

“Relación estructura—función de cirsina, una proteasa aspártica de *Cirsium vulgare*, y compuestos peptídicos derivados con actividad biológica.”

Bqca. Daniela Lufrano



Departamento de Ciencias Biológicas

Facultad de Ciencias Exactas

Universidad Nacional de La Plata

Año 2012

Directora: Dra. Laura S. Bakás

Co-directora: Dra. Sandra E. Vairo Cavalli

**Relación estructura-función de cirsina, una proteasa aspártica
de *Cirsium vulgare*, y compuestos peptídicos derivados con
actividad biológica**

El presente Trabajo de Tesis para optar al título de Doctor de la Facultad de Ciencias Exactas de la Universidad Nacional de La Plata, fue realizado en el Laboratorio de Investigaciones de Proteínas Vegetales (LIProVe) bajo la dirección de la Dra. Laura Bakás y la Dra. Sandra Vairo Cavalli y en el BIOCANT bajo la supervisión del Dr. Carlos Faro y la Dra. Isaura Simões. El mismo fue financiado por el Consejo Nacional de Investigaciones Científicas y Técnicas y por el Programa Erasmus Mundus.

Tesista: Bioquímica Daniela Lufrano

Año 2012

Mis sinceros agradecimientos al CONICET

y

a la Facultad de Ciencias de Exactas de la Universidad Nacional de La Plata



Abreviaturas

α -MF: alfa Mating Factor (factor de apareamiento alfa)

ARN: Ácido Ribonucleico

ADNc: Ácido Desoxiribonucleico complementario

BACE: β secretasa

BLAST: Basic Local Alignment Search Tool

DMSO: Dimetil sulfóxido

Dnasa: Desoxiribonucleasa

DNP: 1-4 Dinitrofenol

DTT: Ditioneitol

E-64: