AY427563 accession). A number of *cis*-regulatory elements involved in seed and endosperm specific gene expression were also identified. A GCN4 motif involved in endosperm expression was identified in the promoter assayed in our work from the *indica* rice variety that is not present in the *japonica* rice variety (AY427563).

Stable transgenic rice plants (*O. sativa* ssp. *japonica* cv. Senia and *O. sativa* ssp. *japonica* cv. Ariete) were generated by *Agrobacterium*-mediated transformation. Plant transformation vectors contained the cassette for expression of a synthetic *AMP* gene (fused to the N-terminal signal peptide sequence of the seed storage protein) under the control of a seed-specific promoter and the *Nos* terminator sequence. Transformations were made for

up to 14 different gene constructs which yielded stable transgenic rice plants **(Table 5.1)**. The transformation efficiencies (T0 plants /number of starting calli) varied among the various transformation events, ranging from 20 to 103% depending on the transgene. Globally the seed-based strategy for transgenic expression of the *AMP* genes in rice does not significantly alter normal growth and development of the transgenic rice plants, including vegetative growth, flowering behaviour and yield.

Table 5.1 Transgenic rice plants generated in this work producing BP134 and Cec A peptide derivatives in their seeds

Site of expression	Promoter	*AMP* gene	Additional signal sequences: N-terminal	Additional signal sequences: ER retention	Plant genotype	Plant Code
Endosperm	*GluB-1*	*BP178*	+	+	*GluB-1::BP178_KDEL*	S11, A11
		*BP188**	+	+	*GluB-1::BP188_KDEL*	R7, A12
		BP192	+	+	*GluB-1::BP192_KDEL*	R4, S8, MC2
		Cec A_KDEL	+	+	*GluB-1::Cec A_KDEL*	S3, MC3, A3
		Cec A	+	-	*GluB-1::Cec A*	A5
Endosperm	*GluB-4*	*BP178*	+	+	*GluB-4::BP178_KDEL*	A14
		*BP188**	+	+	*GluB-4::BP188_KDEL*	R8, A13
		BP192	+	+	*GluB-4::BP192_KDEL*	A7
		Cec A_KDEL	+	+	*GluB-4::Cec A_KDEL*	S2
		Cec A	+	-	*GluB-4::Cec A*	R10, A1
Endosperm	*Glb-1*	*BP178*	+	+	*Glb-1::BP178_KDEL*	R11
		BP192	+	+	*Glb-1::BP192_KDEL*	S6
Embryo	*Ole18*	*Ole18_PRS_Cec A*	-	-	*Ole18::Ole18_PRS_Cec A*	R9, A16
		Ole18	-	-	*Ole18::Ole18*	MCX

* The *BP188_KDEL* endosperm expression was confirmed in R7, A12, R8 and A13 lines, however, peptide production stills to be confirmed. For the other peptides, BP178_KDEL, BP192_KDEL, Cec A_KDEL and Cec A, both the transgene expression and peptide production in the rice endosperm were confirmed.

Some transgenic plants, such as the *pCGluB-4::BP188_KDEL* (R8), *pCOle::Ole18_PRS_Cec A* (R9), and *pCGluB-4::Cec A* (R10) lines, showed higher biomass (root and shoot systems) and increased seed yield than the control plants (WT and EV plants). On the contrary, although the *pGluB-1::BP188_KDEL* (R7) lines were fertile, they exhibited reduced growth and abnormal phenotype. As this phenotype was observed in all independent lines bearing the same construct, the possibility of an insertional adverse effect caused by transgene integration into the genome can be excluded. Thereby, phenotypical differences observed between *pGluB-1::BP188_KDEL* (R7) and *pCGluB-4::BP188_KDEL* (R8) plants might be associated to the promoter used to drive *BP188_KDEL* gene expression (*GluB-1* promoter in R7 lines, *GluB-4* promoter in R8 lines). Additional studies are, however, needed to explain the unintended effects caused by expression of the *BP188_KDEL* gene under the control of the endosperm-specific *GluB-1* promoter on rice development.

Integration of the transgene into the rice genome and its integrity were confirmed by conventional PCR, whereas the transgene copy number was estimated by qPCR. Confirmation that the *AMP* transgenes, namely the *BP178_KDEL*, *BP188_KDEL*, *BP192_KDEL*, *Cec A* and *Cec A_KDEL*, were properly expressed in developing seeds (17-21 days after flowering, DAF) was performed by RT-PCR. In other studies, however, transcripts for transgenes encoding small peptides (<50 aa) (Matoba et al., 2001; Yasuda et al., 2005; Takagi et al., 2008, 2010) could not be detected (also when introduced into rice plants). For instance, the 30 amino-acid long glucagon-like peptide-1 (mGLP-1, 3.8 kDa) was introduced in rice under the control of the *glutelin B-1* promoter, but neither the transgene transcripts nor the transgene product could be detected in mature seeds (Yasuda et al., 2005). Moreover, Yasuda and co-workers (2006) reported that the trimer of mGLP-1 (mGLP-1x3, 11.4 kDa) had the minimal size among different multimers to be expressed in transgenic rice. From the results here presented, it appears that the antimicrobial peptides under study whose size ranges from 28 to 41 amino acids (from 3.14 to 4.49 kDa) can be successfully expressed in transgenic rice using seed-specific rice promoters to drive their expression.

In addition, the *Cec A* gene was successfully expressed into the embryo as a translational fusion to the *18 kDa oleosin* gene.

The detection and isolation of the small and strongly cationic AMPder peptides was a challenging task in the present work. Several problems were encountered which were mainly related to the physicochemical properties of the AMPs (e.g. small size, strong cationic/basic and amphipathic character). Problems for detection of such short and basic peptides by Western blot analysis are commonly found in the literature in different plant species, such as potato (Osusky et al., 2004), and rice (Coca et al., 2006, Company et al., 2013). The development of fast and efficient subcellular fractionation procedures, in combination with chromatographic methods allowed us to demonstrate that the synthetic AMPs were efficiently produced and accumulated in seeds of most transgenic events. Thus, Western blot analysis of protein extracts obtained from protein bodies and oil bodies from AMP-expressing rice plants demonstrated that the AMP of interest accumulated in the expected seed organelle (either protein bodies or oil bodies). The AMPders designed to be accumulated in protein bodies were exclusively detected in the endosperm tissue of the transgenic seed, but not in embryo, leaf or root tissues. Their site of accumulation was not influenced by the presence or absence of the KDEL retention signal. These results suggest that the signal peptide fused to the AMP peptide plays an important role in targeting AMPs to the PBs of rice seeds and contain the molecular determinants required for this process. As for the C-terminal KDEL sequence, this tetrapeptide is not always sufficient to retain transgenically produced proteins in the ER lumen of plants (Jones and Herman, 1993; Okamoto et al., 1994; Pueyo et al., 1995).

Several reports illustrate the utility of using an endosperm specific promoter to drive expression of a gene of interest in transgenic rice. For instance, the *ferritin*, *β-glucronidase* (GUS), *glutelin A-3 Crp* (3 Crp, derived from Japanese cedar pollen allergens) and *glutelins-IIAEKx6* (derived from bovine milk β-lactoglobulin) genes were expressed in rice under the control of endosperm specific promoters (Goto et al., 1999; Qu and Takaiwa, 2004; Takagi et al., 2008, 2010; Wakasa et al., 2011b). However, there are also examples in

which this strategy has been unsuccessful. As an example, Yasuda and co-workers (2006) described that by expressing a chimeric gene which consisted in the *GluB-1* signal peptide-*mGLP-1x6_KDEL* gene under the control of the *glutelin B-1* promoter, the transgene product could not be detected in the PBs.

The possibility that the AMPs under study in our work accumulate in PB regardless of the targeting signals included in the peptide sequence, but due to the interaction of the peptide with seed storage proteins (normally targeted to PBs) should be also considered. Even, protein aggregation can also play a dominant role for protein sorting to the PBs. In this sense, Yang et al., (2003) reported that the sequence of the human lysozyme contains a targeting signal which directs the recombinant protein to the PB II, regardless of the targeting signals.

The fact that certain peptides can not be produced in plants could also be associated to other factors, such as degradation of the transgene products within the cell host, transgene silencing, or to the need of a particular threshold size or peptide structure. Some examples of peptides that could not be produced heterologously in rice seeds are: the 10-repeat novokinin (~11 kDa), the T cell epitope from the Japanese cedar pollen allergens Cry j 1 and Cry j II (3 Crp, 5.9 kDa), and the glucagon-like peptide-1, mGLP-1 (3.8 kDa) (Yasuda et al., 2005; Takagi et al., 2008, 2010; Wakasa et al., 2011a). These observations led some authors to propose that the best way to produce small peptides in transgenic plants is by coupling these peptides to a seed storage protein (Hashizume et al., 2008; Wakasa et al., 2011b; Cabanos et al., 2013). *IIAEK*, derived from bovine milk β-lactoglobulin (6x peptide fusion to glutelin), the *3 Crp* (fused to glutelin acidic subunit) and the *type II collagen* (4x peptide fusion to glutelin A) transgenes, were successfully synthetized in transgenic rice seeds (Takagi et al., 2008; Hashizume et al., 2008; Wakasa et al., 2011b). Also, peptide tandems (e.g. 18 repeated novokinin, mGLP-1x6) were successfully produced in rice seeds (Yasuda et al., 2006; Wakasa et al., 2011a). As mentioned above, the 30 amino-acid long glucagon-like peptide-1 (*mGLP-1*) was directly expressed in rice seed under the control of the *Glutelin B-1* promoter, but the product

peptide could not be detected in mature seed. However, the successful production of the mGLP-1 peptide in rice seeds was achieved when the *mGLP-1* peptide gene was expressed as a chimeric construct with a reporter gene (e.g. *GFP*) (Yasuda et al., 2005).

An unexpected result in our studies was the mobility of the produced AMPs in electrophoretic systems. Based on the Western-blot analysis, the apparent molecular weight of the *in planta* produced AMPs was higher (based on their mobility on Tris-Tricine SDS-PAGE gel) than the expected from corresponding synthetic compound. A similar phenomenon has been reported in transgenic plants expressing constitutively either *Cec A* or a *BP100der* (Coca et al., 2006; Company et al., 2013). In both studies it was argued that this increase in size could be due to multimers formation, suggesting a higher stability of these peptides as dimerized forms. Yasuda et al., (2006) reported that the mGLP-1x6 (KDEL) peptide produced in rice seeds also exhibited an apparent higher molecular weight relative to that of the corresponding standard peptide, thus hypothesizing that the N-terminal signal peptide fused to the peptide was not properly processed. Similarly, Nochi and co workers (2007) expressed a fusion gene encoding the cholera toxin B subunit (CBT) polypeptide fused to the N-terminal signal peptide of glutelin B-1 in rice, and two different polypeptide species of 12 and 15 kDa were immunodetected. These authors suggested that the high molecular weight polypeptide (15 kDa) corresponded to non-processed (or partially processed) fusion polypeptide (protein of interest and N-terminal amino acid sequences)

In our case, different possibilities were considered to explain the abnormal electrophoretic mobility of the AMP peptides, including improper processing of the signal peptide, post-translational modifications in the AMP peptide, or formation of self-assembled peptide multimers. By using LC-MS/MS we were able to demonstrate that the apparent higher molecular mass observed in the seed-produced AMPs was not related to incorrect processing of the transgene product (signal peptide not present).

It is also known that proteins that enter into the secretory pathway are often modified in plant cells by adding sugar chains to their amino acid sequence. Thus, there was the possibility that the high molecular mass of the *in planta* produced AMP was associated to glycosylated forms (and not to n-mers formation). After several attempts, we were unable to determine if the plants produced BP134der and CecAder that were glycosylated (or glycated). Clearly, further studies are needed to elucidate whether the higher molecular mass that AMPs exhibit in Tris-Tricine SDS-PAGE gels correspond to peptide multimerization, abnormal behaviour of the plant-made AMPder in this electrophoretic system, or posttranslational modifications of the recombinant peptide products.

Definitive proof that the ~8 kDa immunodetected polypeptide corresponded to the expected peptides (BP178_KDEL, BP192_KDEL, Cec A_KDEL or Cec A) came by LC-MS/MS data. Thus, analysis of protein extracts obtained from PBs of transgenic seeds bearing *GluB-1::BP192_KDEL* (MC2, S8, R4), *GluB-1::Cec A_KDEL* (MC3, S3), *GluB-1::BP178_KDEL* (A11), *GluB-1::Cec A* (A5), *GluB-4::BP192_KDEL* (A7), *GluB-4::Cec A_KDEL* (S2), *GluB-4::Cec A* (R10, A1), *Glb-1::BP178_KDEL* (R11) or *Glb-1::BP192_KDEL* (S6) gene constructs identified the expected tryptic peptides for AMPs.

In Cec A-lines, the Cec A peptide was also identified in a gel slice band corresponding to 35-45 kDa. The identity and correct synthesis of the Cec A peptide produced using the oleosin fusion technology was also confirmed by LC-MS/MS of protein extracts obtained from oil bodies of transgenic seeds bearing the *ole18_PRS_Cec A* fusion gene.

Unfortunately, the BP188_KDEL peptide could not be detected by LC-MS/MS analysis of seeds from any of the BP188_KDEL-transgenic lines (*GluB-1::BP188_KDEL*, R7 and A12; *GluB-4::BP188_KDEL*, R8). Similarly, the chemically synthesized peptide used as standard was not detected. However, the RT-PCR analysis of the transgenic lines allowed us to detect *BP188_KDEL* transcripts indicating that this transgene was expressed in developing BP188_KDEL-rice seeds. Moreover, the *GluB-4::BP188_KDEL* (R8)

lines exhibited a clear and consistent phenotype of resistance to infection by the bacterial and fungal phytopathogens tested in this work.

The accumulation levels of the AMPs in seed protein extracts showed an important variability among transgenic lines even in those harbouring the same construct. Similar results were reported when T-cell epitope peptides derived from cedar pollen allergens accumulated in transgenic rice seeds (Takagi et al., 2005a). Also, Wakasa and co-workers (2006) reported different accumulation levels of lactostatin in transgenic rice seeds among lines. Most probably, this variability is due to positional effects, transgene copy number, or gene silencing effects. The intrinsic physical condition of the plant might also contribute to the observed differences in AMP accumulation.

Our results also indicated that the accumulation levels of the endogenous seed storage proteins (e.g. glutelins, prolamins) were not significantly altered by the accumulation of the AMPs of interest in the transgenic rice seeds as judged by the band intensity and protein pattern in seed protein extracts. By contrast, Wakasa and co-workers (2013) reported that the production of transgenic rice seeds accumulating the major Japanese cedar pollen allergens, Cry j 1 and Cry j 2 at high levels, was accompanied by an important reduction in the accumulation of most endogenous seed storage proteins. The authors attributed this observation to ER stress due to the accumulation of high levels of foreign proteins in the ER system. Similarly, a high level of accumulation of a hybrid peptide comprising seven predominant human T cell epitopes (derived from Japanese cedar pollen allergens, or 7 Crp) resulted in higher accumulation of 55 kDa proteins (precursors of Glu A and Glu C) and lower accumulation of 25 and 15 kDa proteins (e.g. α-globulin) compared to non-transgenic plants (Takagi et al., 2006). In our case, it is possible that, the accumulation of foreign peptides in the PB of rice seed prevents the ER stress described in other reports. Alternatively, the accumulation levels of the BP134 and Cec A derivatives did not reach the threshold level required to induce ER stress.

Natural and synthetic AMPs such as Cec A, Ace-AMP1, Dm-AMP1, BrD1 and BP100 derivatives, have been successfully produced in the heterologous

system of rice, and the transgenic plants were reported to exhibit enhanced protection against a variety of plant pathogens (Coca et al., 2006; Roy-Barman et al., 2006; Choi et al., 2009; Jha et al., 2009; Imamura et al., 2010; Nadal et al., 2012). In the present work, seed-specific expression of the various AMP genes also provided a high level of protection against infection by the plant pathogens *Dickeya* sp. 1552.10. 1 and *Fusarium verticillioides*. In our experimental conditions, transgenic lines expressing *pCOle18::Ole18_PRS_Cec A*, *pCGluB-4::Cec A*, *pCGluB-1::Cec A*, *pCGluB-4::Cec A_KDEL*, *pCGluB-1::Cec A_KDEL*, *pCGluB-4::BP188_KDEL* and *pCGluB-1::BP192_KDEL*, *pCGluB-4::BP178_KDEL* and *pCGlb-1::BP178_KDEL* constructs showed a clear and consistent phenotype of resistance to infection by the bacterial pathogen *Dickeya* sp. 1552.10.1, compared to control plants (WT, empty vector). By contrast, rice seeds expressing the *Glb-1::BP192_KDEL* gene together with lines expressing either the *BP178_KDEL* gene or the *BP188_KDEL* gene under the control of *GluB-1* promoter were found to be sensitive to *Dickeya* sp. infections (at levels similar to those observed in control seeds, WT and empty controls). Transgenic seeds bearing the *Cec A* gene (with or without KDEL), *BP188_KDEL* and *BP178_KDEL* under the control of a *glutelin* promoter (*GluB-1*, *GluB-4*) also showed better germination capability when inoculated with the fungal plant pathogen *Fusarium verticillioides* compared to fungal-infected control seeds (WT, empty vector). By contrast, all transgenic lines expressing the *BP192_KDEL* gene showed susceptibility to *F. verticillioides* infection at levels similar to the control seeds). Here, it should be mentioned that BP134 derivatives were designed and synthesized based on their antibacterial activity, while its antifungal activity was not primarily pursued.

Differences in the degree of protection against pathogen infection between transgenic plants carrying the same construct were also observed, which could be explained by transgene dosage and integration site into the rice genome, or to transgene expression levels. Generally, no clear relationship between phenotype (resistance/susceptibility) and transgene copy number or level of AMP production was observed. In agreement with this, Fearing and co-workers (1997) demonstrated no connection between transgene

dosages and enhanced transgene expression in cereal crops. By contrast, Jha et al., (2009) reported a significant correlation between Dm-AMP1 levels and the degree of disease resistance observed in each transgenic line.

Globally, the results obtained in seed infection bioassays, indicated that the levels of expression of the *AMP* genes that are achieved using seed targeted strategies are sufficient to confer resistance against the tested plant pathogens. This is also an indication that the *in planta* produced AMPs were biologically active against the phytopathogens. In addition, a possible additive effect of the stress resulting from the transgenic expression of an *AMP* gene (product) might also induce the endogenous defense responses in the plant. However, further studies on the expression of defense marker genes are needed to explore this possibility.

The success in obtaining transgenic plants with stable expression of an *AMP* gene of interest through successive generations is also desirable when the use of the rice seed as biofactory is pursued. In this respect, AMP expression in the transgenic rice seeds obtained in this work was proven to be inherited, at least up to the third generation. Moreover, peptide accumulation was found to be maintained during storage of the rice seeds (up to two to three years), the peptide being also detected during the early stage of seed germination.

For the subsequent stages of the work, seed amplification of the selected transgenic lines (up to six T3 homozygous lines for each line and transformation even) was carried out. Criteria for selection of transgenic lines were the following: (1) high level of AMP accumulation in the expected subcellular compartment, (2) high level of resistance to infection by *Dickeya* sp. 1552.10.1 and *Fusarium verticillioides* and, (3) generally a single copy of the transgene (exceptionally more than one). Single copy events were preferred because the integration of multiple copies of transgenes in the plant genome could result in silencing and/or rearrangements of the transgene. Among the transgenic plants selected, all the single-copy lines yielded T3 homozygous rice plants. However, only 28.6% of the transgenic lines presenting more than one copy of the transgene yielded T3

homozygous plants. Most of these lines showed high levels of peptide accumulation, thus, pointing to a possible effect of toxicity due to AMP accumulation. This issue needs to be further explored.

In summary, at least three homozygous transgenic rice lines from each transformation event corresponding to fourteen different constructs were obtained, most of them harbouring a single copy of both the selectable marker and the *AMPder* transgene. The stable integration and inheritance of the *AMP* transgenes through successive generations (at least up to the T3 generation) was confirmed.

5.3 Characterization and yield of recombinant peptides in transgenic rice seeds

An important issue for the development of suitable plant expression platforms for recombinant protein production is the procedure used for their obtention and purification from plant tissues, which directly impacts the quality and yield of the resulting protein product and, subsequent applications (pesticides, pharma or food preservatives, etc.). Thus, the procedure should render a biologically active product while avoiding the need of time-consuming steps. The procedure also should be cost-effective and have scalability easily adaptable to production needs. For these reasons, and important effort was made during the course of this work to develop suitable methods for AMP production in rice seeds with the easiest purification procedures and lowest costs possible. It was then necessary to determine the yield of production in order to estimate the feasibility of using rice seeds as biofactories for AMPs, and to demonstrate that the recovered products were biologically active.

AMP accumulates in the rice endosperm

Regarding to the AMP accumulation into PB of the starchy endosperm, a simplified procedure based on two-step centrifugation was implemented for preparation of partially-purified protein bodies. An initial low speed centrifugation step allowed removing cell debris and most starch from the initial seed homogenate, and a second centrifugation enabled the obtention

of an enriched AMPder-containing PB fraction. By using this protocol, partially purified protein bodies were recovered in a few minutes.

Once the subcellular organelles were isolated, purification of the AMPs (free of host proteins) has proven to be a major difficulty. After disruption of the isolated protein bodies, a poor solubility of the peptides in the protein extract was evidenced. This problem can be related to the extreme physicochemical properties exhibited by our peptides, including a strong cationic net charge, and amphipathic character that confer a tendency to self-aggregation and facilitates interactions with seed storage proteins (e.g. acidic glutelins).

For the establishment of suitable purification methods, different types of chromatographic separation techniques were initially evaluated using chemically synthesized peptides as references. Unfortunately, when strong cation exchange chromatography was used, the synthetic peptides remained tightly bounded to the sorbent (e.g. 60-80% of elution salt buffer was needed to partially elute the synthetic peptides) or, by contrast, they aggregated becoming too large to enter trough the column sorbent. As a consequence, the recovery of the AMPs was too low. These results proved the difficulty to elute the peptides from a strong cation exchange sorbent.

After several trials, the synthetic peptides were successfully purified by Weak Cation Exchange-Reversed Phase chromatography. The AMPs were found to be soluble in aqueous 25 mM ammonium acetate, pH 6.0 buffer which increases the positive charge of peptides (stronger interaction with the negative charges of the weak cation column sorbent). AMP elution from the chromatographic matrix was achieved by using 50% of acetic acid which was then easily removed by lyophilisation (this avoids the need of dialysis). The yield of peptide recovery was ~60% relative to the initial amount of peptide. The remaining amount of synthetic peptide might either remain tightly bounded to the column sorbent or, by contrast, might aggregate becoming too large to enter trough the column sorbent. As for the first possibility, additional washing steps with 100% of acetic acid, formic acid, acetonitrile or NaCl 2 M did not yielded additional peptide fractions.

Once the procedure for AMP purification was established using synthetic AMP peptides as the reference materials, the purification of the recombinant peptides from rice seeds was approached. For this, total protein extracts from Cec A and BP178_KDEL-protein body fractions were precipitated with a mixture of acetone+TCA+DTT. As for Cec A, it precipitated together with the seed storage proteins, whereas the BP178_KDEL was found to be soluble in the precipitation mixture. Taking advantage of the observed solubility of BP178_KDEL peptide in this condition, the protein extracts obtained from PBs were precipitated (acetone+TCA+DTT) and the seed storage proteins were removed by centrifugation. The BP178_KDEL-containing acetone fraction was lyophilized, and the BP178_KDEL peptide was recovered. Therefore, a single purification step was required to obtain a high quality BP178_KDEL peptide from PB enriched fractions.

Several efforts were made to purify the Cec A peptide from protein body fractions with no success. Although 25 mM ammonium acetate (pH 6.0) buffer served to solubilize the lyophilized chemically synthesized peptides, it failed to solubilize the lyophilized Cec A-seed storage protein extract, as visible insoluble aggregates were formed.

Unsuccessful attempts were also made to obtain Cec A by organic solvents (acetic acid, acetonitrile) or saline buffers (and at different pH) as the solubilising agents. In all the cases, the Cec A peptide precipitated together with storage proteins (probably by interacting with glutelins, as they are acid-soluble).

Partial solubilization of protein extracts obtained from protein bodies was accomplished using 0.1%TFA and the Cec A was then partially purified by Weak Cation Exchange-Reverse Phase column (following the protocol developed using the synthetic AMPder). Unexpectedly, Cec A was eluted in the wash fraction consisting of 100% methanol instead of eluting in the acetic acid fraction (as found for the synthetic Cec A peptide). These results indicated properties of the *in planta*-produced Cec A peptide (e.g. changes in peptide polarity) different from those of the reference synthetic peptides.

Also of interest, was the fact of the apparent molecular weight of Cec Ader (based on its electrophoretic mobility in Tris-Tricine SDS-PAGE) of the *in planta*-produced Cec A was ~8 kDa (expected Cec A monomer size, ~4 kDa). In addition, when 50 % acetic acid was used after the methanol elution step, a ~37 kDa polypeptide reacting with the anti-Cec A antibody was observed (that Cec A was present in this high molecular weight protein fraction was confirmed by LC/LC Mass spectrometry). A reasonable explanation for this observation is that Cec A interacts with seed storage proteins, particularly with glutelins, because evidence of CecA multimers in rice protein extracts is also presented (CecA-related polypeptides of 4, 8 and 12 kDa are immunodetected). Finally, although peptide glycosylation could explain a change in peptide polarity (and solubility) and electrophoretic mobility (higher apparent molecular weight) of the in planta-produced Cec A, no evidence on post-translational modifications were obtained in this work.

Oleosin_Cec A fusion protein accumulates in oil bodies

Oil bodies were purified from seeds of rice plants harbouring an *oleosin_Cec A* fusion gene by two consecutive cycles of two layer flotation. Cleavage of the oleosin_PRS_Cec A fusion protein with the TEV protease releases a Cec A peptide in which an extra serine residue is attached to its N-terminal end (TEV protease recognition signal, PTTENLYFQ↓S-Cec A). The S-Cec A was then purified from TEV digestion mixtures using Weak Cation Exchange-Reversed Phase chromatography. We observed that the efficiency of the TEV digestion greatly varied.

In our strategy for production of AMPs in rice embryo, the Cec A peptide was fused to the C-terminal end of the 18 kDa oleosin protein, which is exposed to the cytosol, and adopts an amphipathic α-helical structure. Thus, theoretically, the fusion protein should be accessible to proteolytic digestion in intact oil bodies (i. e. no need of oil body disruption and/or protein purification) (Parmenter et al., 1995).

After TEV digestion, either the crude OB fraction or the OB protein extract were solubilized in TFA 0.1% and the S-Cec A peptide was eluted in the 100% acetic fraction together with the fraction of undigested fusion protein. These results suggest: 1) the incomplete proteolysis of the fusion protein

(Parmenter et al., 1995), 2) interaction of the released S-Cec A with the oleosin_Cec A fusion protein, and 3) formation of undesirable fusion protein aggregates that might difficult the access of the TEV protease to the cleavage site of the fusion protein.

The S-Cec A peptide released from the OB of the *Ole18::Ole18_PRS_Cec A* lines is biologically active

Although resistance of transgenic lines expressing the AMPders *Dickeya* sp. 1552.10.1 was previously demonstrated by means of direct seed bioassays, evidences that the *in planta* purified AMPs maintain antimicrobial activity became necessary. Evidence of its bioactivity came from the *in vitro* antibacterial and bactericidal activity assays of the purified expression product. It was proven that the S-Cec A peptide released from the oleosin_Cec A fusion protein exhibited a high antibacterial and bactericidal activity. This finding supports that the S-Cec A maintains an active conformation allowing the peptide to interact with the bacterial membrane. Moreover, the serine residue at the NH_2-terminal amino acid sequence of Cec A does not interfere with Cec A antimicrobial activity. In agreement with this, synthetic peptides in which a serine residue is attached to their N-terminal sequence (e.g. BP134ders) maintained either bactericidal or antibacterial activities **(Section 4.1.3, Results, Table, 4.2)**.

Although seed bioassays revealed resistance to *Dickeya* sp. 1552.10.1 in rice seeds expressing the *oleosin_Cec A* fusion gene, in our hands, no antimicrobial activity could be observed in oleosin_Cec A preparations that have not been digested with TEV. This suggests that the Cec A display activity only upon being released from the oleosin protein carrier. Although the reason for this apparent discrepancy between seed bioassays and oleosin_Cec A activity remains unknown. Several possibilities can be considered to explain the activity of the oleosin_Cec A, such as proteolytic digestion of the oleosin_CecA fusion protein by endogenous proteases within the seed or improper folding of the fusion protein may be acquired under the *in vitro* conditions used to asses its antibacterial activity.

In addition to the confirmation of the antimicrobial activity of S-Cec A, the *in vitro* activity assays with *Dickeya* sp. 1552.10.1 allowed us to infer the amount of S-Cec A peptide purified from oil bodies by comparison of tis activity to the one of synthetic Cec A. In this way, a S-Cec A concentration of at least at 4 µM in the oil body protein sample (69 µg of oil body proteins in 25 µL) was estimated. Taking into account the molecular weight of the S-Cec A (4.091 kDa), this means that at least 409.1 ng of free S-Cec A was present in the TEV-digested protein extract containing the Ole18_Cec A fusion protein. The released S-Cec A represented about 0.6% of the total amount of proteins obtained from the oil body fraction. Knowing that the starting material for macroextraction of oil bodies was 1500 seeds, and that 245 mg of total protein extracts were processed, the S-Cec A content was estimated to be around 0.976 µg per seed. The estimated S-Cec A amount is higher than that obtained from the Western-Blot analysis quantification. However, there is the possibility that the *in planta*-produced S-Cec A displayed higher antibacterial activity than the chemically synthesized Cec A. Underestimation of the S-Cec A content in protein extracts based on immunoblot detection should also be considered (poor electrophoretic transfer efficiency, low immunogenicity, etc..). In the future, taking into account that complete processing of the oleosin_Cec A fusion protein does not occur, better yields can be expected by improving TEV protease digestion conditions.

In other reports, the oleosin_fusion platform has been exploited to produce and accumulate foreign recombinant proteins and high value polypeptides in the oil body surface of seeds from plants, namely safflower, sesame, soybean, rapeseed and arabidopsis (van Rooijen and Moloney, 1995; Boothe et al., 2010, SemBioSys Genetic Inc.). For instance, Parmenter et al., (1995) reported the production of an oleosin_hirudin fusion protein in *Brassica napus* seeds (under the *Arabidopsis oleosin* promoter). In this case, the oleosin_hirudin fusion protein (fusion protein, 26 kDa; hirudin, 6.9 kDa) accumulated at levels up to 1% in total seed protein extracts. Considering that hirudin corresponds to a 26.5% of the fusion protein, the total amount of released hirudin would account for 0.26% of total seed protein extract. A

biologically active recombinant insulin was found to accumulate up to 0.13% (referred to total seed proteins) when expressed as an oleosin fusion in *Arabidopsis thaliana* (Nykiforuk et al., 2006). Moreover, the product of various single-chain antibodies fused to trimeric polyoleosin (~75 kDa), accumulated in the *A. thaliana* oil bodies at levels in the range of 0.25-0.9% of total seed protein extracts (Winichayakulet a., 2012). Our data, however, can not be compared with the above mentioned reports because their yields were referred to total seed proteins, and the plant species used in those studies differ in the content of oil bodies and oleosin proteins.

Yield of AMP derivatives produced in transgenic rice seeds

Concerning the strategy of producing an AMP derivative without a fusion to a protein carrier, the approximate amount of AMP accumulating in protein bodies of the transgenic rice seeds was estimated by Western-blot analysis, and for those lines showing the highest accumulation levels. This was the case for rice seeds producing the BP178_KDEL (3.24 kDa), Cec A (4.00 kDa) and Cec A_KDEL (4.49 kDa) peptides. Seeds from *Ole18::Ole18_PRS_Cec A* (R9), *GluB-4::Cec A* (R10), *GluB-1::Cec A_KDEL* (MC3) and *Glb-1::BP178_KDEL* (R11) rice plants accumulated 10.8, 1.3, 1.4 and 1.8 pmols/mg of seed, respectively. Based on these results here presented on the production of AMP derivatives in rice seeds, several considerations can be made.

First, the amount of the AMP derivatives produced in the present work is comparable, when considering the molar amounts of the peptide/protein referred to the total fusion protein, to that reported in other studies in which repeat tandems of peptides (or fusion proteins) are produced in rice seeds. For instance, the Type II-collagen (XCII250-270, tandem of four monomers fused to Glu A), the ScFvT84.66 (single-chain Fv antibody), the IL-10 (interleukin-10), the chimeric toleragen 3 Crp (as a fusion protein with cholera toxin B subunit) and the toleragen Cry j I and Cry j II (part of A1aB1b, a soybean seed storage protein), accumulated in rice seeds at 0.03, 0.1, 2.5, 1.2 and 6.3-10.0 pmols/mg of seed, respectively **(Section 1.6, Introduction, Table 1.6)**. These values are in the same range than the ones obtained in our work for the AMPs produced in seeds. There are also examples in the literature of peptides that could not be produced in rice seeds, such as the

chimeric toleragen 3 Crp (5.9 kDa), 10x novokinin (11 kDa) and the glucagon-like peptide mGLP-1 (3.8 kDa) (Takagi et al., 2008, 2010; Wakasa et al., 2011a; Yasuda et al., 2005),

Second, the addition of the KDEL retention signal to the Cec A sequence appears not to influence its accumulation level in rice seeds (Cec A and Cec A_KDEL accumulated at 1.3 and 1.4 pmols/mg of seed, respectively). Contrary to this, other reports indicate that the presence of the KDEL retention increases both the stability and the yield of the recombinants proteins produced in rice seeds (Takagi et al., 2005a; Takaiwa et al., 2007).

Finally, an underestimation of the peptide yield based on immunoblot detection of short and basic AMPs produced in plant tissues might also occur (Osusky et al., 2004; Coca et al., 2006; Company et al., 2013). It is also true that the majority of rice seed-based production systems for peptides described by several authors have been developed as fusion proteins (to storage proteins) or as repeated tandems (Takagi et al., 2005b; Yasuda et al., 2005, 2006; Takagi et al., 2008, 2010, Wakasa et al., 2011a), a different strategy from the developed by us.

The advantages and disadvantages of the various strategies used in the present work for the production of the AMP in rice seeds are described in **Table 5.2**.

Table 5.2 Comparison of the three strategies for the production of AMPs in rice seed biofactories

Strategy	Localization	Original peptide	Product	Design for adaptation[1]	Maximum yield (ng/grain)	Target organelle isolation	Product purification	Processing
Embryo	Oil bodies	Cec A	Oleosin Cec A	Minimum	863	Easy	Moderate	Complex and expensive
Endosperm	Protein bodies	Cec A	Cec A, Cec A_KDEL	Moderate	103, 129	Easy	Complex	Not needed
Endosperm	Protein bodies	BP134	BP178_KDEL	Complex	117	Easy	Minimum	Not needed

(1) Modifications on the AMP sequence made on the basis of the specific requirements for their expression in rice plants

Among the strategies assayed for the production of Cec A in rice seeds, the highest AMP accumulation level was accomplished by expressing the *Cec A* gene as a translational fusion to the *18 kDa oleosin* gene. This strategy allowed levels of Cec A production ten times higher than those obtained by using direct Cec A accumulation in the rice endosperm. However, the processing of the fusion protein is more complex and expensive.

The yield obtained for production of BP178_KDEL, Cec A or Cec A_KDEL peptides by using the endosperm strategy was similar in all three cases. Production of BP178_KDEL in the rice endosperm has an important advantage as it requires minimum purification and processing steps due to its solubility in acetone. This fact makes BP178_KDEL free of the endogenous storage proteins. Clearly, the production of BP178_KDEL in the endosperm of rice seed permits to obtain a highly purified AMP without the need to use complex purification and processing procedures.

In summary, from the results here presented, it seems that the endosperm tissue of the rice seed can be considered as an appropriated platform for the production of AMP derivatives, with no need of complex downstream processing steps. The system also allows for long-term storage of the AMP-accumulating rice seeds. In the case of the oleosin_AMP fusion strategy, it will be of interest to explore the efficacy of this strategy in oil crops such as safflower, sesame, rapeseed, soybean or sunflower, these plant species having higher oil body content than the rice seed. As an additional advantage, the production and accumulation of the heterologous AMP in seeds of transgenic rice plants generally does not affect the normal growth and development of the plant. Thus, the strategies here presented solve the limitation that was encountered when genes encoding highly cationic α-helical peptides are constitutively expressed in transgenic rice plants, due to AMP toxicity towards the host plant (Coca et al., 2006; Nadal et al., 2012; Company et al., 2013).

Future lines of work might focus on the obtention of higher yields of the peptide by designing and building-up improved AMP gene constructions, or by approaching the simultaneous expression of different AMP genes

accumulating in different subcellular compartments (i.e. accumulation of an AMP in the endosperm protein bodies and another AMP (or the same one) as an oleosin_AMP fusion protein in the oil bodies of the same rice seed. Clearly, the possible adverse effects on the plant cell (e.g. ER stress, cytoxicity) of this AMP overproduction should be examined on a case-by-case.

5.4 Rice seeds as biofactories for the production of AMPs and their applications

The results obtained in our work on the generation of transgenic plants expressing the AMPders genes under the control of seed-specific promoters illustrates the potential of these strategies for their use in crop protection. However, the feasibility of using rice seeds as biofactories for obtaining AMPs offers new possibilities for their application, not only in crop protection, but also in several other fields (biopharmaceuticals, biopreservatives in cosmetics and food preservation).

Antibiotics such as streptomycin are routinely used in agriculture for bacterial disease control with MIC values of 2-9 µM and operational doses for field treatment of around 100-200 µM (Monroc et al., 2006a; Montesinos et al., 2012). Of interest, at the same doses, the AMP derivatives produced in the present work are not expected to display cytotoxic effects. On the other hand, and to prepare the standard rates of 1000 L of fungicide/bactericide product per hectare of crop (based on the above mentioned doses), 100-200 g of active ingredient will be needed. Taking into account the amount of AMP derivatives produced in transgenic rice seeds in the present work, yields of 35 g/ha for endosperm and 300 g/ha for the embryo strategies can be estimated (assuming a standard production for a grain crop of around 6000 kg/ha). From this, it seems that the oleosin_Cec A strategy can be useful for agricultural purposes because a hectare of rice biofactory would be sufficient for management of a hectare of crop, which is in the limit of feasibility.

Regarding to the AMP yield obtained using the endosperm strategies, the BP178_KDEL, BP192_KDEL, Cec A and Cec A_KDEL production, although the

yield was lower than in the case of the oleosin fusion strategy, it still offers interesting opportunities for pharmaceutical application, biopreservatives in cosmetics and food preservation, as great value products can be obtained. They can be used either alone or in combination with other antibiotics.

Finally, the above mentioned figures are interesting when compared to the production cost of a chemically synthesized peptide such as cecropin A or melittin ranging from 100-500 €/mg. In addition, the production and processing costs of the AMPs produced in seeds of transgenic plants may be lower than in other systems for heterologous protein expression such as microbial or animal.

6. CONCLUSIONS

1- A series of peptides derived from the BP134 peptide were chemically synthesized and evaluated for their antibacterial properties. The rationale behind these modifications was to facilitate the expression of their corresponding gene sequences in transgenic rice while incorporating the required targeting signals to direct accumulation of the *in planta*-produced peptide in seed organelles, namely protein bodies and oil bodies. Peptide modifications had a strong impact in activity against bacterial pathogens of agronomic importance, as well as in phytotoxic and hemolytic activity. Among them, the peptides BP178, BP188 and BP192 having a C-terminal KDEL extension exhibited optimal features such as strong antibacterial and bactericidal effects, and low hemolytic and phytotoxic activities. They were selected for expression in transgenic rice.

2- Transgenic rice plants were generated expressing a codon-optimized *AMP*, namely *BP178*, *BP188*, *BP192* and *cecropin A* all of them with a C-terminal KDEL extension, and the *cecropin A* under the control of a seed-specific promoter (*glutelin B-1*, *glutelin B-4*, *26 kDa α-globulin* or *18 kDa oleosin*).

3- Most transgenic plants were morphologically similar to wild-type plants with respect to vegetative growth, flowering behaviour and seed yield. Even though, the lines expressing the *BP188_KDEL* under the control of the *glutelin B-4* promoter and the *cecropin A* under the control of either the *glutelin B-4* or the *18 kDa oleosin* promoter showed more vigorous shoot and root system as well as increased seed yield compared to wild-type plants.

4- The *BP178, BP188*, *BP192* and *cecropin A* all of them with a C-terminal KDEL extension, and the *cecropin A* transgenes were successfully expressed in rice seeds. The optimized subcellular fractionation in combination with Western blot and LC-MS/MS analyses confirmed that the BP178_KDEL, BP192_KDEL, Cec A_KDEL and Cec A peptides accumulated in the protein

bodies of the endosperm. Equally, the oleosin_Cec A fusion protein accumulated in the embryo oil bodies of rice seeds. Their accumulation in rice seeds was maintained stable for two-three years, even when stored at room temperature. In the case of the BP188_KDEL plants, although transgene expression was confirmed, the accumulation of the transgene product could not be verified.

5- The biological activity of the AMPs produced in transgenic rice was confirmed. Generally, the expression of the *BP134der* and *CecAder* transgenes under the control of a rice seed-specific promoter results in higher levels of protection against the bacterial phytopathogen *Dickeya* sp. 1552.10.1.

6- The recombinant AMP peptides were recovered from seeds. The BP178_KDEL peptide was recovered from the protein body-enriched fraction by a single purification step whereas the cecropin A peptide was partially purified using Weak Cation Exchange-Reverse Phase chromatography. Multimerization of the cecropin A peptide accumulating in rice seeds was observed. Moreover, the cecropin A peptide appears to interact with host proteins in the protein body fraction, most probably glutelins. Cleavage of the oleosin_CecA fusion protein (produced in the oil bodies) with the TEV protease, released a S-Cec A peptide which was then purified using Weak Cation Exchange-Reverse Phase chromatography. The S-Cec A peptide exhibited a high antibacterial and bactericidal activity against *Dickeya* sp. 1552.10.1.

7- The accumulation levels of AMP peptides in seeds were dependent on the strategy used. Seeds of the transgenic lines expressing the *18 kDa oleosin_Cec A* fusion gene under the control of the *18 kDa oleosin* promoter (*Ole18::Ole18_PRS_Cec A*, R9 lines) accumulated up to 10.8 pmols of a Cec A peptide/mg of seed (oil body strategy). Concerning the transgenic plants expressing the *Cec A* or the *Cec A_KDEL* gene under the control of the *GluB-4* or the *GluB-1* promoter, respectively (*GluB-4::Cec A*, R10; *GluB-1::Cec A_KDEL*, MC3), accumulated 1.3, 1.4 pmols AMP derivative/mg of seed in each case (protein body strategy). The rice plants expressing the

BP178_KDEL gene under the control of the *Glb-1* promoter (*Glb-1::BP178_KDEL*, R11) accumulated 1.8 pmols of an AMP derivative/mg of seed (protein body strategy). Regarding to the Cec A production, the highest accumulation level was accomplished when the Cec A peptide was expressed as a fusion protein with the oleosin on the oil body. As for AMP derivatives, the production of BP178 in the rice seed endosperm was found to be the most convenient strategy, as its recovery from transgenic rice seeds requires minimum purification and processing steps.

8- Taking together the results obtained in the present work it can be finally concluded that the production of AMPs in seeds of transgenic rice plants offers a promising perspective for their use as biofactories. The endosperm tissue of rice seed appears to be the most appropriate platform for the production of biologically active AMP derivatives, because there is no need of downstream processing steps to obtain the AMP. The strategy for production of Cec A in oil bodies allows a high level of production of cecropin A, but requires complex processes for its purification.

7. REFERENCES

A

Aerts, A.M., François, I.E., Meert, E.M., Li, Q.T., Cammue, B.P. and Thevissen, K. 2007. The antifungal activity of RsAFP2, a plant defensin from *Raphanus sativus*, involves the induction of reactive oxygen species in *Candida albicans*. J Mol Microbiol Biotechnol. 13:243-247.

Agerberth, B., Lee, J.Y., Bergman, T., Carlquist, M., Boman, H. G., Mutt, V. and Jörnvall, H. 1991. Amino acid sequence of PR-39. Isolation from pig intestine of a new member of the family of proline-arginine-rich antibacterial peptides.Eur J Biochem. 202:849-854.

Agrios, G.N. 2005. Plant Pathology, Academic Press, London.

Ahmad, A., Pereira, E.O., Conley, A.J., Richman, A.S. and Menassa, R. 2010. Green biofactories: recombinant protein production in plants. Recent Pat Biotechnol. 4:242-259.

Alan, A.R. and Earle, E.D. 2002. Sensitivity of bacterial and fungal plant pathogens to the lytic peptides, MSI-99, magainin II, and cecropin B. Mol Plant Microbe Interact. 15:701-708.

Ali, G.S. and Reddy, A.S. 2000. Inhibition of fungal and bacterial plant pathogens by synthetic peptides: in vitro growth inhibition, interaction between peptides and inhibition of disease progression. Mol Plant Microbe Interact. 13:847-859.

Allefs, S.J. H.M., De Jong E.R., Florak, D.E.A. and Hoogendoorn, C. 1996. *Erwinia* soft rot resistance of potato cultivars expressing antimicrobial peptide tachyplesin I. Mol Breed. 2:97-105.

Almeida, M.S., Cabral, K.M., Zingali, R.B. and Kurtenbach, E. 2000. Characterization of two novel defense peptides from pea (*Pisum sativum*) seeds. Arch Biochem Biophys. 378:278-286.

Anderssen, E.L., Diep, D.B., Nes, I.F., Eijsink, V.G. and Nissen-Meyer, J. 1998. Antagonistic activity of Lactobacillus plantarum C11: two new two-peptide bacteriocins, plantaricins EF and JK, and the induction factor plantaricin A. Appl Environ Microbiol. 64:2269-2272.

Andreu D, Rivas L. 1998. Animal antimicrobial peptides: an overview. Biopolymers. 47:415-433.

Arce, P., Moreno, M., Gutierrez, M., Gebauer, M., Dell'Orto, P., Torres, H., Acuña, I, Oliger, P., Venegas, A., Jordana, X., Kalazich, J. and Holuigue, L. 1999. Enhanced resistance to bacterial infection by *Erwinia carotovora* subsp. *atroseptica* in transgenic potato plants expressing the attacin or the cecropin SB-37 genes. American Journal of Potato Res. 76:169-177.

Arenas, G., Marshall, S., Espinoza, V., Ramírez, I. and Peña-Cortés, H. 2006. Protective effect of an antimicrobial peptide from *Mytilus edulis* chilensis expressed in *Nicotiana tabacum* L. Electron J Biotechnol. 9:2.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (eds.). 1998. Current Protocols in Molecular Biology. John Wiley & Sons, New York. ISBN: 978-0-471-50338-5.

B

Badea, A. Eudes, F., Laroche, A., Graf, R., Doshi, K., Amundsen, E., Nilsson, D. and Puchalski, B. 2013. Antimicrobial peptides expressed in wheat reduce susceptibility to *Fusarium* head blight and powdery mildew. Can J Plant Sci. 93:199-208.

Badosa, E., Ferre, R., Planas, M., Feliu, L., Besalú, E., Cabrefiga, J., Bardají, E. and Montesinos, E. 2007. A library of linear undecapeptides with bactericidal activity against phytopathogenic bacteria. Peptides. 28:2276-2285.

Badosa, E., Ferré, R., Frances, J., Bardají, E., Feliu, L., Planas, M., Montesinos, E. 2009. Sporicidal activity of synthetic antifungal undecapeptides and control of *Penicillium* rot of apples. Appl Environ Microbiol. 75:5563-5569.

Badosa, E., Moiset, G., Montesinos, L., Talleda, M., Bardají, E., Feliu, L., Planas, M. and Montesinos, E. 2013. Derivatives of the antimicrobial peptide BP100 for expression in plant systems. Plos One. 8:e85515.

Bagella, L., Scocchi, M. and Zanetti, M. 1995.cDNA sequences of three sheep myeloid cathelicidins. FEBS Lett. 376:225-228.

Banzet, N., Latorse, M.P, Bulet, P., François, E., Derpierre, C. and Dubald, M. 2002. Expression of insect cystein-rich antifungal peptides in transgenic tobacco enhances resistance to a fungal disease. Plant Sci. 162:995-1006.

Barbante, A., Irons, S., Hawes, C., Frigerio, L., Vitale, A. and Pedrazzini, E. 2008. Anchorage to the cytosolic face of the endoplasmic reticulum membrane: a new strategy to stabilize a cytosolic recombinant antigen in plants. Plant Biotechnol J. 6:560-575.

Barta, A., Sommergruber, K., Thompson, D., Hartmuth, K., Matzke, M.A. and Matzke, A.J. 1986. The expression of a *nopaline synthase* - human growth hormone chimaeric gene in transformed tobacco and sunflower callus tissue. Plant Mol Biol. 6:347-357.

Bechinger, B. 2004. Structure and function of membrane-lytic peptides. Crit Rev Plant Sci. 23:271-292.

Bechtel, D.B. and Juliano, B.O. 1980. Formation of Protein Bodies in the Starchy Endosperm of Rice (*Oryza sativa* L.): A Re-investigation. Ann Bot. 45:503-509.

Bellamy, W., Takase, M., Wakabayashi, H., Kawase, K., Tomita, M. 1992. Antibacterial spectrum of lactoferricin B, a potent bactericidal peptide derived from the N-terminal region of bovine lactoferrin. J Appl Bacteriol. 73:472-479.

Benchabane, M., Goulet, C., Rivard, D., Faye, L., Gomord, V. and Michaud, D. 2008. Preventing unintended proteolysis in plant protein biofactories. Plant Biotechnol. J. 6:633-648.

Berry, C.L., Brassinga, A.K., Donald, L.J., Fernando, W.G., Loewen, P.C. and de Kievit, T.R. 2012. Chemical and biological characterization of sclerosin, an antifungal lipopeptide. Can J Microbiol. 58:1027-1034.

Bhargava, A., Osusky, M., Hancock, E.R., Forward, B.S., Kay, W.W. and Misra, S. 2007. Antiviral indolicidin variant peptides: Evaluation for broad-spectrum disease resistance in transgenic *Nicotiana tabacum*. Plant Sci. 172:515-523.

Bhat, S.R. and Srinivasan, S. 2002. Molecular and genetic analyses of transgenic plants: Considerations and approaches. Plant Sci. 163:673-681.

Birkemo, G.A., Lüders, T., Andersen, Ø., Nes, I.F. and Nissen-Meyer, J. 2003. Hipposin, a histone-derived antimicrobial peptide in Atlantic halibut (*Hippoglossus hippoglossus* L.). Biochim Biophys Acta. 1646:207-215.

Blondelle, S.E, Takahashi, E., Houghten, R.A. and Pérez-Payá, E. 1996.Rapid identification of compounds with enhanced antimicrobial activity by using conformationally defined combinatorial libraries. Biochem J. 313:141-147.

Blondelle, S.E., Pinilla, C. and Boggiano, C. 2003. Synthetic combinatorial libraries as an alternative strategy for the development of novel treatments for infectious diseases. Methods Enzymol. 369:322-344.

Boman, H.G. and Steiner, H. 1981. Humoral immunity in *Cecropia pupae*. Curr Top Microbiol Immunol. 94-95:75-91.

Boothe, J., Nykiforuk, C., Shen, Y., Zaplachinski, S., Szarka S., Kuhlman P., Murray E., Morck D. and Moloney M.M. 2010. Seed-based expression systems for plant molecular pharming. Plant Biotechnol. J. 8:588-606.

Broekaert, W.F., Mariën, W., Terras, F.R., De Bolle, M.F., Proost, P., Van Damme, J., Dillen, L., Claeys, M., Rees, S.B., Vanderleyden, J. and Cammue, B.P. 1992. Antimicrobial peptides from *Amaranthus caudatus* seeds with sequence homology to the cysteine/glycine-rich domain of chitin-binding proteins.Biochemistry. 31:4308-4314.

Brogden, K.A. 2005. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? Nat Rev Microbiol. 3:238-250.

Bulet, P. and Stöcklin, R. 2005. Insect antimicrobial peptides: structures, properties and gene regulation. Protein Pept Lett. 12:3-11.

C

Cabanos, C., Ekyo, A., Amari, Y., Kato, N., Kuroda, M., Nagaoka, S., Takaiwa, F., Utsumi, S. and Maruyama, N. 2013. High-level production of lactostatin, a hypocholesterolemic peptide, in transgenic rice using soybean A1aB1b as carrier. Transgenic Res. 22:621-629.

Cagampang, G.B., Perdon, A.A. and Juliano, B.O. 1976. Changes in salt-soluble proteins of rice during grain development. Phytochemistry. 15:1425-1429.

Caligari, P.D.S., Yapabandara, Y.M.H.B., Paul, E.M., Perret, J., Roger, P. and Dunwell, J.M. 1993. Field performance of derived generations of transgenic tobacco. Theor Appl Genet. 86:875-879.

Carmona, M.J., Molina, A., Fernandez, J.A., Lopez-Fando, J.J. and Garcia-Olmedo, F. 1993. Expression of the α-thionin gene from barley in tobacco confers enhanced resistance to bacterial pathogens. Plant J. 3:457-462.

Carrington, J.C. and Dougherty, W.G. 1988. A viral cleavage site cassette: identification of amino acid sequences required for tobacco etch virus polyprotein processing. Proc Natl Acad Sci U S A. 85:3391-3395.

Cary, J.W., Rajasekaran, K., Jaynes, J.M. and Cleveland TE. 2000. Transgenic expression of a gene encoding a synthetic antimicrobial peptide results in inhibition of fungal growth in vitro and in planta. Plant Sci. 154:171-181.

Casteels, P., Ampe, C., Jacobs, F., Vaeck, M. and Tempst, P. 1989. Apidaecins: antibacterial peptides from honeybees. EMBO J. 8:2387-2391.

Cavallarin, L., D. and San Segundo, B. 1998. Cecropin A-derived peptides are potent inhibitors of fungal plant pathogens. Mol Plant Microbe Interact. 11:218-27.

Cereghino, G.P., Cereghino, J.L., Ilgen, C. and Cregg, J.M. 2002. Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. Curr Opin Biotechnol. 13:329-332.

Chakrabarti, A., Ganapathi, T.R., Mukherjee, P.K. and Bapat, V.A. 2003. MSI-99, a magainin analogue, imparts enhanced disease resistance in transgenic tobacco and banana. Planta. 216:587-596.

Charlet, M., Chernysh, S., Philippe, H., Hetru, C., Hoffmann, J.A. and Bulet, P. 1996. Innate immunity.Isolation of several cysteine-rich antimicrobial peptides from the blood of a mollusc, *Mytilus edulis*. J Biol Chem. 271:21808-21813.

Chen, D.H., Chyan, C.L., Jiang, P.L., Chen, C.S. and Tzen, J.T. 2012. The same oleosin isoforms are present in oil bodies of rice embryo and aleurone layer while caleosin exists only in those of the embryo. Plant Physiol Biochem. 60:18-24.

Chen, F.C. and Kuehnle, A.R. Obtaining Transgenic *Anthurium* through *Agrobacterium*-mediated Transformation of Etiolated Internodes. 1996. J Amer Soc Hort Sci. 121:47-51.

Chen, G.Q. and Lin, J-T. 2010. Use of quantitative polymerase chain reaction for determining copy numbers of transgenes in *Lesquerella fendleri*. Am. J. Agric. Biol. Sci. 5: 415-421.

Chen, J., Falla, T. J., Liu, H., Hurst, M. A., Fujii, C. A., Mosca, D. A., Embree, J. R., Loury, D. J., Radel, P. A., Cheng Chang, C., Gu, L. and Fiddes, J. C. 2000. Development of protegrins for the treatment and prevention of oral mucositis: structure-activity relationshipsof synthetic protegrin analogues. Biopolymers. 55:88-98.

Chen, L., Marmey, P., Taylor, N.J., Brizard, J.P., Espinoza, C., D'Cruz, P., Huet, H., Zhang, S., de Kochko, A., Beachy, R.N. and Fauquet, C.M. 1998. Expression and inheritance of multiple transgenes in rice plants. Nat Biotechnol. 16:1060-1064.

Choi, M.S, Kim, Y.H., Park, H.M., Seo, B., Jung, J.K., Kim, S.T.,Shin, D.B., Yun, H.T., Choi, I.S, Kim, C.K. and Lee, J.Y. 2009. Expression of BrD1, a plant defensin from *Brassica rapa*, confers resistance against brown planthopper (*Nilaparvata lugens*) in transgenic rices. Mol Cells. 28:131-137.

Christou, P., Swain, W.F., Yang, N.S. and McCabe, D.E. 1989. Inheritance and expression of foreign genes in transgenic soybean plants. Proc Natl Acad Sci U S A. 86:7500-7504.

Christou, P. 1997. Rice transformation: bombardment. Plant Mol Biol. 35:197-203.

Chu, Y., Bhattacharya, A., Wu, C., Knoll, J.E. and Ozias-Akins, P. 2013. Improvement of peanut (*Arachis hypogaea* L.) transformation efficiency and determination of transgene copy number by relative quantitative real-time PCR. In Vitro Cell Dev Biol Plant. 49:266-275.

Chuang, R.L., Chen, J.C., Chu, J. and Tzen, J.T. 1996. Characterization of seed oil bodies and their surface oleosin isoforms from rice embryos. J Biochem. 120:74-81.

Coca, M., Bortolotti, C., Rufat, M., Peñas, G., Eritja, R., Tharreau, D., del Pozo, A.M., Messeguer, J. and San Segundo B. 2004. Transgenic rice plants expressing the antifungal AFP protein from *Aspergillus giganteus* show enhanced resistance to the rice blast fungus *Magnaporthe grisea*. Plant Mol Biol. 54:245-259.

Coca, M., Peñas, G., Gómez, J., Campo, S., Bortolotti, C., Messeguer, J. and Segundo B.S. 2006. Enhanced resistance to the rice blast fungus *Magnaporthe grisea* conferred by expression of a *cecropin A* gene in transgenic rice. Planta. 223:392-406.

Conley, A.J., Joensuu, J.J., Richman, A. and Menassa, R. 2011. Protein body inducing fusions for high-level production and purification of recombinant proteins in plants. Plant Biotechnol J. 9:419-433.

Cuthbertson, B.J., Shepard, E.F., Chapman, R.W. and Gross, P.S. 2002. Diversity of the penaeidin antimicrobial peptides in two shrimp species. Immunogenetics. 54:442-445.

Cuthbertson, B.J., Büllesbach, E.E. and Gross, P.S. 2006. Discovery of synthetic penaeidin activity against antibiotic-resistant fungi. Chem Biol Drug Des. 68:120-127.

D

DeGray, G., Rajasekaran, K., Smith, F., Sanford J. and Daniell H. 2001.Expression of an antimicrobial peptide via the chloroplast genome to control phytopathogenic bacteria and fungi.Plant Physiology. 127:852-862.

De la Fuente-Salcido, N., Salcedo-Hernández, R., Alanís-Guzmán, M.G., Bideshi, D.K. and Barboza-Corona, J.E. 2007. A new rapid fluorogenic method for measuring bacteriocin activity.J Microbiol Methods. 70:196-199.

De Preter, K., Speleman, F., Combaret, V., Lunec, J., Laureys, G., Eussen, B.H., Francotte, N., Board, J., Pearson, A.D., De Paepe, A., Van Roy, N. and Vandesompele, J. 2002. Quantification of MYCN, DDX1, and NAG gene copy number in neuroblastoma using a real-time quantitative PCR assay. Mod Pathol. 15:159-166.

De Smet, K. and Contreras, R. 2005. Human antimicrobial peptides: defensins, cathelicidins and histatins. Biotechnol Lett. 27:1337-1347.

Delaney, DE. 2002. Choice of crop species and development of transgenic product lines. pp 55-78. In: Hood, E.E. and Howard, J.A. (eds.). Plants as factories for protein production. Kluwer Academics Publishers, Dordrecht.

De Lucca, A.J. and Walsh, T.J. 1999. Antifungal peptides: novel therapeutic compounds against emerging pathogens. Antimicrob Agents Chemother. 43:1-11.

Déstefano-Beltrán, L. 1991. Introduction into tobacco plants of genes which encode some of the natural components of the humoral immune response of *Hyalophora cecropia*. PhD diss. La. State Univ., Baton Rouge.

Destoumieux, D., Munoz, M., Bulet, P. and Bachère, E. 2000. Penaeidins, a family of antimicrobial peptides from penaeid shrimp (Crustacea, Decapoda). Cell Mol Life Sci. 57:1260-1271.

Dougherty, W.G., Carrington, J.C., Cary, S.M. and Parks, T.D. 1988. Biochemical and mutational analysis of a plant virus polyprotein cleavage site. 7:1281-1287.

Dougherty, W.G. and Parks, T.D. 1989. Molecular genetic and biochemical evidence for the involvement of the heptapeptide cleavage sequence in determining the reaction profile at two Tobacco Etch Virus cleavage sites in cell-free assays. Virology. 172:145-155.

Du, T., Wang, Y., Hu, Q.X., Chen, J., Liu, S., Huang, W.J. and Lin, M.L. 2005. Transgenic *Paulownia* expressing *shiva-1* gene has increased resistance to paulownia witches' broom disease. Journal of Integrative Plant Biol. 47:1500-1506.

Duan, X., Li, X., Xue, Q., Abo-el-Saad, M., Xu, D. and Wu, R. 1996.Transgenic rice plants harboring an introduced potato proteinase inhibitor II gene are insect resistant. Nat Biotechnol. 14:494-498.

Duvick, J.P., Rood, T., Rao, A.G. and Marshak, D.R. 1992. Purification and characterization of a novel antimicrobial peptide from maize (*Zea mays* L.) kernels. J Biol Chem. 267:18814-18820.

E

Ehrenstein G and Lecar H. 1977.Electrically gated ionic channels in lipid bilayers. Q Rev Biophys. 10:1-34.

Ellepola, S.W., Choi, S.M. and Ma, C.Y. 2005. Conformational study of globulin from rice (*Oryza sativa*) seeds by Fourier-transform infrared spectroscopy. Int J Biol Macromol. 37:12-20.

Epand, R.M. and Epand, R.F. 2011. Bacterial membrane lipids in the action of antimicrobial agents. J Pept Sci. 17:298-305.

Erlendsson, L.S., Muench, M.O., Hellman, U., Hrafnkelsdóttir, S.M., Jonsson, A., Balmer, Y., Mäntylä, E. and Orvar, B.L. 2010. Barley as a green factory for the production of functional Flt3 ligand. Biotechnol J. 5:163-171.

F

Faye, L. and Gomord, V. 2010. Success stories in molecular farming- a brief review. Plant Biotechnol J. 8:525-528.

Fearing, P.L., Brown, D., Vlachos, D., Meghji, M. and Privalle, L. 1997. Quantitative analysis of CryIA (b) expression in Bt maize plants, tissues, and silage and stability of expression over successive generations. Mol Breed. 3:169-176.

Fehlbaum, P., Bulet, P., Michaut, L., Lagueux, M., Broekaert, W.F., Hetru, C. and Hoffmann, J.A. 1994. Insect immunity. Septic injury of Drosophila induces the synthesis of a potent antifungal peptide with sequence homology to plant antifungal peptides. J Biol Chem. 269:33159-33163.

Fernandez de Caleya, R., Gonzalez-Pascual, B., García-Olmedo, F. and Carbonero, P. 1972. Susceptibility of phytopathogenic bacteria to wheat purothionins *in vitro*. Appl Microbiol. 23:998-1000.

Ferré, R., Badosa, E., Feliu, L., Planas, M., Montesinos, E. and Bardají E. 2006. Inhibition of plant-pathogenic bacteria by short synthetic cecropin A-melittin hybrid peptides. Appl Environ Microbiol. 72:3302-3308.

Fischer, U., Kuhlmann, M., Pecinka, A., Schmidt, R. and Mette, M.F. 2008. Local DNA features affect RNA-directed transcriptional gene silencing and DNA methylation. Plant J. 53:1-10.

Flavell, R.B. 1994. Inactivation of gene expression in plants as a consequence of specific sequence duplication. Proc Natl Acad Sci U S A. 91:3490-3496.

François, I.E., De Bolle, M.F., Dwyer, G., Goderis, I.J., Woutors, P.F., Verhaert, P.D., Proost, P., Schaaper, W.M., Cammue, B.P and Broekaert, W.F. 2002. Transgenic expression in *Arabidopsis* of a polyprotein construct leading to production of two different antimicrobial proteins. Plant Physiol. 128:1346-1358.

Friedrich, C.L., Moyles, D., Beveridge, T.J. and Hancock, R.E. 2000. Antibacterial action of structurally diverse cationic peptides on gram-positive bacteria. Antimicrob Agents Chemother. 44: 2086–2092.

Fujiwara, Y., Aiki, Y., Yang, L., Takaiwa, F., Kosaka, A., Tsuji, N.M., Shiraki, K. and Sekikawa K. 2010. Extraction of human interleukin-10 from transgenic rice seeds.Protein Expr Purif. 72:125-130.

Fukuda, M., Wen, L., Satoh-Cruz, M., Kawagoe, Y., Nagamura, Y., Okita, T.W., Washida, H., Sugino, A., Ishino, S., Ishino, Y., Ogawa, M., Sunada, M., Ueda, T. and Kumamaru, T. 2013. A guanine nucleotide exchange factor for Rab5 proteins is essential for intracellular transport of the proglutelin from the Golgi apparatus to the protein storage vacuole in rice endosperm. Plant Physiol. 162:663-674.

Fukuta, S., Kawamoto, K., Mizukami, Y., Yoshimura, Y., Ueda, J. and Kanbe, M. 2012. Transgenic tobacco plants expressing antimicrobial peptide bovine lactoferricin show enhanced resistance to phytopathogens. Plant Biotechnol. 29:383-389.

Fukuzawa, N., Tabayashi, N., Okinaka, Y., Furusawa, R., Furuta, K., Kagaya, U. and Matsumura, T. 2010. Production of biologically active Atlantic salmon interferon in transgenic potato and rice plants. J Biosci Bioeng. 110:201-207.

G

Galili, G. 2004. ER-derived compartments are formed by highly regulated processes and have special functions in plants. Plant Physiol. 136:3411-3413.

Gallo, R.L., Kim, K.J., Bernfield, M., Kozak, C.A., Zanetti, M., Merluzzi, L. and Gennaro, R. 1997. Identification of CRAMP, a cathelin-related antimicrobial peptide expressed in the embryonic and adult mouse. J Biol Chem. 272:13088-13093.

Gao, A.G, Hakimi, S.M., Mittanck, C.A., Wu, Y., Woerner, B.M., Stark, D.M., Shah, D.M., Liang, J. and Rommens, C.M. 2000. Fungal pathogen protection in potato by expression of a plant defensin peptide. Nat Biotechnol. 18:1307-1310.

García-Olmedo, F., Molina, A., Alamillo, J.M. and Rodríguez-Palenzuela, P. 1998. Plant defense peptides. Biopolymers. 47:479-491.

Ge, X., Chen, J., Li, N., Lin, Y., Sun, C. and Cao, K. 1999. Resistance function of rice lipid transfer protein LTP110. J Biochem Mol Biol. 36:603-607.

Gibbings, J.G., Cook, B.P., Dufault, M.R., Madden, S.L., Khuri, S., Turnbull, C.J. and Dunwell, J.M. 2003. Global transcript analysis of rice leaf and seed using SAGE technology. Plant Biotechnol J. 1:271-285.

Giddings, G., Allison, G., Brooks, D. and Carter, A. 2000. Transgenic plants as factories for biopharmaceuticals. Nature Biotechnology. 18:1151-1155.

Giuliani, A., Pirri, G. and Nicoletto, S.F. 2007. Antimicrobial peptides: an overview of a promising class of therapeutics. Cent Eur J Biol.2:1-33.

Glukhov, E., Stark, M., Burrows, L.L. and Deber, C.M. 2005. Basis for selectivity of cationic antimicrobial peptides for bacterial versus mammalian membranes. J Biol Chem. 280:33960-33967.

Goff, S.A., Ricke, D., Lan, T.H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., Hadley, D., Hutchison, D., Martin, C., Katagiri, F., Lange, B.M., Moughamer, T., Xia, Y., Budworth, P., Zhong, J., Miguel, T., Paszkowski, U., Zhang, S., Colbert, M., Sun, W.L., Chen, L., Cooper, B., Park, S., Wood, T.C., Mao, L., Quail, P., Wing, R., Dean, R., Yu, Y., Zharkikh, A., Shen, R., Sahasrabudhe, S., Thomas, A., Cannings, R., Gutin, A., Pruss, D., Reid, J., Tavtigian, S., Mitchell, J., Eldredge, G., Scholl, T., Miller, R.M., Bhatnagar, S., Adey, N., Rubano, T., Tusneem, N., Robinson, R., Feldhaus, J., Macalma, T., Oliphant, A. and Briggs, S. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science. 296:92-100.

Gómez-Ariza, J., Campo, S., Rufat, M., Estopà, M., Messeguer, J., San Segundo, B. and Coca, M. 2007. Sucrose-mediated priming of plant defense responses and broad-spectrum disease resistance by overexpression of the maize pathogenesis-related PRms protein in rice plants. Mol Plant Microbe Interact. 20:832-842.

Gomord, V. and Faye, L. 2004. Posttranslational modification of therapeutic proteins in plants. Curr Opin Plant Biol. 7: 171-178.

Gomord, V., Fitchette, A.C., Menu-Bouaouiche, L., Saint-Jore-Dupas, C., Plasson, C., Michaud, D. and Faye L. 2010. Plant-specific glycosylation patterns in the context of therapeutic protein production. Plant Biotechnol. J. 8:564-587.

Goto, F., Yoshihara, T., Shigemoto, N., Toki, S. and Takaiwa, F. 1999. Iron fortification of rice seed by the soybean ferritin gene. Nat Biotechnol. 17:282-286.

Guan, Z.J., Guo, B., Huo, Y.L., Guan, Z.P., Dai, J.K. and Wei, Y.H. 2013.Recent advances and safety issues of transgenic plant-derived vaccines.Appl Microbiol Biotechnol. 97:2817-2840.

Güell, I., Cabrefiga, J., Badosa, E., Ferre, R., Talleda, M., Bardají, E., Planas, M., Feliu, L. and Montesinos, E.2011. Improvement of the efficacy of linear undecapeptides against plant-pathogenic bacteria by incorporation of D-amino acids. Appl Environ Microbiol. 77:2667-2675.

Güell, I., Micaló, L., Cano, L., Badosa, E., Ferre, R., Montesinos, E., Bardají, E., Feliu, L. and Planas, M. 2012. Peptidotriazoles with antimicrobial activity against bacterial and fungal plant pathogens. Peptides. 33:9-17.

Gun Lee, D., Shin, S.Y., Maeng, C.Y., Jin, Z.Z., Kim, K.L. and Hahm, K.S. 1999.Isolation and characterization of a novel antifungal peptide from *Aspergillus niger*. Biochem Biophys Res Commun. 263:646-651.

Günter Blobel, 1999 (Novel price in Physiology or Medicine:*"for the discovery that proteins have intrinsic signals that govern their transport and localization in the cell".*

H

Hadjicharalambous, C., Sheynis, T., Jelinek, R., Shanahan, M.T., Ouellette, A.J. and Gizeli, E. 2008. Mechanisms of alpha-defensin bactericidal action: comparative membrane disruption by Cryptdin-4 and its disulfide-null analogue.Biochemistry. 47: 12626-12634.

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol. 166:557-580.

Hancock, R.E.W., Brown, M.H. and Piers, K. 1992. Cationic peptides and method of preparation. US Patent application serial No. 07/913, 492, filed August 21, 1992.

Hancock, R.E. and Lehrer, R. 1998. Cationic peptides: a new source of antibiotics. Trends Biotechnol. 16:82-88.

Hancock R.E. Host defence (cationic) peptides: what is their future clinical potential? 1999. Drugs. 57:469-473.

Hancock, R.E. and Diamond, G.The role of cationic antimicrobial peptides in innate host defences. 2000. Trends Microbiol. 8:402-410.

Hancock, R.E. 2001. Cationic peptides: effectors in innate immunity and novel antimicrobials. Lancet Infect Dis. 1:156-164.

Hancock, R.E. and Sahl, H.G. 2006. Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. Nat Biotechnol. 24:1551-1557.

Harashima, S. 1994. Heterologous protein production by yeast host-vector systems. Bioprocess Tecnol. 19:137-158.

Harder, J., Bartels, J., Christophers, E. and Schröder, J.M. 1997. A peptide antibiotic from human skin. Nature. 387(6636):861.

Harwig, S.S., Swiderek, K.M., Kokryakov, V.N., Tan, L., Lee, T.D., Panyutich, E.A., Aleshina, G.M., Shamova, O.V. and Lehrer, R.I. 1994. Gallinacins: cysteine-rich antimicrobial peptides of chicken leukocytes. FEBS Lett. 342:281-285.

Hashizume, F., Hino, S., Kakehashi, M., Okajima, T., Nadano, D., Aoki N. and Matsuda T. 2008.Development and evaluation of transgenic rice seeds accumulating a type II-collagen tolerogenic peptide.Transgenic Res. 17:1117-1129.

He, Y., Ning, T., Xie, T., Qiu, Q., Zhang, L., Sun, Y., Jiang, D., Fu, K., Yin, F., Zhang, W., Shen, L., Wang, H., Li, J., Lin, Q., Sun, Y., Li, H., Zhu, Y. and Yang, D. 2011. Large-scale production of functional human serum albumin from transgenic rice seeds. Proc. Natl. Acad. Sci. USA. 108:19078-19083.

He, Y., Wang, S. and Ding, Y. 2013. Identification of novel glutelin subunits and a comparison of glutelin composition between *japonica* and *indica* rice (*Oryza sativa* L.). J Cereal Sci. 57:362-371.

Hibino, T., Kidzu, K., Masumura, T., Ohtsuki, K., Tanaka, K., Kawabata, M. and Fujii, S. 1989. Amino acid composition of rice prolamin polypeptides. Agri Biol Chem. 53:513-518.

Hiei, Y., Ohta, S., Komari, T. and Kumashiro, T. 1994. Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. Plant J. 6:271-282.

Hiei, Y., Komari, T. and Kubo T. 1997.Transformation of rice mediated by *Agrobacterium tumefaciens*. Plant Mol Biol. 35:205-218.

Higo, K., Ugawa, Y., Iwamoto, M. and Korenaga, T. 1999. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Res. 27:297-300.

Hobbs, S.L., Kpodar, P. and DeLong, C.M. 1990. The effect of T-DNA copy number, position and methylation on reporter gene expression in tobacco transformants. Plant Mol Biol. 15:851-864.

Hofbauer, A. and Stoger, E. 2013. Subcellular accumulation and modification of pharmaceutical proteins in different plant tissues. Curr Pharm Des. 19:5495-502.

Hood, E.E., Gelvin, S.B., Melchers, S. and Hoekema, A. 1993. New *Agrobacterium* helper plasmids for gene transfer to plants (EHA105). Trans Res. 2:208-218.

Hood, E.E., Witcher, D.R., Maddock, S., Meyer, T., Baszczynski, C., Bailey, M., Flynn, P., Register, J., Marshall, L., Bond, D., Kulisek, E., Kusnadi, A., Evangelista, R., Nikolov, Z., Wooge, C., Mehigh, R.J., Hernan, R., Kappel, W.K., Ritland, D., Ping Li, C. and Howard, J.A. 1997. Commercial production of avidin from transgenic maize: characterization of transformant, production, processing, extraction and purification. Mol Breed. 3:291-306.

Hood, E.E., Woodard, S.L. and Horn, M.E. 2002. Monoclonal antibody manufacturing in transgenic plants-myths and realities. Curr Opin Biotechnol. 13:630-635.

Horn, M.E., Woodard, S.L. and Howard, J.A. 2004. Plant molecular farming: systems and products. Plant Cell Rep. 22:711-720.

Houston, D.F. and Mohammad, A. 1970. Purification and partial characterization of a major globulin from rice endosperm. Cereal Chem. 47:5-12.

Hsieh, K. and Huang, A.H. 2004. Endoplasmic reticulum, oleosins, and oils in seeds and tapetum cells. Plant Physiol. 136:3427-3434.

Huang, N. 2004. High-level protein expression system uses self-pollinating crops as hosts. BioProcess International. 2:54-59.

Huang, H.W. 2006. Molecular mechanism of antimicrobial peptides: the origin of cooperativity. Biochim Biophys Acta. 1758:1292-1302.

Huang, A.H.C. 2011. Subcellular oil droplets and oleosins in plants. The AOCS Lipid Library.

Hubert, F., Noel, T. and Roch, P. 1996. A member of the arthropod defensin family from edible Mediterranean mussels (*Mytilus galloprovincialis*). Eur J Biochem. 240:302-306.

I

Imamura, T., Yasuda, M., Kusano, H., Nakashita, H., Ohno, Y., Kamakura, T., Taguchi, S. and Shimada, H. 2010. Acquired resistance to the rice blast in transgenic rice accumulating the antimicrobial peptide thanatin. Transgenic Res. 19:415-424.

Ingelbrecht, I., van Houdt, H., van Montagu, M. and Depicker, A. Posttranscriptional silencing of reporter transgenes in tobacco correlates with DNA methylation. Proc Natl Acad Sci U S A. 91:10502-10506.

Ingham, D.J., Beer, S., Money, S., and Hansen, G. 2001. Quantitative real-time PCR assay for determining transgene copy number in transformed plants. Biotechniques. 31:132-140.

Ingham, A.B. and Moore, R.J. 2007. Recombinant production of antimicrobial peptides in heterologous microbial systems. Biotechnol Appl Biochem. 47:1-9.

J

Jabrane, A., Sabri, A., Compère, P., Jacques, P., Vandenberghe, I., Van Beeumen, J. and Thonart, P. 2002. Characterization of serracin P, a phage-tail-like bacteriocin, and its activity against *Erwinia amylovora*, the fire blight pathogen. Appl Environ Microbiol. 68:5704-5710.

Jan, P.S., Huang, H.Y., Cheng, H.M. 2010. Expression of a synthesized gene encoding cationic peptide cecropin B in transgenic tomato plants protects against bacterial diseases. Appl Biochem Biotechnol. 76:769-775.

Jaynes, J.M, Nagpala, P., Destefano-Beltran, L., Huang, J.H, Kim, J.H., Denny, T. and Cetiner, S. 1993. Expression of a Cecropin B lytic peptide analog in transgenic tobacco confers enhanced resistance to bacterial wilt caused by *Pseudomonas solanacearum*. Plant Sci. 89:43-53.

Jenssen, H., Hamill, P. and Hancock, R.E. 2006. Peptide antimicrobial agents. Clin. Microbiol. Rev. 19:491-511.

Jha S, Tank HG, Prasad BD, Chattoo BB. 2009. Expression of Dm-AMP1 in rice confers resistance to *Magnaporthe oryzae* and *Rhizoctonia solani*. Transgenic Res. 18:59-69.

Jha, S. and Chattoo, B.B. 2010. Expression of a plant defensin in rice confers resistance to fungal phytopathogens. Transgenic Res. 19:373-384.

Jiang, L., Yang, L., Zhang, H., Guo, J., Mazzara, M., Van den Eede, G. and Zhang, D. 2009. International collaborative study of the endogenous reference gene, sucrose phosphate synthase (SPS), used for qualitative and quantitative analysis of genetically modified rice. J Agric Food Chem. 57:3525-3532.

Joensuu, J.J., Brown, K., Conley, A.J., Clavijo, A., Menasse, R. and Brandle, J.E. 2008. Plant recombinant antibodies for the prevention of foot-and-mouth disease virus.In Vitro Cell Dev Biol Plant. 44:349-350.

Jones, A.M. and Herman, E.M. 1993. KDEL-Containing Auxin-Binding Protein Is Secreted to the Plasma Membrane and Cell Wall. Plant Physiol. 101:595-606.

Jones, R.W., Ospin-Giraldo, M.D. and Clemente, T. 2004. Prosystemin-antimicrobial-peptide fusion reduces tomato late blight lesion expansion. Mol Breed. 14:83-89.

Jorgensen, R.A., Cluster, P.D., English, J., Que, Q. and Napoli, C.A. Chalcone synthase cosuppression phenotypes in petunia flowers: comparison of sense vs. antisense constructs and single-copy vs. complex T-DNA sequences. Plant Mol Biol. 31:957-973.

Ju, Z.W., Hettiarachchy, N.S. and Rath, N. 2001. Extraction, denaturation and hydrophobic Properties of Rice Flour Proteins. J Food Sci. 66:229-232.

Juliano, B.O. 1985. Polysaccharides, protein, and lipids of rice, pp. 59-174. In: Juliano, B.O. (ed.). Rice: Chemistry and Technology, second ed. American Association of Cereal Chemists, Inc., St. Paul (USA) M.N.

Jung, S., Dingley, A.J., Augustin, R., Anton-Erxleben, F., Stanisak, M., Gelhaus, C., Gutsmann, T., Hammer, M.U., Podschun, R., Bonvin, A.M., Leippe, M., Bosch, T.C. and Grötzinger, J. 2009. Hydramacin-1, structure and antibacterial activity of a protein from the basal metazoan *Hydra*. J Biol Chem. 284:1896-1905.

Jung, Y.J., Lee, S.Y., Moon, Y.S. and Kang, K.K. 2012. Enhanced resistance to bacterial and fungal pathogens by overexpression of a human cathelicidin antimicrobial peptide (hCAP18/LL-37) in Chinese cabbage. Plant Biotechnol Rep. 6:39-46.

K

Kagawa, H., Hirano, H. and Kikuchi, F. 1988. Variation of glutelin seed storage protein in rice (Oryza sativa L). Jpn. J. Breed. 38: 327-332.

Kawagoe, Y., Suzuki, K., Tasaki, M., Yasuda, H., Akagi, K., Katoh, E., Nishizawa, N.K., Ogawa, M. and Takaiwa, F. 2005. The critical role of disulfide bond formation in protein sorting in the endosperm of rice. Plant Cell. 17:1141-1153.

Kaiserer L, Oberparleiter C, Weiler-Görz R, Burgstaller W, Leiter E, Marx F. 2003. Characterization of the *Penicillium chrysogenum* antifungal protein PAF. Arch Microbiol. 180:204-210.

Kalfa, V.C., Jia, H.P., Kunkle, R.A., McCray, P.B., Tack, B.F. and Brogden, K.A. 2001. Congeners of SMAP29 kill ovine pathogens and induce ultrastructural damage in bacterial cells. Antimicrob Agents Chemother. 45:3256-3261.

Kamysz, W., Krolicka, A., Bogucka, K., Ossowski, T., Lukasiak, J. and Lojkowska, E. 2005. Antibacterial activity of synthetic peptides against plant pathogenic *Pectobacterium* species. J Physiopathol. 153:313-317.

Kaplan, C.W., Sim, J.H., Shah, K.R., Kolesnikova-Kaplan, A., Shi, W. and Eckert, R. 2011. Selective membrane disruption: mode of action of C16G2, a specifically targeted antimicrobial peptide. Antimicrob Agents Chemother. 55:3446-3452.

Kathuria, H., Giri, J., Tyagi, H. and Tyagi, A.K. 2007.Advances in transgenic rice biotechnology.CRC Crit Rev Plant Sci. 26:65-103.

Kawagoe, Y., Suzuki, K., Tasaki, M., Yasuda, H., Akagi, K., Katoh, E., Nishizawa, N.K., Ogawa, M. and Takaiwa, F. 2005. The critical role of disulfide bond formation in protein sorting in the endosperm of rice.Plant Cell. 17:1141-1153.

Kawakatsu, T., Yamamoto, M.P., Hirose, S., Yano, M. and Takaiwa, F. 2008. Characterization of a new rice glutelin gene *GluD-1* expressed in the starchy endosperm. J Exp Bot. 59:4233-4245.

Kawakatsu, T., Hirose, S., Yasuda, H. and Takaiwa, F. 2010. Reducing rice seed storage protein accumulation leads to changes in nutrient quality and storage organelle formation. Plant Physiol. 154:1842-1854.

Khan, I., Twyman, R.M., Arcalis, E. and Stoger, E. 2012. Using storage organelles for the accumulation and encapsulation of recombinant proteins. Biotechnol J. 7:1099-1108.

Kilcoyne, M., Shah, M., Gerlach, J.Q., Bhavanandan, V., Nagaraj, V., Smith, A.D., Fujiyama, K., Sommer, U., Costello, C.E., Olszewski, N. and Joshi, L. 2009. O-glycosylation of protein subpopulations in alcohol-extracted rice proteins. J Plant Physiol. 166:219-232.

Kim, W.T. and Okita, T.W. 1988. Structure, expression, and heterogeneity of the rice seed prolamins. Plant Physiol. 88:649-655.

Kishimoto, T., Watanabe, M., Mitsui, T. And Hori, H. 1999. Glutelin basic subunits have a mammalian mucin-type O-linked disaccharide side chain. Arch Biochem Biophys. 370:271-277.

Ko, K., Norelli, J.L., Reynoird, J.P., Aldwinckle, H.S. and Brown, S.K. 2002. *T4 lysozyme* and *attacin* genes enhance resistance of transgenic 'Galaxy' apple against *Erwinia amylovora*. J Amer Soc Hort Sci. 127:515-519.

Komatsu, S. and Hirano, H. 1992. Rice seed globulin: a protein similar to wheat seed glutenin. Phytochemistry. 31:3455-3459.

Krause, A., Neitz, S., Mägert, H.J., Schulz, A., Forssmann, W.G., Schulz-Knappe, P. and Adermann, K. 2000. LEAP-1, a novel highly disulfide-bonded human peptide, exhibits antimicrobial activity. FEBS Lett. 480:147-150.

Kreil, G. 1973. Structure of melittin isolated from two species of honey bees. FEBS Lett. 33: 241-244.

Krishnan, H.B., Franceschi, V.R. and Okita, T.W. 1986. Immunochemical studies on the role of the Golgi complex in protein-body formation in rice cells. Planta. 169:471-480.

Krishnan, H.B. and Pueppke, S.G. 1993. Nucleotide sequence of an abundant rice seed globulin: homology with the high molecular weight glutelins of wheat, rye and triticale. Biochem Biophys Res Commun.193:460-466.

Kuehnle, A.R., Fujii, R., Chen, F.C., Alvarez, A., Sujii, N., Fukuri, R. and Aragon, S.L. 2004. Peptide biocides for engineering bacterial blight tolerance and susceptibility in cut-flower *Anthurium*. HortScience. 39:1327-1331.

Kumamaru, T., Satoh, H., Iwata, N., Omura, T., Ogawa, M. and Tanaka, K. 1988. Mutants for rice storage proteins. Screening of mutant for storage proteins in the starchy endosperm. Theor Appl Genet. 76:11-16.

Kuo, Y.C., Tan, C.C., Ku, J.T., Hsu, W.C., Su, S.C., Lu, C.A. and Huang, L.F. 2013. Improving Pharmaceutical Protein Production in *Oryza sativa*. Int. J. Mol. Sci. 14:8719-8739.

Kusnadi, A.R., Nikolov, Z.L. and Howard, J.A. 1997. Production of recombinant proteins in transgenic plants: Practical considerations. Biotechnol.Bioeng. 56:473-484.

L

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 227:680-685.

Lee, D.G., Park, Y., Jin, I., Hahm, K.S., Lee, H.H., Moon, Y.H. and Woo, E.R. 2004. Structure-antiviral activity relationships of cecropin A-magainin 2 hybrid peptide and its analogues. J Pept Sci. 10:298-303.

Lescot, M., Déhais, P., Moreau, Y., De Moor, B., Rouzé, P. and Rombauts S. 2002. PlantCARE: A database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Research.30:325-327.

Li, B., Shi, Y., Ibrahim, M., Liu, H., Shan, C., Wang, Y., Kube, M., Xie, G.L. and Sun, G. 2012. Genome sequence of the rice pathogen *Dickeya zeae* strain ZJU1202. J Bacteriol. 194:4452-4453.

Li, C., Blencke, H.M., Smith, L.C., Karp, M.T. and Stensvåg, K. 2010. Two recombinant peptides, SpStrongylocins 1 and 2, from *Strongylocentrotus purpuratus*, show antimicrobial activity against Gram-positive and Gram-negative bacteria. Dev Comp Immunol. 34:286-292.

Li, Q., Lawrence, C.B., Xing, H.Y., Babbitt, R.A., Bass, W.T., Maiti, I.B. and Everett, N.P. 2001. Enhanced disease resistance conferred by expression of an antimicrobial magainin analog in transgenic tobacco. Planta. 212:635-639.

Li, Q., Lawrence, C.B., Maelor Davies, H. and Everett, N.P. 2002. A tridecapeptide possesses both antimicrobial and protease-inhibitory activities. Peptides. 23:1-6.

Li, X., Wu, Y., Zhang, D.Z., Gillikin, J.W., Boston, R.S., Franceschi, V.R. and Okita, T.W. 1993. Rice prolamine protein body biogenesis: a BiP-mediated process. Science. 262:1054-1056.

Li, Z., Hansen, J.L., Liu, Y., Zemetra, R.S. and Berger, P.H. 2004. Using real-time PCR to determine transgene copy number in wheat. Plant Molecular Biol. Reporter. 22:179-188.

Liao, W., Ma, D., Liu, S. and Han, Z. 2009.mRNA cloning, tissue distribution and expression of duck avian Beta-defensin 9 in *E. coli*. Scientia Agricultura Sinica. 42: 1406-1412.

Lico, C., Desiderio, A., Banchieri, S. and Benvenuto, E. 2005.Plants as biofactories: Production of pharmaceutical recombinant proteins. pp 577-593. In: Tuberosa, R., Phillips, R.L. and Gale, M., (eds.). The wake of the double helix: from the green revolution to the gene revolution. Bologna: Avenue Media.

Lico, C., Santi, L., Twyman, R.M., Pezzotti, M. and Avesani L. 2012. The use of plants for the production of therapeutic human peptides. Plant Cell Rep. 31:439-451.

Lin C, Nie P, Lu W, Zhang Q, Li J, Shen Z. 2010. A selectively terminable transgenic rice line expressing human lactoferrin. Protein Expr Purif. 74:60-64.

Liu, Q., Ingersoll, J., Owens, S., Salih, R., Meng, R. And Hammerschlag, F. 2001. Response of transgenic “Royal Gala” apple (*Malus* x *domestica* Borkh.) shoots carrying a modified cecropin *MB39* gene, to *Erwinia amylovora*. Plant Cell Rpt. 20:306-312.

Liu, Q., Feng, Y., Zhao, X., Dong, H. and Xue Q. 2004. Synonymous codon usage bias in *Oryza sativa*. Plant Sci. 167:101-105.

López-García, B., Pérez-Payá, E. and Marcos, J.F. 2002. Identification of novel hexapeptides bioactive against phytopathogenic fungi through screening of a synthetic peptide combinatorial library. Appl. Environ. Microbiol. 68:2453-2460.

López-García, B., San Segundo, B. and Coca, M. 2012. Antimicrobial peptides for plant disease control. From discovery to application, pp. 263-294. In: Rajasekaran K, Cary J, Jaynes J, Montesinos E. (eds.). Small Wonders: Peptides for Disease Control. Volume 1095. American Chemical Society/ Oxford University Press Inc. Washington, USA.

M

Ma, D.Y., Liu, S.W., Han, Z.X., Li, Y.J. and Shan, A.S. 2008.Expression and characterization of recombinant gallinacin-9 and gallinacin-8 in *Escherichia coli*. Protein Expr Purif. 58:284-291.

Makovitzki, A., Viterbo, A., Brotman, Y., Chet, I. and Shai, Y. 2007. Inhibition of fungal and bacterial plant pathogens in vitro and in planta with ultrashort cationic lipopeptides. Appl Environ Microbiol. 73:6629-6636.

Malissard, M., Borsig, L., DiMarco, S., Grutter, M.G., Kragl, U., Wandrey, C. and Berger, E.G.1996. Recombinant soluble beta-1,4-galactosyltransferases expressed in *Saccharomyces cerevisiae*. Purification, characterization and comparison with human enzyme. Eur. J. Biochem. 239:340-348.

Mandard, N., Sy, D., Maufrais, C., Bonmatin, J.M., Bulet, P., Hetru, C., Vovelle, F. 1999. Androctonin, a novel antimicrobial peptide from scorpion *Androctonus australis*: solution structure and molecular dynamics simulations in the presence of a lipid monolayer. J Biomol Struct Dyn. 17:367-380.

Marcos, J.F., Muñoz, A., Pérez-Payá, E., Misra, S. and López-García, B. 2008. Identification and rational design of novel antimicrobial peptides for plant protection.Annu Rev Phytopathol. 46:273-301.

Marcus, J.P., Goulter, K.C., Green, J.L., Harrison, S.J. and Manners, J.M. 1997. Purification, characterisation and cDNA cloning of an antimicrobial peptide from *Macadamia integrifolia*. Eur J Biochem. 244:743-749.

Mason, G., Provero, P., Vaira, A.M. and Accotto, G.P. 2002. Estimating the number of integrations in transformed plants by quantitative real-time PCR. BMC Biotechnol. 2:20.

Masumura, T., Kidzu, K., Sugiyama, Y., Mitsukawa, N., Hibino, T., Tanaka, K. and Fujii, S. 1989. Nucleotide sequence of a cDNA encoding a major rice glutelin. Plant Mol. Biol. 12:723-725.

Matoba, N., Doyama, N., Yamada, Y., Maruyama, N., Utsumi, S. and Yoshikawa, M. 2001. Design and production of genetically modified soybean protein with anti-hypertensive activity by incorporating potent analogue of ovokinin (2-7). FEBS Lett. 497:50-54.

Matsuzaki, K., Sugishita, K., Fujii, N. and Miyajima, K. 1995. Molecular basis for membrane selectivity of an antimicrobial peptide, magainin 2, in phospholipid bilayers. Biochemistry. 34:3423-3429.

Mattanovich, D., Branduardi, P., Dato, L., Gasser, B., Sauer, M. and Porro, D. 2012. Recombinant protein production in yeasts. Methods Mol Biol. 824:329-358.

Melo, M.N., Ferre, R. and Castanho, M.A. 2009. Antimicrobial peptides: linking partition, activity and high membrane-bound concentrations. Nat Rev Microbiol. 7:245-250.

Mentag, R., Luckevich, M., Morency, M.J. and Séguin, A. 2003. Bacterial disease resistance of transgenic hybrid poplar expressing the synthetic antimicrobial peptide D4E1. Tree Physiol. 23:405-411.

Meyer, C.E. and Reusser, F. 1967. A polypeptide antibacterial agent isolated from *Trichoderma viride*.Experientia. 23:85-86.

Mitsukawa, N., Konishi, R., Kidzu, K., Ohtsuki, K., Masumura, T. and Tanaka, K. 1999. Amino acid sequencing and cDNA cloning of rice seed storage proteins, the 13 kDa prolamins, extracted from type I protein bodies. Plant Biotechnol. 16:103-113.

Moloney, M. 1997. Oil-body proteins as carriers of high-value peptides in plants. Patent US5650554 A.

Monroc, S., Badosa, E., Besalú, E., Planas, M., Bardají, E., Montesinos, E. and Feliu, L. 2006a. Improvement of cyclic decapeptides against plant pathogenic bacteria using a combinatorial chemistry approach. Peptides. 27:2575-2584.

Monroc, S., Badosa, E., Feliu, L., Planas, M., Montesinos, E. and Bardají, E. 2006b.De novo designed cyclic cationic peptides as inhibitors of plant pathogenic bacteria. Peptides. 27:2567-2574.

Montesinos, E. 2007. Antimicrobial peptides and plant disease control. FEMS Microbiol Lett. 270:1-11.

Montesinos, E. and Bardají, E. 2008. Synthetic antimicrobial peptides as agricultural pesticides for plant-disease control. Chem Biodivers. 5:1225-1237.

Montesinos, E., Badosa, E., Cabrefiga, J., Planas, M., Feliu, L. and Bardají, E. 2012. Antimicrobial peptides for plant disease control. From discovery to application, pp. 235-261. In: Rajasekaran K, Cary J, Jaynes J, Montesinos E. (eds.). Small Wonders: Peptides for Disease Control. Volume 1095. American Chemical Society/ Oxford University Press Inc.Washington, USA.

Mor, A., and Nicolas, P. 1994. Isolation and structure of novel defensive peptides from frog skin. Eur J Biochem. 219:145-154.

Morassutti, C., De Amicis, F., Skerlavaj, B., Zanetti, M. and Marchetti, S. 2002. Production of a recombinant antimicrobial peptide in transgenic plants using a modified VMA intein expression system. FEBS Lett. 22:141-146.

Moreno, A.B., Peñas, G., Rufat, M., Bravo, J.M., Estopà, M., Messeguer, J. and San Segundo, B. 2005. Pathogen-induced production of the antifungal AFP protein from *Aspergillus giganteus* confers resistance to the blast fungus *Magnaporthe grisea* in transgenic rice. Mol Plant Microbe Interact. 18:960-972.

Muench, D.G., Ogawa, M. and Okita T.W. 1999. The prolamins of rice. 93-108. In: Shewry P.R., Casey, R. (eds.). Seed proteins. Kluwer Academic Publishers, Dordrecht.

Muñoz, A., B. López-Garcia, and J. F. Marcos. 2006. Studies on the mode of action of the antifungal hexapeptide PAF 26. Antimicrob.Agents Chemother. 50:3847-3855.

Murray, M.G. and Thompson, W.F. 1980. Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. 8:4321-4325.

N

Nadal, A., Montero, M., Company, N., Badosa, E., Messeguer, J., Montesinos, L., Montesinos, E. and Pla M. Constitutive expression of transgenes encoding derivatives of the synthetic antimicrobial peptide BP100: impact on rice host plant fitness. BMC Plant Biol. 12:159.

Nagamine, A., Matsusaka, H., Ushijima, T., Kawagoe, Y., Ogawa, M., Okita, T.W. and Kumamaru, T. 2011.A role for the cysteine-rich 10 kDa prolamin in protein body I formation in rice. Plant Cell Physiol. 52:1003-1016.

Nakamura, T., Furunaka, H., Miyata, T., Tokunaga, F., Muta, T., Iwanaga, S., Niwa, M., Takao, T. and Shimonishi, Y. 1988. Tachyplesin, a class of antimicrobial peptide from the hemocytes of the horseshoe crab (*Tachypleus tridentatus*). Isolation and chemical structure.J Biol Chem. 263:16709-16713.

Nakase, M., Alvarez, A.M., Adachi, T., Aoki, N., Nakamura, R. and Matsuda, T. 1996. Immunochemical and biochemical identification of the rice seed protein encoded by cDNA clone A3-12. Biosci Biotechnol Biochem. 60:1031-1032.

Neuhoff, V., Arold, N., Taube, D. and Ehrhardt, W. 1988. Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. Electrophoresis. 9:255-262.

Nguyen, L.T., Schibli, D.J. and Vogel, H.J. 2005. Structural studies and model membrane interactions of two peptides derived from bovine lactoferricin. J Pept Sci. 11:379-389.

Ning, T., Xie, T., Qiu, Q., Yang, W., Zhou, S., Zhou, L., Zheng, C., Zhu, Y. and Yang D. 2008. Oral administration of recombinant human granulocyte-macrophage colony stimulating factor expressed in rice endosperm can increase leukocytes in mice. Biotechnol Lett. 30:1679-1686.

Nochi, T., Takagi, H., Yuki, Y., Yang, L., Masumura T., Mejima M., Nakanishi U., Matsumura A., Uozumi A., Hiroi T., Morita S., Tanaka K., Takaiwa F and Kiyono H. 2007. Rice-based mucosal vaccine as a global strategy for cold-chain and needle-free vaccination. Proc. Natl. Acad. Sci. USA. 104:10986-10991.

Nuttall J, Vine N, Hadlington JL, Drake P, Frigerio L, Ma JK. 2002. ER-resident chaperone interactions with recombinant antibodies in transgenic plants. Eur. J. Biochem. 269:6042-6051.

Nykiforuk, C.L., Boothe, J.G., Murray, E.W., Keon, R.G., Goren, H.J., Markley, N.A. and Moloney, M.M. 2006. Transgenic expression and recovery of biologically active recombinant human insulin from *Arabidopsis thaliana* seeds. Plant Biotechnol J. 4:77-85.

O

Obembe, O.O., Popoola, J.O., Leelavathi, S, and Reddy, S.V. 2011. Advances in plant molecular pharming.Biotechnol Adv. 29:210-222.

Oerke, E.C. 2006. Crop losses to pests. J Agric Sci. 144:31-43.

Ogawa, M., Kumamaru, T., satog, H., Iwata, N., Omura, T., Kasai, Z. and Tanaka, K. 1987. Purification of protein body-I of rice seed and its polypeptide composition. Plant Cell Physiol. 28:1517-1527.

Ogawa, M., Kumamaru, T., Satoh, H., Omura, T., Park, T., Shintaku, K. and Baba, K. 1989. Mutants for rice storage proteins. Theor Appl Genet. 78:305-310.

Oh, D., Shin, S.Y., Lee, S., Kang, J.H., Kim, S.D., Ryu, P.D., Hahm, K.S. and Kim, Y. 2000. Role of the hinge region and the tryptophan residue in the synthetic antimicrobial peptides, cecropin A (1-8)-magainin 2 (1-12) and its analogues, on their antibiotic activities and structures. Biochemistry. 39:11855-18864.

Ohshima, M., Mitsuhara, I., Okamoto, M., Sawano, S., Nishiyama, K., Kaku, H., Natori, S. and Ohashi, Y. 1999. Enhanced resistance to bacterial diseases of transgenic tobacco plants overexpressing sarcotoxin IA, a bactericidal peptide of insect. J Biochem. 125:431-435.

Okamoto, M., Mitsuhara, I., Ohshima, M., Natori, S. and Ohashi, Y. 1998. Enhanced expression of an antimicrobial peptide sarcotoxin IA by GUS fusion in transgenic tobacco plants. Plant Cell Physiol. 39:57-63.

Okamoto, T., Nakayama, H., Seta, K., Isobe, T. and Minamikawa, T. 1994. Posttranslational processing of a carboxy-terminal propeptide containing a KDEL sequence of plant vacuolar cysteine endopeptidase (SH-EP). FEBS Lett. 351:31-34.

Oldach, K.H., Becker, D. and Lörz, H. 2001. Heterologous expression of genes mediating enhanced fungal resistance in transgenic wheat. Mol Plant Microbe Interact. 14:832-838.

Oono, Y., Wakasa, Y., Hirose, S., Yang, L., Sakuta, C. and Takaiwa, F. 2010. Analysis of ER stress in developing rice endosperm accumulating beta-amyloid peptide. Plant Biotechnol J. 8:691-718.

Oren, Z. and Shai, Y. 1998. Mode of action of linear amphipathic alpha-helical antimicrobial peptides. Biopolymers. 47:451-463.

Osborn, R.W., De Samblanx, G.W., Thevissen, K., Goderis, I., Torrekens, S., Van Leuven, F., Attenborough, S., Rees, S.B. and Broekaert, W.F. 1995. Isolation and characterisation of plant defensins from seeds of Asteraceae, Fabaceae, Hippocastanaceae and Saxifragaceae. FEBS Lett. 368:257-262.

Osusky, M., Zhou, G., Osuska, L., Hancock, R.E., Kay, W.W. and Misra, S. 2000. Transgenic plants expressing cationic peptide chimeras exhibit broad-spectrum resistance to phytopathogens. Nat Biotechnol. 18:1162-1166.

Osusky, M., Osuska, L., Hancock, R.E., Kay, W.W. and Misra, S. 2004. Transgenic potatoes expressing a novel cationic peptide are resistant to late blight and pink rot. Transgenic Res. 13:181-190.

Osusky, M., Osuska, L., Kay, W. and Misra, S. 2005. Genetic modification of potato against microbial diseases: in vitro and in planta activity of a dermaseptin B1 derivative, MsrA2. Theor Appl Genet. 111:711-722.

Owens, L.D. and Heutte, T.M. 1997. A single amino acid substitution in the antimicrobial defense protein cecropin B is associated with diminished degradation by leaf intercellular fluid. Mol Plant Microbe Interact. 10:525-528.

P

Pan, S.J. and Reeck, G.R. 1988. Isolation and characterization of rice α-globulin. Cereal Chem. 65:316-319.

Parachin, N.S., Mulder, K.C., Viana, A.A., Dias, S.C. and Franco, O.L. 2012. Expression systems for heterologous production of antimicrobial peptides. Peptides. 38:446-456.

Park, C.H., Valore, E.V., Waring, A.J. and Ganz, T. 2001. Hepcidin, a urinary antimicrobial peptide synthesized in the liver. J Biol Chem. 276:7806-7810.

Park, I.Y., Park, C.B., Kim, M.S. and Kim, S.C. 1998. Parasin I, an antimicrobial peptide derived from histon H2A in the catfish*, Parasilurus asotus*. FEBS Lett. 437:258-262.

Park, Y.D, Papp, I., Moscone, E.A., Iglesias, V.A., Vaucheret, H., Matzke, A.J. and Matzke, M.A. 1996. Gene silencing mediated by promoter homology occurs at the level of transcription and results in meiotically heritable alterations in methylation and gene activity. Plant J. 9:183-194.

Parmenter, D.L., Boothe, J.G., van Rooijen, G.J., Yeung, E.C., Moloney, M.M. 1995. Production of biologically active hirudin in plant seeds using oleosin partitioning. Plant Mol. Biol. 29:1167-1180.

Pasupuleti, M., Schmidtchen, A. and Malmsten, M. 2012. Antimicrobial peptides: key components of the innate immune system. Crit Rev Biotechnol. 32:143-171.

Peach, C. and Velten, J. 1991. Transgene expression variability (position effect) of CAT and GUS reporter genes driven by linked divergent T-DNA promoters. Plant Mol Biol. 17:49-60.

Perdon, A.A. and Juliano B.O. 1978. Properties of a major α-globulin of rice endosperm. Phytochemistry. 17:351-353.

Phelan, R.W., Barret, M., Cotter, P.D., O'Connor, P.M., Chen, R., Morrissey, J.P., Dobson, A.D., O'Gara, F. and Barbosa, T.M. 2013. Subtilomycin: A New Lantibiotic from *Bacillus subtilis* Strain MMA7 Isolated from the Marine Sponge *Haliclona simulans*. Mar Drugs. 11:1878-1898

Ponti, D., Mangoni, M.L., Mignogna, G., Simmaco, M. and Barra, D. 2003. An amphibian antimicrobial peptide variant expressed in *Nicotiana tabacum* confers resistance to phytopathogens. Biochem J. 370:121-127.

Pridmore, D., Rekhif, N., Pittet, A.C., Suri, B. and Mollet, B. 1996. Variacin, a new lanthionine-containing bacteriocin produced by *Micrococcus* varians: comparison to lacticin 481 of *Lactococcus lactis*. Appl Environ Microbiol. 62:1799-1802.

Prodromou, C. and Pearl, L.H. 1992. Recursive PCR: a novel technique for total gene synthesis. Protein Eng. 5:827-829.

Pueyo, J.J., Chrispeels, M.J. and Herman, E.M. 1995. Degradation of transport-competent destabilized phaseolin with a signal for retention in the endoplasmic reticulum occurs in the vacuole. Planta. 196:586-596.

Q

Qu, L.Q. and Takaiwa, F. 2004. Evaluation of tissue specificity and expression strength of rice seed component gene promoters in transgenic rice. Plant Biotechnol J. 2:113-125.

Qu, L.Q., Xing, Y.P., Liu, W.X, Xu, X.P. and Song YR. 2008.Expression pattern and activity of six glutelin gene promoters in transgenic rice. J Exp Bot. 59:2417-2424.

R

Rajasekaran, K., Cary, J.W., Jaynes, J.M. and Cleveland, T.E. 2005. Disease resistance conferred by the expression of a gene encoding a synthetic peptide in transgenic cotton (*Gossypium hirsutum* L.) plants. Plant Biotechnol J. 3:545–554.

Ramanathan, V. and Veluthambi, K. 1995. Transfer of non-T-DNA portions of the *Agrobacterium tumefaciens* Ti plasmid pTiA6 from the left terminus of TL-DNA.Plant Mol Biol. 28:1149-1154.

Ramessar, K., Capell, T. and Christou, P. 2008. Molecular farming in cereal crops. Phytochem. Rev. 7:579-592.

Rao, A.G. 1999. Conformation and antimicrobial activity of linear derivatives of tachyplesin lacking disulfide bonds. Arch Biochem Biophys. 361:127-134.

Rathinakumar, R., Walkenhorst, W. F. and Wimley, W. C. 2009. Broad-spectrum antimicrobial peptides by rational combinatorial design and high-throughput screening: the importance of interfacial activity. J Am Chem Soc. 131:7609-7617.

Reddy, K.V., Yedery, R.D. and Aranha, C. 2004. Antimicrobial peptides: premises and promises. Int J Antimicrob Agents. 24:536-547.

Reed, J.D, Edwards, D.L. and Gonzalez C.F. 1997. Synthetic peptide combinatorial libraries: a method for the identification of bioactive peptides against phytopathogenic fungi. Mol Plant Microbe Interact. 10:537-549.

Reyes, R., Ramírez, G. and, Delgado, N.M. 2004.Fluorescent berberine binding as a marker of internal glycosaminoglycans sulfate in bovine oocytes and sperm cells. Arch. Androl. 50: 327-332.

Reynoird, J.P, Mourgues, F., Norelli, J., Aldwinckle, H.S., Brisset, M.N. and Chevreau, E. 1999. First evidence for improved resistance to fire blight in transgenic pear expressing the attacin E gene from *Hyalophora cecropia*, Plant Sci. 149:23-31.

Riggs, A.D. 1981. Bacterial production of human insulin.Diabetes Care. 4:64-68.

Rogers, L.A. 1928. The inhibiting effect of Streptococcus lactis on *Lactobacillus bulgaricus*. J Bacteriol. 16:321-325.

Roy-Barman, S., Sautter, C. and Chattoo, B.B. 2006. Expression of the lipid transfer protein *Ace*-AMP1 in transgenic wheat enhances antifungal activity and defense responses. Transgenic Res. 15:435-446.

Rybicki, E.P. 2010. Plant-made vaccines for humans and animals. Plant Biotechnol J. 8:620-637.

S

Sambrook, J. and Russel, D. 2001. Molecular cloning: A laboratory manual, 3nd ed. Cold Spring Harbour, N.Y.: Cold Spring Harbor Laboratory Press. ISBN 978-087969577-4.

Schägger, H. 2006. Tricine-SDS-PAGE. Nat Protoc. 1:16-22.

Saito, Y., Shigemitsu, T., Yamasaki, R., Sasou, A., Goto, F., Kishida, K., Kuroda, M., Tanaka, K., Morita, S., Satoh, S. and Masumura, T. 2012. Formation mechanism of the internal structure of type I protein bodies in rice endosperm: relationship between the localization of prolamin species and the expression of individual genes. Plant J. 70:1043-1055.

Sallaud, C., Meynard, D., van Boxtel, J., Gay, C., Bès, M., Brizard, J.P., Larmande, P., Ortega, D., Raynal, M., Portefaix, M., Ouwerkerk, P.B., Rueb, S., Delseny, M. and Guiderdoni E. 2003. Highly efficient production and characterization of T-DNA plants for rice (*Oryza sativa* L.) functional genomics. Theor Appl Genet. 106:1396-1408.

Salomón, R.A. and Farías, R.N. 1992. Microcin 25, a novel antimicrobial peptide produced by *Escherichia coli*. J Bacteriol.174:7428-7435.

Schaaper, W.M., Posthuma, G.A., Plasman, H.H., Sijtsma, L., Fant, F., Borremans, F.A., Thevissen, K., Broekaert, W.F., Meloen, R.H. and van Amerongen, A. 2001. Synthetic peptides derived from the beta2-beta3 loop of *Raphanus sativus* antifungal protein 2 that mimic the active site. J Pept Res. 57:409-418.

Schillberg, S., Zimmermann, S., Voss, A. and Fischer, R. 1999. Apoplastic and cytosolic expression of full-size antibodies and antibody fragments in Nicotiana tabacum. Transgenic Res. 8:255-263.

Schillberg, S., Fisher.R and Emans, N. 2003. Molecular farming of recombinant antibodies in plants. Cell. Mol. Life Sci. 60:433-445.

Selsted, M.E., Novotny, M.J., Morris, W.L., Tang, Y.Q., Smith, W. and Cullor, J.S. 1992. Indolicidin, a novel bactericidal tridecapeptide amide from neutrophils. J Biol Chem. 267:4292-4295.

Shai, Y. 2002. Mode of action of membrane active antimicrobial peptides.Biopolymers. 66:236-248.

Sharma, A., Sharma, R., Imamura, M., Yamakawa, M. and Machii, H. 2000. Transgenic expression of cecropin B, an antimicrobial peptide from *Bombyx mori,* confers enhanced resistance to bacterial leaf blight in rice. FEBS Lett 484:7-11.

Sharma, A.K., Jani, D. and Tyagi, A.K. 2004. Transgenic plants as bioreactors. Ind J Biotechnol. 3:274-390.

Sharma, A.K. and Sharma, M.K. 2009. Plants as bioreactors: Recent developments and emerging opportunities. Biotechnol Adv. 27:811-832.

Shewry, P.R., Naiper, J.A. and Tatham A.S. 1995. Seed Storage Proteins-Structures and Biosynthesis. Plant Cell. 7:945-956.

Shibata, K., Suzawa, T., Soga, S., Mizukami, T., Yamada, K., Hanai, N. and Yamasaki, M. 2003. Improvement of biological activity and proteolytic stability of peptides by coupling with a cyclic peptide. Bioorg Med Chem Lett. 13:2583-2586.

Shin, Y.J., Hong, S.Y., Kwon, T.H., Jang, Y.S. and Yang, M.S. 2003.High level of expression of recombinant human granulocyte-macrophage colony stimulating factor in transgenic rice cell suspension culture. Biotechnol Bioeng. 82:778-783.

Shorrosh, B.S., Wen, L., Zen, K.C., Huang, J.K., Pan, J.S., Hermodson, M.A., Tanaka, K., Muthukrishnan, S. and Reeck, G.R. 1992. A novel cereal storage protein: molecular genetics of the 19 kDa globulin of rice. Plant Mol Biol. 18:151-154.

Shou, H., Frame, B. R., Whitham, S.A. and Wang, K. 2004. Assessment of transgenic maize events produced by particle bombardment or *Agrobacterium*-mediated transformation. Mol Breed. 13:201-208.

Shunmann, P.H.D., Coia, G., and Whaterhouse, P.M. 2002. Biopharming the SimpliRED HIV diagnostic reagent in barley, potato and tobacco. Mol. Breed. 9:113-121.

Shyur, L. F., Wen, T. N. and Chen, C. S. 1994.Purification and characterization of rice prolamins. Bot. Bull. Acad. Sin. 35:65-71.

Silva, P.I., Daffre, S. and Bulet, 2000. Isolation and characterization of gomesin, an 18-residue cysteine-rich defense peptide from the spider *Acanthoscurria gomesiana* hemocytes with sequence similarities to horseshoe crab antimicrobial peptides of the tachyplesin family. J Biol Chem. 275:33464-33470.

Simmaco, M., Mignogna, G., Barra, D. and Bossa, F. 1993. Novel antimicrobial peptides from skin secretion of the European frog *Rana esculenta*. FEBS Lett. 324:159-161.

Simmaco, M., Mignogna, G., Canofeni, S., Miele, R., Mangoni, M.L. and Barra, D. 1996. Temporins, antimicrobial peptides from the European red frog *Rana temporaria*. Eur J Biochem. 15:788-792.

Song, P., Cai, C.Q., Skokut, M., Kosegi, B.D. and Petolino, J.F. 2002. Quantitative real-time PCR as a screening tool for estimating transgene copy number in WHISKERS™-derived transgenic maize. Plant Cell Rep., 20:948-954.

Song, Z., Lu B. and Chen, J. 2001. A study of pollen viability and longevity in *Oryza rufipogon*, *Oryza sativa* and their hybrid. Int Rice Res Notes. 26:31-42.

Southern, E.M. 1975.Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol. 98:503-517.

Srivastava, V., Vasil, V. and Vasil, I.K. 1996. Molecular characterization of the fate of transgenes in transformed wheat (*Triticum aestivumL*.). Theor Appl Genet. 92:1031-1037.

Stam, M., Mol, J.N.M. and Kooter, J.M. 1997. The silence of genes in transgenic plants. Ann Bot.79:3-12.

Stansly, P. G., Shepherd, R. G. and White, H. J. 1947. Polymyxin: a new chemotherapeutic agent. Bull Johns Hopkins Hosp. 81:43-54.

Steiner, H., Hultmark, D., Engström, A., Bennich, H. and Boman, H.G. 1981. Nature. Sequence and specificity of two antibacterial proteins involved in insect immunity. 292:246-248.

Stoger, E., Sack, M., Perrin, Y., Vaquero, C., Torres, E., Twyman, R.M., Christou, P. and Fisher, R. 2002. Practical considerations for pharmaceutical antibody production in different crop systems. Mol. Breeding. 9:149-158.

Stoger, E., Ma, J.K., Fisher, R. and Christou, P. 2005. Showing the seeds of success: pharmaceutical proteins from plants. Curr Opin Biotechnol. 16:167-173

Su, C.F., Kuo, I.C., Chen, P.W., Huang, C.H., Seow, S.V., Chua, K.Y. and Yu, S.M. 2012. Characterization of an immunomodulatory Der p 2-FIP-fve fusion protein produced in transformed rice suspension cell culture. Transgenic Res. 21:177-192.

Sugiarto, H. and Yu, P.L. 2006. Identification of three novel ostricacins: an update on the phylogenetic perspective of beta-defensins. Int J Antimicrob Agents. 27:229-235.

Suzuki, K., Kaminuma, O., Yang, L., Takai, T., Mori, A., Umezu-Goto, M., Ohtomo, T., Ohmachi, Y., Noda, Y., Hirose, S., Okumura, K., Ogawa, H., Takada, K., Hirasawa, M., Hiroi, T. and Takaiwa F. 2011. Prevention of allergic asthma by vaccination with transgenic rice seed expressing mite allergen: induction of allergen-specific oral tolerance without bystander suppression. Plant Biotechnol J. 9:982-990.

Synge, R.L. 1945. Gramicidin S: over-all chemical characteristics and amino-acid composition. Biochem J. 39:363-367.

T

Tailor, R.H., Acland, D.P., Attenborough, S., Cammue, B.P., Evans, I.J., Osborn, R.W., Ray, J.A., Rees, S.B. and Broekaert, W.F. 1997. A novel family of small cysteine-rich antimicrobial peptides from seed of *Impatiens balsamina* is derived from a single precursor protein. J Biol Chem. 272:24480-24487.

Takagi, H., Saito, S., Yang, L., Nagasaka, S., Nishizawa, N. and Takaiwa, F. 2005a. Oral immunotherapy against a pollen allergy using a seed-based peptide vaccine. Plant Biotechnol J. 3:521-533.

Takagi, H., Hiroi, T., Yang, L., Tada, Y., Yuki, Y., Takamura, K., Ishimitsu, R., Kawauchi, H., Kiyono, H. and Takaiwa F. 2005b. A rice-based edible vaccine expressing multiple T cell epitopes induces oral tolerance for inhibition of Th2-mediated IgE responses. Proc Natl Acad Sci U S A. 102:17525-17530.

Takagi H, Hiroi T, Yang L, Takamura K, Ishimitsu R, Kawauchi H, Takaiwa F. 2008. Efficient induction of oral tolerance by fusing cholera toxin B subunit with allergen-specific T-cell epitopes accumulated in rice seed.Vaccine. 26:6027-6030.

Takagi, H., Hiroi, T., Hirose, S., Yang, L. and Takaiwa, F. 2010.Rice seed ER-derived protein body as an efficient delivery vehicle for oral tolerogenic peptides.Peptides. 31:1421-1425.

Takaiwa, F., Kikuchi, S. and Oono, K. 1987. Major type of glutelin mRNA can be divided into two classes. Mol. Gen. Genet. 208:15-22.

Takaiwa, F., Oono, K., Wing, D. and Kato, A. 1991. Sequences of three members and expression of a new major subfamily of glutelin gene from rice. Plant Mol. Biol. 17:875-885.

Takaiwa, F., Yamanouchi, U., Yoshihara, T., Washida, H., Tanabe, F., Kato, A. and Yamada, K. 1996. Characterization of common cis-regulatory elements responsible for the endosperm-specific expression of members of the rice glutelin multigene family. Plant Mol Biol. 30:1207-1221.

Takaiwa, F., Ogawa, M. and Okita, T.W. 1999.Rice glutelins. pp 401-426. In: Seed Proteins (Swery, P and Casey R. (eds.). Dordrecht: Kluwer Academic Publishers.

Takaiwa, F., Takagi, H., Hirose, S., Wakasa, Y. 2007. Endosperm tissue is a good production platform for artificial recombinant proteins in transgenic rice. Plant Biotechnol. J. 5:84-92.

Takaiwa, F., Hirose, S., Takagi, H., Yang, L. and Wakasa, Y. 2009.Deposition of a recombinant peptide in ER-derived protein bodies by retention with cysteine-rich prolamins in transgenic rice seed.Planta. 229:1147-1158.

Tam, J.P., Lu, Y.A., Yang, J.L. and Chiu, K.W. 1999. An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides.Proc Natl Acad Sci U S A. 96:8913-8918.

Tanaka, K., Sugimoto, T., Ogawa, M. and Kasai Z. 1980. Isolation and characterization of two types of protein bodies in rice endosperm. Agric. Biol. Chem. 44:1633-1639.

Tasiemski, A., Schikorski, D., Le Marrec-Croq, F., Pontoire-Van Camp, C., Boidin-Wichlacz, C. and Sautière, P.E. 2007. Hedistin: A novel antimicrobial peptide containing bromotryptophan constitutively expressed in the NK cells-like of the marine annelid, *Nereis diversicolor*. Dev Comp Immunol. 31:749-762.

Terras, F.R., Eggermont, K., Kovaleva, V., Raikhel, N.V., Osborn, R.W., Kester, A., Rees, S.B., Torrekens, S., van Leuven, F., Vanderleyden, J., Cammue, B.P.A. and Broekaert W.F. 1995. Small cysteine-rich antifungal proteins from radish: their role in host defense. Plant Cell. 7:573-588.

Thomas, S., Karnik, S., Barai, R.S., Jayaraman, V.K.and Idicula-Thomas, S. 2010. CAMP: a useful resource for research on antimicrobial peptides. Nucl Acid Res. Vol.38, Database issue.

Tizaoui, K. and Kchouk, M.E. 2012. Genetic approaches for studying transgene inheritance and genetic recombination in three successive generations of transformed tobacco. Genet Mol Biol. 35:640-649.

Toke, O. 2005. Antimicrobial peptides: new candidates in the fight against bacterial infections. Biopolymers. 80:717-735.

Torres, E., Vaquero, C., Nicholson, L., Sack, M., Stöger E., Drossard, J., Christou, P., Fisher, R. and Perrin, Y. 1999.Rice cell culture as an alternative production system for functional diagnostic and therapeutic antibodies.Transgenic Res. 8:441-449.

Torres, E., Gonzalez-Melendi, P., Stöger, E., Shaw, P., Twyman, RM., Nicholson, L., Vaquero, C., Fischer, R. Christou P. and Perrin Y. 2001. Native and artificial reticuloplasmins co-accumulate in distinct domains of the endoplasmic reticulum and in post-endoplasmic reticulum compartments. Plant Physiol. 127:1212-1223.

Twyman, R.M., Stoger, E., Schillberg, S., Christou, P. and Fisher R. 2003. Molecular farming in plants: host systems and expression technology. Trends Biotechnol. 21:570-578.

Tzen, J.T. and Huang, A.H. 1992. Surface structure and properties of plant seed oil bodies. J Cell Biol. 117:327-335.

V

Vain, P., James, A., Worland, B. and Snape, W. 2002. Transgene behaviour across two generations in a large random population of transgenic rice plants produced by particle bombardment. Theor Appl Genet. 105:878-889.

van Roojien, G.J. and Moloney, M.M. 1995. Plant seed oil-bodies as carriers for foreign proteins.Biotechnology (NY). 13:72-77

Viana, A.A., Pelegrini, P.B. and Grossi-de-Sá, M.F. 2012. Plant biofarming: novel insights for peptide expression in heterologous systems.Biopolymers. 98:416-427.

Vidal, J.R., Kikkert, J.R., Malnoy, M.A., Wallace, P.G., Barnard, J. and Reisch, B.I. 2006. Evaluation of transgenic 'Chardonnay' (*Vitis vinifera*) containing magainin genes for resistance to crown gall and powdery mildew.Transgenic Res. 15:69-82.

Vila-Perelló, M., Sánchez-Vallet, A., García-Olmedo, F., Molina, A. and Andreu, D. 2003. Synthetic and structural studies on *Pyrularia pubera* thionin: a single-residue mutation enhances activity against Gram-negative bacteria. FEBS Lett. 536:215-219.

Villanueva, M.A. and Malek-Hedayat, S. 1987. Elimination of endogenous peroxidase artifacts in immunoblots of plant extracts. Plant Sci. 52:141-146.

Vitale, A. and Ceriotti, A. 2004. Protein quality control mechanisms and protein storage in the endoplasmic reticulum. A conflict of interests? Plant Physiol. 136:3420-3426.

von Heijne, G. 1983. Patterns of amino acids near signal-sequence cleavage sites. Eur J Biochem. 133:17-21.

von Heijne, G. 1985. Signal sequences. The limits of variation. J Mol Biol. 184:99-105.

Vutto, N.L., Gapeeva, T.A., Pundik, A.N., Tret'iakova, T.G. and Volotovskiĭ, I.D. 2010. Transgenic Belarussian-bred potato plants expressing genes for antimicrobial peptides of the cecropin-melittin type. Genetika. 46:1626-1634.

Vyacheslavova AO, Berdichevets IN, Tiurin AA, Shimshilashvili KhR, Mustafaev O, Goldenkova-Pavlova IV. 2012. Expression of heterologous genes in plant systems: new possibilities. Genetika. 11:1067-1079.

W

Wakasa, Y., Yasuda, H. and Takaiwa, F. 2006. High accumulation of bioactive peptide in transgenic rice seeds by expression of introduced multiple genes. Plant Biotechnol J. 4:499-510.

Wakasa, Y., Zhao, H., Hirose, S., Yamauchi, D., Yamada, Y., Yang, L., Ohinata, K., Yoshikawa, M. and Takaiwa, F. 2011a. Antihypertensive activity of transgenic rice seed containing an 18-repeat novokinin peptide

localized in the nucleolus of endosperm cells. Plant Biotechnol J. 9:729-735.

Wakasa, Y., Tamakoshi, C., Ohno, T., Hirose, S., Goto, T., Nagaoka, S. and Takaiwa F. 2011b. The hypocholesterolemic activity of transgenic rice seed accumulating lactostatin, a bioactive peptide derived from bovine milk β-lactoglobulin. J Agric Food Chem. 59:3845-3850.

Wakasa, Y., Takagi, H., Hirose, S., Yang, L., Saeki, M., Nishimura, T., Kaminuma, O., Hiroi, T. and Takaiwa, F. 2013. Oral immunotherapy with transgenic rice seed containing destructed Japanese cedar pollen allergens, Cry j 1 and Cry j 2, against Japanese cedar pollinosis. Plant Biotechnol J. 11:66-76.

Wang, G., Li, X. and Wang, Z. 2009. APD2: the updated antimicrobial peptide database and its application in peptide design. Nucleic Acids Res. 37:933-937.

Wang, G. 2010. Structure, dynamics and mapping of membrane-binding residues of micelle-bound antimicrobial peptides by natural abundance (13) C NMR spectroscopy. Biochim Biophys Acta. 1798:114-121.

Wang, S., Takahashi, H., Kajiura, H., Kawakatsu, T., Fujiyama, K. and Takaiwa F. 2013. Transgenic rice seeds accumulating recombinant hypoallergenic birch pollen allergen Bet v 1 generate giant protein bodies. Plant Cell Physiol. 54:917-933.

Wang, Y., Nowak, G., Culley, D., Hadwiger, L.A. and Fristensky, B. 1999. Constitutive Expression of Pea Defense Gene DRR206 Confers Resistance to Blackleg (*Leptosphaeria maculans*) Disease in Transgenic Canola (*Brassica napus*). Mol Plant Microbe Interact. 12:410-418.

Wang, Y., Hong, J., Liu, X., Yang, H., Liu, R., Wu, J., Wang, A., Lin, D. and Lai, R. 2008. Snake cathelicidin from *Bungarus fasciatus* is a potent peptide antibiotic. PLoS One. 3:e3217.

Wang, Y., Deng, H., Zhang, X., Xiao, H., Jiang, Y., Song, Y., Fang, L., Xiao, S., Zhen, Y. and Chen, H. 2009. Generation and immunogenicity of Japanese encephalitis virus envelope protein expressed in transgenic rice. Biochem Biophys Res Commun. 380:292-297.

Wang, Z. and Wang, G. 2004. APD: the Antimicrobial Peptide Database. Nucl Acid Res. Vol.32, No.1, pp. D590-D592.

Washida, H., Wu, C.Y., Suzuki, A., Yamanouchi, U., Akihama, T., Harada, K. and Takaiwa, F. 1999. Identification of cis-regulatory elements required for endosperm expression of the rice storage protein glutelin gene GluB-1. Plant Mol Biol. 40:1-12.

Wei, G.X. and Bobek, L.A. 2005. Human salivary mucin MUC7 12-mer-L and 12-mer-D peptides: antifungal activity in saliva, enhancement of activity with protease inhibitor cocktail or EDTA, and cytotoxicity to human cells. Antimicrob Agents Chemother. 49:2336-2342.

Wen, T.N. and Luthe, D.S. 1985.Biochemical characterization of rice glutelin. Plant Physiol. 78:172-177.

Weng, H., Pan, L., Yang, C., Zhang, C., Liu, Z. and Zhang, D. 2004. Estimating number of transgene copies in transgenic rapeseed by real-time PCR assay with HMG I/Y as an endogenous reference gene. Plant Mol Biol Rep. 22:289-300.

Winichayakul, S., Pernthaner, A., Livingston, S., Cookson, R., Scott, R. and Roberts N.2012. Production of active single-chain antibodies in seeds using trimeric polyoleosin fusion. J Biotechnol. 161:407-413.

Witcher, D.R., Hood, E.E., Peterson D., Bailey M., Bond, D., Kusnadi, A., Evangelista R., Nikolov, Z., Wooge, C., Mehigh, R., Kappel, W., Register, J. and Howard J.A. 1998. Commercial production of β-glucuronidase (GUS): a model system for the production of proteins in plants. Mol Breed. 4:301-312.

Wnendt, S., Ulbrich, N. and Stahl, U. 1990. Cloning and nucleotide sequence of a cDNA encoding the antifungal-protein of *Aspergillus giganteus* and preliminary characterization of the native gene. Nucleic Acids Res. 18:3987.

Wu, C., Washida, H., Onodera, Y., Harada, K. and Takaiwa, F. 2000. Quantitative nature of the Prolamin-box, ACGT and AACA motifs in a rice glutelin gene promoter: minimal cis-element requirements for endosperm-specific gene expression. Plant J. 23:415-421.

Wu, C.Y., Suzuki, A., Washida, H. and Takaiwa, F. 1998. The GCN4 motif in a rice glutelin gene is essential for endosperm-specific expression and is activated by Opaque 2 in transgenic rice plants. Plant J. 14:673-683.

Wu, J., Yu, L., Li, L., Hu, J., Zhou, J. and Zhou, X. 2007. Oral immunization with transgenic rice seeds expressing VP2 protein of infections bursal disease virus induces protective immune responses in chickens. Plant Biotechnol J. 5:570-578.

Wu, L.S., Wang, L.D., Chen, P.W., Chen, L.J. and Tzen, J.T. 1998. Genomic cloning of 18 kDa oleosin and detection of triacylglycerols and oleosin isoforms in maturing rice and postgerminative seedlings. J Biochem. 123:386-391.

Wu, M., Maier, E., Benz, R. and Hancock, R.E. 1999. Mechanism of interaction of different classes of cationic antimicrobial peptides with planar bilayers and with the cytoplasmic membrane of *Escherichia coli*. Biochemistry. 38:7235-7242.

Wulff, E.G., Sørensen, J.L., Lübeck, M., Nielsen, K.F., Thrane, U. and Torp, J. 2010. *Fusarium* spp. associated with rice Bakanae: ecology, genetic diversity, pathogenicity and toxigenicity. Environ Microbiol. 12:649-657.

X

Xing, H., Lawrence, C.B., Chambers, O., Davies, H.M., Everett, N.P. and Li, Q.Q. 2006. Increased pathogen resistance and yield in transgenic plants expressing combinations of the modified antimicrobial peptides based on indolicidin and magainin. Planta. 223:1024-1032.

Xu, J.H. and Messing J. 2009. Amplification of prolamin storage protein genes in different subfamilies of the Poaceae. Theor Appl Genet. 119:1397-1412.

Y

Yamagata H, Sugimoto T, Tanaka K, Kasai Z. 1982. Biosynthesis of storage proteins in developing rice seeds. Plant Physiol. 70:1094-1100.

Yamagata, H. and Tanaka, K. 1986. The site of synthesis and accumulation of rice storage proteins. Plant Cell Physiol. 27:135-145.

Yang, D., Guo, F., Liu, B., Huang, N. and Watkins, S.C. 2003. Expression and localization of human lysozyme in the endosperm of transgenic rice. Planta. 216:597:603.

Yang, J., Barr, L.A., Fahnestock, S.R. and Liu, Z.B. 2005. High yield recombinant silk-like protein production in transgenic plants through protein targeting. Transgenic Res. 14:313-324.

Yang, L., Weiss, T.M., Lehrer, R.I. and Huang, H.W. 2000. Crystallization of antimicrobial pores in membranes: magainin and protegrin. Biophys J. 79:2002-2009.

Yang, L., Ding, J., Zhang, C., Jia, J., Weng, H., Liu, W. and Zhang, D. 2005. Estimating the copy number of transgenes in transformed rice by real-time quantitative PCR. Plant Cell Rep. 23:759-763.

Yang, L., Suzuki, K., Hirose, S., Wakasa, Y. and Takaiwa, F. 2007. Development of transgenic rice seed accumulating a major Japanese cedar pollen allergen (Cry j 1) structurally disrupted for oral immunotherapy. Plant Biotechnol J. 5:815-826.

Yang, L., Kajiura, H., Suzuki, K., Hirose, S., Fujiyama, K. and Takaiwa, F. 2008. Generation of a transgenic rice seed-based edible vaccine against house dust mite allergy. Biochem Biophys Res Commun. 11:334-339.

Yang, L., Hirose, S., Suzuki, K., Hiroi, T. and Takaiwa, F. 2012. Expression of hypoallergenic Der f 2 derivatives with altered intramolecular disulphide bonds induces the formation of novel ER-derived protein bodies in transgenic rice seeds. J Exp Bot. 63:2947-2959.

Yasuda, H., Tada, Y., Hayashi, Y., Jomori, T. and Takaiwa F. Expression of the small peptide GLP-1 in transgenic plants. 2005. Transgenic Res. 14:677-684.

Yasuda, H., Hayashi, Y., Jomori, T. and Takaiwa, F. 2006. The correlation between expression and localization of a foreign gene product in rice endosperm. Plant Cell Physiol. 47:756-763.

Yeaman, M.R. and Yount, N.Y. 2003. Mechanisms of antimicrobial peptide action and resistance. Pharmacol Rev. 55:27-55.

Yevtushenko, D.P., Romero, R., Forward, B.S., Hancock, R.E., Kay, W.W. and Misra, S. 2005. Pathogen-induced expression of a cecropin A-melittin antimicrobial peptide gene confers antifungal resistance in transgenic tobacco. J Exp Bot. 56:1685-1695.

Yevtushenko, D.P. and Misra, S. 2007. Comparison of pathogen-induced expression and efficacy of two amphibian antimicrobial peptides, MsrA2 and temporin A, for engineering wide-spectrum disease resistance in tobacco. Plant Biotechnol J. 5:720-734.

Yevtushenko, D.P and Misra, S. 2012. Transgenic expression of antimicrobial peptides in plants: strategies for enhanced disease resistance, improved productivity, and production of therapeutics, pp. 445–458. In: Rajasekaran K, Cary J, Jaynes J, Montesinos E. (eds.). Small Wonders: Peptides for Disease Control. Volume 1095. American Chemical Society/ Oxford University Press Inc. Washington, USA.

Yi, C.X., Zhang, J., Chan, K.M., Liu, X.K. and Hong, Y. 2008. Quantitative real-time PCR assay to detect transgene copy number in cotton (*Gossypium hirsutum*). Anal Biochem. 375:150-152.

Yoshida, T., Kimura, E., Koike, S., Nojima, J., Futai, E., Sasagawa, N., Watanabe, Y. and Ishiura S. 2011.Transgenic rice expressing amyloid β-peptide for oral immunization.Int J Biol Sci. 2011 Mar 25:7(3):301-7.

Yount, N. Y. and Yeaman, M.R. 2005. Immunocontinuum: perspectives in antimicrobial peptide mechanisms of action and resistance. Protein Pept. Lett. 12:49-67.

Yu, J., Hu, S., Wang, J., Wong, G.K., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X., Cao, M., Liu, J., Sun, J., Tang, J., Chen, Y., Huang, X., Lin, W., Ye, C., Tong, W., Cong, L., Geng, J., Han, Y., Li, L., Li, W., Hu, G., Huang, X., Li, W., Li, J., Liu, Z., Li, L., Liu, J., Qi, Q., Liu, J., Li, L., Li, T., Wang, X., Lu, H., Wu, T., Zhu, M., Ni, P., Han, H., Dong, W., Ren, X., Feng, X., Cui, P., Li, X., Wang, H., Xu, X., Zhai, W., Xu, Z., Zhang, J., He, S., Zhang, J., Xu, J., Zhang, K., Zheng, X., Dong, J., Zeng, W., Tao, L., Ye, J., Tan, J., Ren, X., Chen, X., He, J., Liu, D., Tian, W., Tian, C.,Xia, H., Bao, Q., Li, G., Gao, H., Cao, T., Wang, J., Zhao, W., Li, P., Chen, W., Wang, X., Zhang, Y., Hu, J., Wang, J., Liu, S., Yang, J., Zhang, G., Xiong, Y., Li, Z., Mao, L., Zhou, C., Zhu, Z.,Chen, R., Hao, B., Zheng, W., Chen, S., Guo, W., Li, G., Liu, S., Tao, M., Wang, J., Zhu, L., Yuan, L. and Yang, H. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science. 296:79-92.

Z

Zasloff, M. 1987. Magainins, a class of antimicrobial peptides from *Xenopus* skin: isolation, characterization of two active forms, and partial cDNA sequence of a precursor. Proc Natl Acad Sci U S A. 84: 5449-5453.

Zasloff, M. 2002. Antimicrobial peptides of multicellular organisms. Nature. 415:389-395.

Zhang, L., Rozek, A., Hancock, R.E. 2001. Interaction of cationic antimicrobial peptides with model membranes. J Biol Chem 276:35714-35722.

Zhang, Y., Yang, B.Y. and Chen, S.Y. 2006. Inheritance analysis of herbicide-resistant transgenic soybean lines. Yi Chuan Xue Bao. 33:1105-1111.

Zheng, Z., Sumi, K., Tanaka, K. and Murai, N. 1995. The bean seed Storage protein [beta]-phaseolin is synthesized, processed, and accumulated in the vacuolar type-II protein bodies of transgenic rice endosperm. Plant Physiol. 109:777-786.

Zhou, M., Hu, Q., Li, Z., Li, D., Chen, C.F. and Luo, H. 2011. Expression of a novel antimicrobial peptide Penaeidin4–1 in creeping bentgrass (*Agrostis stolonifera* L.) enhances plant fungal disease resistance. PLoS One. 6:e24677.

Zuo, J. and Chua, N-H. 2000. Chemical inducible systems for regulated expression of plant genes. Curr. Opin. Biotechnol. 11:146-151.

8. ANNEX

Annex 8.1: Nucleotide sequence alignment of the nucleotide sequences of the promoter and N-terminal signal peptide region of *GluB-1*, *GluB-4*, *Glb-1* and *Ole18* genes isolated in this work against their respective nucleotide sequences deposited in EMBL/GenBank/DDBJ database

```
GluB-1          GGTACCTCTAGACAGATTCTTGCTACCAACAACTTCACAAAGTAGTAGTCAACCAAAACT 60
AY427569              TCTAGACAGATTCTTGCTACCAACAACTTCACAAAGTAGTAGTCAACCAAAACT 60
                      ********************************************************

GluB-1          ATGCTAAGGAATCACCTCACTTCCGCCCATGACCGTGAGCACGACTGTTCAAACAGTTTG 120
AY427569        ATGCTAAGGAATCACCTCACTTCCGCCCATGACCGTGAGCACGACTGTTCAAACAGTTTG 120
                ************************************************************

GluB-1          TTAATCTCTACAAAGAAGGTACACTTTACCTACACAACGCCACTAACCTGAGTTACCCAG 180
AY427569        TTAATCTCTACAAAGAAGGTACACTTTACCTACACAACGCCACTAACCTGAGTTACCCAG 180
                ************************************************************

GluB-1          CCCATGCAAAATAGCCACGTCTTGTGACTTAAGGGATTTCGCGACAAGGCATTTCGAAAG 240
AY427569        CCCATGCAAAATAGCCACGTCTTGTGACTTAAGGGATTTCGCGACAAGGCATTTCGAAAG 240
                ************************************************************

GluB-1          CCCACACAAGGACACCTTATGAAAACTGGAGGGGTCCCACAGACCAACAACAAGTTAGGT 300
AY427569        CCCACACAAGGACACCTTATGAAAACTGGAGGGGTCCCACAGACCAACAACAAGTTAGGT 300
                ************************************************************

GluB-1          CCCAAACCATGTTGTGCCAGGAAAAATCCAAGGGGTCCTCCCCAACACCACCCCGACAAA 360
AY427569        CCCAAACCATGTTGTGCCAGGAAAAATCCAAGGGGTCCTCCCCAACACCACCCCGACAAA 360
                ************************************************************

GluB-1          TCCACTTGTCCATTGGCATCAAGATTTGCCTGACCTAGCTAATTACTCAGCCAGGCATGT 420
AY427569        TCCACTTGTCCATTGGCATCAAGATTTGCCTGACCTAGCTAATTACTCAGCCAGGCATGT 420
                ************************************************************

GluB-1          CACAATTCACCCATGTGGTCACACATGTTATGGTTGGATGAAATTCTAAAGGAA-----T 475
AY427569        CACAATTCACCCATGTGGTCACACATGTTA-GGTTGGA-GAAATTCTAAAGGAAAGGAAT 478
                ****************************** ******* ***************     *

GluB-1          CGGTCCATATGAGCAAGAACGAGAAAACCATACCACCAGTACTTCTACCGAAATACGAGT 535
AY427569        CGGTCCATATGAGCAAGACCGAGAAA-CCATACCACCAGTACTTCTACCGAAATACGAGT 537
                ****************** ******* **********************************

GluB-1          TTAGTAAACTCATTTGTTTTCAAGGCACC-GACCCAGGTGTGTCGGGTTTTCCAGGGATT 594
AY427569        TTAGTAAACTCATTTGTTTTCAAGGCACCCGACCCAGGTGTGTCGGGTTTTCCAGGGATT 597
                ****************************** ******************************

GluB-1          TTGTAAACCCAAGTTTTACCCATAGTTGATCATTCAAATTTTGAGGAGGGTCATTGGTAT 654
AY427569        TTGTAAACCCAAGTTTTACCCATAGTTGATCATTCAAATTTTGAGGAGGGTCATTGGTAT 657
                ************************************************************

GluB-1          CCGTACCTGAGGGCACGAATACTGAGACCTAGCATTGTAGTCGACCAAGGAGGTTAATGC 714
AY427569        CCGTACCTGAGGGCACGAATACTGAGACCTAGCATTGTAGTCGACCAAGGAGGTTAATGC 717
                ************************************************************

GluB-1          AGCAATTGTAGGTGGGGCCTGTTGGTTATATTGCAAACTGCGGCCAACATTTCATGTGTA 774
AY427569        AGCAATTGTAGGTGGGGCCTGTTGGTTATATTGCAAACTGCGGCCAACATTTCATGTGTA 777
                ************************************************************

GluB-1          ATTTAGAGATGTGCATTTTGAGAAATGAAATACTTAGTTTCAAATTATGGGCTCAAAATA 834
AY427569        ATTTAGAGATGTGCATTTTGAGAAATGAAATACTTAGTTTCAAATTATGGGCTCAAA-TA 836
                ************************************************************ **

GluB-1          ATCAAAGGTGACCTACCTTGCTTGATATCTTGAGCTTCTTCCTCGTATTCCGCGCACTAG 894
AY427569        ATGAAAGGTGACCTACCTTGCTTGATATCTTGAGCTTCTTCCTCGTATTCCGCGCACTAG 896
                ** ********************************************************

GluB-1          GAC-TCTTCTGGCTCCGAAGCTACACGTGGAACGAGATAACTCAACAAAACGACCAAGGA 953
AY427569        GAGATCTTCTGGCTCCGAAGCTACACGTGGAACGAGATAACTCAACAAAACGACCAAGGA 956
                **  ********************************************************
```

```
GluB-1          AAAGCTCGTATTAGTGAGTACTAAGTGTGCCACTGAATAGATCTCGATTTTTGAGGAATT 1013
AY427569        AAAGCTCGTATTAGTGAGTACTAAGTGTGCCACTGAATAGATCTCGATTTTTGAGGAATT 1016
                ************************************************************

GluB-1          TTAGAAGTTGAACAGAGTCAATCGAACAGACAGTTGAAGAGATATGGATTTTCTAAGATT 1073
AY427569        TTAGAAGTTGAACAGAGTCAATCGAACAGACAGTTGAAGAGATATGGATTTTCTAAGATT 1076
                ************************************************************

GluB-1          AATTGATTCTCTGTCTAAAGAAAAAAAGTATTATTGAATTAAATGGAAAAAGAAAAAGGA 1133
AY427569        AATTGATTCTCTGTATAAAGAAAAAAAGTATTATTGAATTAAATGGAAAAAGAAAAAGGA 1136
                ************** ********************************************

GluB-1          AAAAGGGGATGGCTTCTGCCTTTTTGGGCTGAAGGCGGCGTGTGGCCCAGCGTGCTGCGT 1193
AY427569        AAAAGGGGATGGCTTCTGC-TTTTTGGGCTGAAGGCGGCGTGTGGCC-AGCGTGCTGCGT 1194
                ****************** ************************** ************

GluB-1          GGGCACAGCCGAGCGAACACACGACGGAGCAGCTACGACGAACGGGGGACCGAGTGGACC 1253
AY427569        GCGGACAGC-GAGCGAACACACGACGGAGCAGCTACGACGAACGGGGGACCGAGTGGACC 1253
                * * ***** ************************************************

GluB-1          GGACGAGGATGTGGCCTAGGACGAGTGCACAAGGCTAGTGGACTCGGTCCCCCGCGCGGT 1313
AY427569        GGACGAGGATGTGGCCTAGGACGAGTGCACAAGGCTAGTGGACTCGGTCCCC-GCGCGGT 1312
                ************************************************** *******

GluB-1          ATCCCGAGTGGGTCCACTCGTCTGCAAACACGATTCACATAGAGCGGGAGCACGCGGGGA 1373
AY427569        ATCCCGAGTGG-TCCACT-GTCTGCAAACACGATTCACATAGAGCGGGCAGACGCGGG-A 1369
                *********** ****** ****************************   ******* *

GluB-1          GCCGTCCTAGGTGCACGGGAAGCAAATCCGTGCGCCTGGGTGGATTTGAGTGACACGGGC 1433
AY427569        GCCGTCCTAGGTGCACCGGAAGCAAATCCGT-CGCCTGGGTGGATTTGAGTGACACGG-C 1427
                *************** ************** ************************** *

GluB-1          CCACGTGTAGCCTCACAGCTCTCCGTGGTCAGATGTGTAAAATTATCATAATATGTGTTT 1493
AY427569        CCACGTGTAGCCTCACAGCTCTCCGTGGTCAGATGTGTAAAATTATCATAATATGTGTTT 1487
                ************************************************************

GluB-1          TTCAAATAGTTAAATAATATATATAGGCAAGTTATATGGGTCAATAAGCAGTAAAAAGGC 1553
AY427569        TTCAAATAGTTAAATAATATATATAGGCAAGTTATATGGGTCAATAAGCAGTAAAAAGGC 1547
                ************************************************************

GluB-1          TTATGACATGGTAAAATTACTTACACCAATATGCCTTACTGTCTGATATATTTTACATGA 1613
AY427569        TTATGACATGGTAAAATTACTTACACCAATATGCCTTACTGTCTGATATATTTTACATGA 1607
                ************************************************************

GluB-1          CAACAAAGTTACAAGTACGTCATTTAAAAATACAAGTTACTTATCAATTGTAGTGTATCA 1673
AY427569        CAACAAAGTTACAAGTACGTCATTTAAAAATACAAGTTACTTATCAATTGTAGTGTATCA 1667
                ************************************************************

GluB-1          AGTAAATGACAACAAACCTACAAATTTGCTATTTTGAAGGAACACTTAAAAAAATCAATA 1733
AY427569        AGTAAATGACAACAAACCTACAAATTTGCTATTTTGAAGGAACACTTAAAAAAATCAATA 1727
                ************************************************************

GluB-1          GGCAAGTTATATAGTCAATAAACTGCAAGAAGGCTTATGACATGGAAAAATTACATACAC 1793
AY427569        GGCAAGTTATATAGTCAATAAACTGCAAGAAGGCTTATGACATGGAAAAATTACATACAC 1787
                ************************************************************

GluB-1          CAATATGCTTTATTGTCCGGTATATTTTACAAGACAACAAAGTTATAAGTATGTCATTTA 1853
AY427569        CAATATGCTTTATTGTCCGGTATATTTTACAAGACAACAAAGTTATAAGTATGTCATTTA 1847
                ************************************************************

GluB-1          AAAATACAAGTTACTTATCAATTGTCAAGTAAATGAAAACAAACCTACAAATTTGTTATT 1913
AY427569        AAAATACAAGTTACTTATCAATTGTCAAGTAAATGAAAACAAACCTACAAATTTGTTATT 1907
                ************************************************************

GluB-1          TTGAAGGAACACCTAAATTATCAAATATAGCTTGCTACGCAAAAATGACAACATGCTTAC 1973
AY427569        TTGAAGGAACACCTAAATTATCAAATATAGCTTGCTACGCAAAA-TGACAACATGCTTAC 1966
                ******************************************* ****************

GluB-1          AAGTTATTATCATCTTAAAGTTAGACTCATCTTCTCAAGCATAAGAGCTTTATGGTGCAA 2033
AY427569        AAGTTATTATCATCTTAAAGTTAGACTCATCTTCTCAAGCATAAGAGCTTTATGGTGCAA 2026
                ************************************************************
```

```
GluB-1          AAACAAATATAATGACAAGGCAAAGATACATA----TTAAGAGTATGGATAGACATTTCT 2089
AY427569        AAACAAATATAATGACAAGGCAAAGATACATACATTAAGAGTATGGACAGACATTTCT 2086
                ********************************    ************ **********

GluB-1          TTAACAAACTCCATTTGTATTACTCCAAAAGCACCAGAAGTTTGTCATGGCTGAGTCATG 2149
AY427569        TTAACAAACTCCATTTGTATTACTCCAAAAGCACCAGAAGTTTGTCATGGCTGAGTCATG 2146
                ************************************************************

GluB-1          AAATGTATAGTTCAATCTTGCAAAGTTGCCTTTCCTTTTGTACTGT--TTTAACACTACA 2207
AY427569        AAATGTATAGTTCAATCTTGCAAAGTTGCCTTTCCTTTTGTACTGTGTTTTAACACTACA 2206
                ********************************************  ************

GluB-1          AGCCATATATTGTCTGTACGTGCAACAAACTATATCACCATGTATCCCAAGATGCTTTTT 2267
AY427569        AGCCATATATTGTCTGTACGTGCAACAAACTATATCACCATGTATCCCAAGATGCTTTTT 2266
                ************************************************************


GluB-1          TATTGCTATATAAACTAGCTTGGTCTGTCTTTGAACTCACATCAATTAGCTTAAGTTTCC 2327
AY427569        TATTGCTATATAAACTAGCTTGGTCTGTCTTTGAACTCACATCAATTAGCTTAAGTTTCC 2326
                ************************************************************

GluB-1          ATAAGCAAGTACAAATAGCTATGGCGAGTTCCGTTTTCTCTCGGTTTTCTATATACTTTT 2387
AY427569        ATAAGCAAGTACAAATAGCTATGGCGAGTTCCGTTTTCTCTCGGTTTTCTATATACTTTT 2386
                ************************************************************

GluB-1          GTGTTCTTCTATTATGCCATGGTTCTATGGCCAAGAAGCTTGAGCTCGGATCC 2428
AY427569        GTGTTCTTCTATTATGCCATGGTTCTATGGCCAAG                   2427
                ***********************************
```

```
Glb-1           GGTACC------TGGAGGGAGGAGAGGGGAGA--TGGTGAGAGAGGAGGAAGAAGAGGAG 58
AY427575              GGCGCCTGGAGGGAGGAGAGGGGAGAGATGGTGAGAGAGGAGGAAGAAGAGGAG 54
                            ********************  **************************

Glb-1           GGGTGACAATGATATGTGGGGCCATGTGGGCCCCACCATTTTTTAATTCATTCTTTTGTT 118
AY427575        GGGTGACAATGATATGTGGG-CCATGTGGCCCCCACCATTTTTTAATTCATTCTTTTGTT 113
                ******************** ******** ******************************

Glb-1           GAAACTGACATGTGGGTCCCATGAGATTTATTATTTTTCGGATCGAATTGCCACGTAAGC 178
AY427575        GAAACTGACATGTGGGTCCCATGAGAATTATTATTTTTCGGATCGAATTGCCACGTAAGC 173
                ************************** ********************************

Glb-1           ACTACGTCAATGCTACGTCAGATGAAGACCGAGTCAAATTAGCCACGTAAGCGCCACGTC 238
AY427575        GCTACGTCAATGCTACGTCAGATGAAGACCGAGTCAAATTAGCCACGTAAGCGCCACGTC 233
                 ************************************************************

Glb-1           AGCCAAAACCACCATCCAAACCGCCGAGGGACCTCATCTGCACTGGTTTTGATAGTTGAG 298
AY427575        AGCCAAAACCACCATCCAAACCGCCGAGGGACCTCATCTGCACTGGTTTTGATAGTTGAG 293
                ************************************************************

Glb-1           GGACCCGTTGTATCTGGTTTTTCGATTGAAGGACGAAAATCAAATTTGTTGACAAGTTAA 358
AY427575        GGACCCGTTGTATCTGGTTTTTCGATTGAAGGACGAAAATCAAATTTGTTGACAAGTTAA 353
                ************************************************************

Glb-1           GGGACCTTAAATGAACTTATTCCATTTCAAAATATTCTGTGAGCCATATATACCGTGGGC 418
AY427575        GGGACCTTAAATGAACTTATTCCATTTCAAAATATTCTGTGAGCCATATAT-CCGTGGGC 412
                ************************************************** ********

Glb-1           TTCCAATCCTCCTCAAATTAAAGGGCCTTTTTAAAATAGATAATTGCCTTCTTTCAGTCA 478
AY427575        TTCCAATCCTCCTCAAATTAAAGGGCCTTTTTAAAATAGATAATTGCCTTCTTTCAGTCA 472
                ************************************************************

Glb-1           CCCATAAAAGTACAAAACTACTACCAACAAGCAACATGCGCAATTACACACATTTTCTGC 538
AY427575        CCCATAAAAGTACAAAACTACTACCAACAAGCAACATGCGCAGTTACACACATTTTCTGC 532
                ****************************************** ******************

Glb-1           ACATTTCCACCACGTCACAAAGAGCTAAGAGTTATCCCTAGGACAATCTCATTAGTGTAG 598
AY427575        ACATTTCCACCACGTCACAAAGAGCTAAGAGTTATCCCTAGGACAATCTCATTAGTGTAG 592
                ************************************************************
```

```
Glb-1           ATACATCCATTAATCTTTTATCAGAGGCAAACGTAAAGCCGCTCTTTATGACAAAAATAG 658
AY427575        ATACATCCATTAATCTTTTATCAGAGGCAAACGTAAAGCCGCTCTTTATGACAAAAATAG 652
                ************************************************************

Glb-1           GTGACACAAAAGTGTTATCTGCCACATACATAACTTCAGAAATTACCCAACACCAAGAGA 718
AY427575        GTGACACAAAAGTGTTATCTGCCACATACATAACTTCAGAAATTACCCAACACCAAGAGA 712
                ************************************************************

Glb-1           AAAATAAAAAAAAAATCTTTTTGCAAGCTCCAAATCTTGGAAACCTTTTTCACTCTTTGCA 778
AY427575        AAAATAAAAAAAAATCTTTTTGCAAGCTCCAAATCTTGGAAACCTTTTTCACTCTTTGCA 772
                ************************************************************

Glb-1           GCATTGTACTCTTGCTCTTTTTCCAACCGATCCATGTCACCCTCAAGCTTCTACTTGATC 838
AY427575        GCATTGTACTCTTGCTCTTTTTCCAACCGATCCATGTCACCCTCAAGCTTCTACTTGATC 832
                ************************************************************

Glb-1           TACACGAAGCTCACCGTGCACACAACCATGGCCACAAAAACCCTATAAAACCCCATCCGA 898
AY427575        TACACGAAGCTCACCGTGCACACAACCATGGCCACAAAAACCCTATAAAACCCCATCCGA 892
                ************************************************************

Glb-1           TCGCCATCATCTCATCATCAGTTCATCACCAACAAACAAAAGAGGAAAAAAAACATATAC 958
AY427575        TCGCCATCATCTCATCATCAGTTCATCACCAACAAACAAAAGAGGAAAAAAAACATATAC 952
                ************************************************************

Glb-1           ACTTCTAGTGATTGTCTGATTGATCATCAATGGCTAGCAAGGTCGTCTTCTTCGCGGCGG 1018
AY427575        ACTTCTAGTGATTGTCTGATTGATCATCAATGGCTAGCAAGGTCGTCTTCTTCGCGGCGG 1012
                ************************************************************

Glb-1           CGCTCATGGCGGCCATGGTGGCCATCTCCGGCGCC 1053
AY427575        CGCTCATGGCGGCCATGGTGGCCATCTCCGGCGCG 1047
                ********************************
```

```
GluB-4          GGTACCTACAGGGTTCCTTGCGTGAAGAAGGGTGGCCTGCGGTTCACCATTAACGGTCAC 60
AY427571              TACAGGGTTCCTTGCGTGAAGAAGGGTGGCCTGCGGTTCACCATTAACGGTCAC 54
                      ********************************************************

GluB-4          GACTACTTCCAGCTAGTACTGGTGACCAACGTCGCGGCGGCAGGGTCAATCAAGTCCATG 120
AY427571        GACTACTTCCAGCTAGTACTGGTGACCAACGTCGCGGCGGCAGGGTCAATCAAGTCCATG 114
                ************************************************************

GluB-4          GAGGTTATGGGTTCCAACACAGCGGATTGGATGCCGATGGCACGTAACTGGGGCGCCCAA 180
AY427571        GAGGTTATGGGTTCCAACACAGCGGATTGGATGCCGATGGCACGTAACTGGGGCGCCCAA 174
                ************************************************************

GluB-4          TGGCACTCACTGGCCTACCTCACCGGTCAAGGTCTATCCTTTAGGGTCACCAACACAGAT 240
AY427571        TGGCACTCACTGGCCTACCTCACCGGTCAAGGTCTATCCTTTAGGGTCACCAACACAGAT 234
                ************************************************************

GluB-4          GACCAAACGCTCGTCTTCACCAACGTCGTGCCACCAGGATGGAAGTTTGGCCAGACATTT 300
AY427571        GACCAAACGCTCGTCTTCACCAACGTCGTGCCACCAGGATGGAAGTTTGGCCAGACATTT 294
                ************************************************************

GluB-4          GCAAGCAAGCTGCAGTTCAAGTGAGAGGAGAAGCCTGAATTGATACCGGAGCGTTTCTTT 360
AY427571        GCAAGCAAGCTGCAGTTCAAGTGAGAGGAGAAGCCTGAATTGATACCGGAGCGTTTCTTT 354
                ************************************************************

GluB-4          TGGGAGTAACATCTCTGGTTGCCTAGCAAACATATGATTGTATATAAGTTTCGTTGTGCG 420
AY427571        TGGGAGTAACATCTCTGGTTGCCTAGCAAACATATGATTGTATATAAGTTTCGTTGTGCG 414
                ************************************************************

GluB-4          TTTATTCTTTCGGTGTGTAAAATAACATACATGCTTTCCTGATATTTTCTTGTATATATG 480
AY427571        TTTATTCTTTCGGTGTGTAAAATAACATACATGCTTTCCTGATATTTTCTTGTATATATG 474
                ************************************************************

GluB-4          TACACACACACGACAAATCCTTCCATTTCTATTATTATTGAACAATTTAATTGCGAGGGC 540
AY427571        TACACACACACGACAAATCCTTCCATTTCTATTATTATTGAACAATTTAATTGCGAGGGC 534
                ************************************************************
```

```
GluB-4          GAGTACTTGTCTGTTTACCTTTTTTTTTTCAGATGGCATTTTATAGTTTAACCTTTCATG 600
AY427571        GAGTACTTGTCTGTTTACCTTTTTTTTTTCAGATGGCATTTTATAGTTTAACCTTTCATG 594
                ************************************************************

GluB-4          GACCGGCAGTAGTTCTAACCATGAATGAAAAGAAATCATAGTCCACACCACGCAGGGACA 660
AY427571        GACCGGCAGTAGTTCTAACCATGAATGAAAAGAAATCATAGTCCACACCACGCAGGGACA 654
                ************************************************************

GluB-4          TTGTGGTCATTTTAGACAAGACGATTTGATTAATGTCTTGTATGATATGGTCGACAGTGA 720
AY427571        TTGTGGTCATTTTAGACAAGACGATTTGATTAATGTCTTGTATGATATGGTCGACAGTGA 714
                ************************************************************

GluB-4          GGACTAACAAACATATGGCATATTTTATTACCGGCGAGTTAAATAAATTTATGTCACAGT 780
AY427571        GGACTAACAAACATATGGCATATTTTATTACCGGCGAGTTAAATAAATTTATGTCACAGT 774
                ************************************************************

GluB-4          AATAAACTGCCTAATAAATGCACGCCAGAAAATATAATGATAAAAAAAAGAAAAGATACA 840
AY427571        AATAAACTGCCTAATAAATGCACGCCAGAAAATATAATGATAAAAAAAAGAAAAGATACA 834
                ************************************************************

GluB-4          TAAGTCCATTGCTTCTACTTTTTTAAAAATTAAATCCAACATTTTCTATTTTTTGGTATA 900
AY427571        TAAGTCCATTGCTTCTACTTTTTTAAAAATTAAATCCAACATTTTCTATTTTTTGGTATA 894
                ************************************************************

GluB-4          AACTTGGAAGTACTAGTTGGATATGCAAAATCATCTAACCTCCATATATTTCATCAATTT 960
AY427571        AACTTGGAAGTACTAGTTGGATATGCAAAATCATCTAACCTCCATATATTTCATCAATTT 954
                ************************************************************

GluB-4          GTTTACTTTACATATGGGAGAGGATAGTATGTCAAAGAAAATGACAACAAGCTTACAAGT 1020
AY427571        GTTTACTTTACATATGGGAGAGGATAGTATGTCAAAGAAAATGACAACAAGCTTACAAGT 1014
                ************************************************************

GluB-4          TTCTTATTTTAAAAGTTCCGCTAACTTATCAAGCATAGTGTGCCACGCAAAACTGACAAC 1080
AY427571        TTCTTATTTTAAAAGTTCCGCTAACTTATCAAGCATAGTGTGCCACGCAAAACTGACAAC 1074
                ************************************************************

GluB-4          AAACCAACAAATTTAAGGAGCGCCTAACTTATCATCTATGACATACCGCACAAAATGATA 1140
AY427571        AAACCAACAAATTTAAGGAGCGCCTAACTTATCATCTATGACATACCGCACAAAATGATA 1134
                ************************************************************

GluB-4          ACATACTAGAGAAACTTTATTGCACAAAAGGAAATTTATCCATAAGGCAAAGGAACATCT 1200
AY427571        ACATACTAGAGAAACTTTATTGCACAAAAGGAAATTTATCCATAAGGCAAAGGAACATCT 1194
                ************************************************************

GluB-4          TAAGGCTTTGGATATACATTTACCAACAAGCATTGTTTGTATTACCCCTAAAGCGCAAGA 1260
AY427571        TAAGGCTTTGGATATACATTTACCAACAAGCATTGTTTGTATTACCCCTAAAGCGCAAGA 1254
                ************************************************************

GluB-4          CATGTCATCCATGAGTCATAGTGTGTATATCTCAACATTGCAAAGCTACCTTTTTTCTAT 1320
AY427571        CATGTCATCCATGAGTCATAGTGTGTATATCTCAACATTGCAAAGCTACCTTTTTTCTAT 1314
                ************************************************************

GluB-4          TATACTTTTCGCATTATAGGCTAGATATTATCTATACATGTCAACAAACTCTATCCCTAC 1380
AY427571        TATACTTTTCGCATTATAGGCTAGATATTATCTATACATGTCAACAAACTCTATCCCTAC 1374
                ************************************************************

GluB-4          GTCATATCTGAAGATTCTTTTCTTCACTATATAAGTTGGCTTCCCTGTCATTGAACTCAC 1440
AY427571        GTCATATCTGAAGATTCTTTTCTTCACTATATAAGTTGGCTTCCCTGTCATTGAACTCAC 1434
                ************************************************************

GluB-4          ATCAACCAGCCCAAGTTTCCAATAACATCCTCAAATAGCTATGGCGACCATAGCTTTCTC 1500
AY427571        ATCAACCAGCCCAAGTTTCCAATAACATCCTCAAATAGCTATGGCGACCATAGCTTTCTC 1494
                ************************************************************

GluB-4          TCGGTTATCTATCTACTTTTGTGTTCTTCTCCTATGCCATCCACTCATGGCC 1552
AY427571        TCGGTTATCTATCTACTTTTGTGTTCTTCTCCTATGCCATCCCTCTATGGCC 1546
                ****************************************** ** ******
```

```
Ole18           GAATTCGATGGTCAGCCAATACATTGATCCGTTGCCAATCATGCAAAGTATTTTGGCTGT 60
AY427563        ------GATGGTCAGCCAATACATTGATCCGTTGCCAATCATGCAAAGTATTTTGGCTGT 54
                      ************************************************** 

Ole18           GGCCGAGTGCCGGAATTGATAATTGTGTTCTGACTAAATTAAATGACCAGAAGTCGCTAT 120
AY427563        GGCCGAGTGCCGGAATTGATAATTGTGTTCTGACTAAATTAAATGACCAGAAGTCGCTAT 114
                ************************************************************ 

Ole18           CTTCCAATGTATCCGAAACCTGGATTAAGCAATCCTGTTCTGTTCTCTAGCCCCTCCTGC 180
AY427563        CTTCCAATGTATCCGAAACCTGGATTAAACAATCCTGTTCTGTTCTCTAGCCCCTCCTGC 174
                ***************************** ****************************** 

Ole18           ATGGCCGGATTGTTTTTTTGACATGTTTTCTTGACTGAGGCCTGTTTGTTCTAAACTTTT 240
AY427563        ATGGCCGGATTGTTTTTTTGACATGTTTTCTTGACTGAGGCCTGTTTGTTCTAAACTTTT 234
                ************************************************************ 

Ole18           TCTTCAAACTTTTAACTTTTTCATCACATCAGAACTTTTCTACACACATAAACTTTTAAC 300
AY427563        TCTTCAAACTTTTAACTTTTTCATCACATCAGAACTTTTCTACACATATAAACTTTTAAC 294
                ********************************************* ************** 

Ole18           TTTTCTGTCACATCGTTCCAATTTCAATCAAACTTTTAATTTTGGCTTGAACTAAACACA 360
AY427563        TTTTCCGTCACATCGTTCCAATTTCAATCAAACTTTCAATTTTGGCGTGAACTAAACACA 354
                ***** ****************************** ********* ************** 

Ole18           CCCTGAGTCTTTTATTGCTCCTCCATACGGGTTGGCTGGTTGAGAATAGGTATTTTCAGA 420
AY427563        CCCTGAGTCTTTTATTGCTCCTCCGTACGGGTTGGCTGGTTGAGAATAGGTATTTTCAGA 414
                ************************ ********************************** 

Ole18           GAGAAAATCTGGATATTGGGAGGAGGAACTTGGCATGAATGGCCACTATATTTAGAGCAA 480
AY427563        GAGAAAATCTAGATATTGGGAGGA---ACTTGGCATGAATGGCCACTATATTTAGAGCAA 471
                ********** ************   ********************************** 

Ole18           TTCTACGGTCTTTGAGGAGGTACCATGAGGTACCAAAATTTTAGTGTAAATTTTAGTATC 540
AY427563        TTCTACGGTCCTTGAGGAGGTACCATGAGGTACCAAAATTTTAGTGTAAATTTTAGTATC 531
                ********** ************************************************** 

Ole18           T------------------------------------------------------------ 541
AY427563        TCATTATAACTAGGTATTATGAGGTACCAAATTTACAATAGAAAAAATAGTACTTCATGG 591
                *

Ole18           -----TCTTAAGGACCGTAAAATTGCTCCTATATTTAAGGG-ATGTTTATATCTATCCAT 595
AY427563        TACTTTCTTAAGTACCGTAAAATTGCTCCTATATTTAAGGGGATGTTTATATCTATCCAT 651
                     ******* **************************** ****************** 

Ole18           ATCCATAATTTGATTTTGATAAGAAAAAATGTGAGCACACCAAGCATGTCCATGACCTTG 655
AY427563        ATCCATAATTTGATTTTGATAAGAAAAAATGTGAGCACACCAAGCATGTCCATGACCTTG 711
                ************************************************************ 

Ole18           CACTCTTGGCTCACTCGTCAACTGTGAAGAACCTAAAAAATGCTCAATATAGCTACAGGT 715
AY427563        CACTCTTGGCTCACTCGTCAACTGTGAAGAACCTCAAAAATGCTCAATATAGCTACAGGT 771
                ********************************** ************************ 

Ole18           GCCTGAAAAAATAACTTTAAAGTTTTGAACATCGATTTCACTAAACAACAATTATTATCT 775
AY427563        GCCTGAAAAAATAACTTTAAAGTTTTGAACATCGATTTCACTAAACAACAATTATTATCT 831
                ************************************************************ 

Ole18           CCCTCTGAAATGTTGCTACCTAAGATGATAGTTTAGAACTCTAGAATCATTGTCGGCGGA 835
AY427563        CCCTCTGAAA------------GATGATAGTTTAGAACTCTAGAATCATTGTCGGCGGA 878
                **********             ************************************ 

Ole18           GAAAGTAAATTATTTTCCCCAAATTTCCAGCTATGAAAAAACCCTCACCAAACACCATCA 895
AY427563        GAAAGTAAATTATTTTCCCCAAATTTCCAGCTATGAAAAAACCCTCACCAAACACCATCA 938
                ************************************************************ 

Ole18           AACAAGAGTTCACCAAACCGCCCATGCGGCCATGCTGTCACGCAACGCACCGCATTGCCT 955
AY427563        AACAAGAGTTCACCAAACCGCCCATGCGGCCATGCTGTCACGCAACGCACCGCATTGCCT 998
                ************************************************************ 
```

```
Ole18           GATGGCCGCTCGATGCATGCATGCTTCCCCGTGCACATATCCGACAGACGCGCCGTGTCA 1015
AY427563        GATGGCCGCTCGATGCATGCATGCTTCCCCGTGCACATATCCGACAGACGCGCCGTGTCA 1058
                ************************************************************

Ole18           GCGAGCTCCTCGACCGACCTGTGTAGCCCATGCAAGCATCCACCCCCGCCACGTACACCC 1075
AY427563        GCGAGCTCCTCGACCGACCTGTGTAGCCCATGCAAGCATCCACCCCCGCCACGTACACCC 1118
                ************************************************************

Ole18           CCTCCTCCTCCCTACGTGTCACCGCTCTCTCCACCTATATATGCCCACCTGGCCCCTCTC 1135
AY427563        CCTCCTCCTCCCTACGTGTCACCGCTCTCTCCACCTATATATGCCCACCTGGCCCCTCTC 1178
                ************************************************************

Ole18           CTCCCATCTCCACTTCACCCGATCGCTTCTTCTTCTTCTTCGTTCCATTCATCTTGCTAG 1195
AY427563        CTCCCATCTCCACTTCACCCGATCGCTTCTTCTTCTTCTTCGTTCCATTCATCTTGCTAG 1238
                ************************************************************

Ole18           CTCGCTTAGCA 1206
AY427563        CTAGCTTAGCA 1249
                ** ********
```

Restriction sites for cloning purpose are highlighted: in red, *Kpn* I; in yellow, *Hind* III; in dark grey, *Sac* I; in pale grey, *Nar* I; in green, *BamH* I; in blue, *EcorR* I and in fuchsia, *Bsm* I (3'-CTTACG -5')in the negative strand). Nucleotide sequence: in orange, modifications made in peptide coding sequence to introduce restriction sites; in red, nucleotide differences (nucleotide insertion, deletion or substitution) between the isolated and the theoretical promoter sequence; -, indicate nucleotide deletion; in green, signal peptide sequence; in pale blue, 5' UTR region. Nucleotides from *AMP derivatives* genes are indicate underlined.

Annex 8.2 Row data on *Dickeya*-inoculated T1 rice seeds expressing an *AMPder*

Construct	Code	Non inoculated	Inoculated
pCGluB-4::Bp188_KDEL(R8)	2	2.0	0.5±0.4
	2.1	2.0	0.8±0.5
	2.2	1.7±0.3	0.2±0.1
	2.3	2.0	0.1±0.1
	2.4	1.3±0.5	0.0
	2.5	1.7±0.3	0.0
	2.6	2.0	0.1±0.1
	2.7	1.8±0.2	0.5±0.3
	2.8	2.0	0.5±0.3
	2.9	2.0	0.4±0.3
	2.10	1.3±0.5	0.0
	2.11	1.6±0.4	0.0
	2.12	2.0	0.8±0.5
	2.13	1.5±0.5	1.0±0.4
	2.14	1.7±0.4	0.3±0.3
	2.15	2.0	1.3±0.5
	2.16	2.0	0.6±0.3
	2.17	2.0	0.1±0.1
	2.25	0.3±0.3	0.0
	2.26	1.5±0.6	0.0
	2.27	1.8±0.3	0.8±0.5
	2.28	1.9±0.2	0.5±0.2
	3.1	1.7±0.4	0.6±0.5
	3.2	2.0	0.0
	3.3	1.7±0.4	0.1±0.1
	3.4	1.3±0.6	0.3±0.2
	3.5	1.3±0.5	0.1±0.1
	4	1.8±0.4	0.2±0.2
	5	1.8±0.4	0.1±0.1
	5.3	1.9±0.2	0.2±0.1
pCOle18::Ole18_PRS_Cec A(R9)	3.1	0.6±0.4	0.0±0.1
	3.2	1.6±0.4	1.3±0.5
	7.1	1.7±0.4	0.4±0.2
	7.2	2.0	1.9±0.1
	8.1	2.0	1.1±0.6
	8.2	2.0	0.6±0.3
	8.3	1.9±0.2	0.1±0.1
	8.4	2.0	0.7±0.4
	8.5	1.9±0.1	0.9±0.1
	9.1	2.0	0.6±0.4
	10.1	2.0	0.8±0.3
	10.2	1.6±0.3	0.4±0.2

Construct	Code	Non inoculated	Inoculated
pCOle18::Ole18_PRS_Cec A(R9)	10.3	1.9±0.1	0.6±0.4
	10.4	1.9±0.2	0.1±0.1
	10.5	1.9±0.1	0.6±0.5
	11.1	2.0	0.0
	11.2	2.0	1.0±0.4
	11.3	2.0	0.8±0.3
	12.1	2.0	1.7±0.3
	12.2	1.9±0.1	0.7±0.5
	15.1	1.2±0.6	0.7±0.4
	15.2	2.0	0.0
	16.2	2.0	1.3±0.5
	17.1	0.8±0.5	0.3±0.2
pCGluB-4::Cec A(R10, A1)	14.1	1.5±0.5	0.4±0.2
	11.1	2.0	1.6±0.4
	1.2	2.0	0.1±0.1
	17.2	1.8±0.3	0.3±0.2
	2.1	1.8±0.3	0.2±0.1
	2.5	1.4±0.5	0.3±0.2
	17.3	1.8±0.3	1.0±0.5
	7.1	1.4±0.5	0.1±0.1
	19.2	1.9±0.2	0.0
	2.6	1.7±0.4	0.2±0.2
	12.3	2.0	0.9±0.4
	2.3	1.5±0.4	0.1±0.1
	9.1	1.8±0.3	0.8±0.5
	19.4	1.8±0.3	0.4±0.3
	2.4	1.7±0.4	0.4±0.2
	19.3	1.7±0.4	0.8±0.5
	14.3	2.0	2.0
	9.2	2.0	0.0±0.1
	12.2	1.9±0.1	1.8±0.3
	14.2	0.8±0.5	0.0
	12.1	1.5±0.3	0.6±0.3
	2.2	1.7±0.3	0.1±0.1
	17.1	1.5±0.4	0.1±0.1
	3.2	1.4±0.4	0.2±0.2
	5.2	2.0	0.3±0.2
pCGlb::BP178_KDEL(R11)	4.1	1.5±0.4	0.3±0.1
	9.1	1.9±0.2	0.9±0.3
	10.1	1.9±0.2	1.0±0.3
	20.2	1.8±0.3	0.5±0.3
	20.3	2.0±0.1	0.3±0.1
	20.5	1.8±0.3	0.3±0.1

Construct	Code	Non inoculated	Inoculated
pCGlb::BP178_KDEL(R11)	20.6	1.9±0.1	0.4±0.2
	20.7	1.5±0.5	0.3±0.2
	2.1	2.0±0.1	0.9±0.3
	7.2	2.0±0.1	0.8±0.3
	8.1	2.0	0.2±0.3
	2.4	1.9±0.2	0.1±0.1
	8.3	2.0	0.6±0.2
pCGluB-1::Cec A_KDEL(MC3)	1.2	1.8±0.4	0.3±0.3
	1.4	2.0	0.1±0.1
	2.3	2.0	0.0±0.1
	2.5	2.0	0.8±0.4
	2.6	1.8±0.3	1.9±0.1
pCGluB-1::Cec A(A5)	3.2	1.2±0.4	0.4±0.2
	5.2	1.5	0.3±0.2
pCGluB-1::BP192_KDEL (MC2, R4, S8)	5.2	2.0	1.6±0.4
	6.2	2.0	0.1±0.1
	6.4	1.4±0.4	1.0±0.6
	9.2	1.7±0.4	0.6±0.3
	5.1	2.0	0.4±0.2
	5.2	2.0	0.9±0.4
	5.3	2.0	0.4±0.2
	5.4	2.0	1.1±0.6
	5.8	2.0	0.5±0.3
	1.1	0.8±0.5	0.1±0.1
	1.2	0.0	0.0
	1.3	1.8±0.2	0.4±0.2
	2.1	0.0	0.0
	2.2	1±0.6	0.0
	2.3	2.0	0.3±0.2

Disease symptoms are expressed as a mean of the seed germination in the presence or absence of *Dickeya* sp. 1552 10.1. The confidence interval is also indicated.

Annex 8.3 Row data on *Dickeya*-inoculated T1 *vs* T3 rice seeds expressing an *AMPder*

Construct	Code	Non inoculated T1	Non inoculated T3	Inoculated T1	Inoculated T3
pCOle18::Ole18_PRS_Cec A(R9)	12.1	2.0	1.9±0.2	1.7±0.3	1.9±0.1
	8.5	1.9±0.1	1.9±0.1	0.9±0.1	1.4±0.5
	11.2	2.0	1.8±0.3	1.0±0.4	2.0
	3.2	1.6±0.4	1.3±0.3	1.3±0.5	1.1±0.6
	16.2	2.0	2.0	1.3±0.5	1.6±0.3
	17.3		1.9±0.2		0.4±0.2
	8.1	2.0	1.6±0.3	1.1±0.6	1.5±0.4
pCGluB-4::Cec A(R10)	11.1	2.0	1.5	1.6±0.4	1.5±0.3
	12.3	2.0	1.3±0.3	0.9±0.4	0.5±0.3
	14.3	2.0	2.0	2.0	1.7±0.3
	17.3	1.8±0.3	1.9±0.2	1.0±0.5	1.1±0.5
pCGlb::BP178_KDEL(R11)	2.1	2.0±0.1	2.0	0.9±0.3	1.2±0.4
	4.1	1.5±0.4	1.9±0.2	0.3±0.1	1.5±0.3
	7.2	2.0±0.1	2.0	0.8±0.3	1.9±0.1
	8.3	2.0	1.9±0.3	0.6±0.2	1.5±0.4
	9.1	1.9±0.2	2.0	0.9±0.3	1.7±0.3
pCGluB-1::BP192_KDEL(MC2, R4, S8)	6.2	2.0	1.9±0.2	0.1±0.1	1.2±0.3
	6.4	1.4±0.4	1.7±0.4	1.0±0.6	1.3±0.5
	9.2	1.7±0.4	2.0	0.6±0.3	2.0
	5.1	2.0	2.0	0.4±0.2	1.4±0.4
	5.2	2.0	1.9±0.2	0.9±0.4	1.4±0.5
	5.3	2.0	1.8±0.3	0.4±0.2	1.3±0.5
	5.8	2.0	2.0	1.1±0.6	1.8
	1.3	1.8±0.2	1.8±0.3	0.4±0.2	1.1±0.5
	2.3	2.0	2.0	0.3±0.2	1.7±0.3

Disease symptoms are expressed as a mean of the seed germination in the presence or absence of *Dickeya* sp. 1552 10.1. The confidence interval is also indicated.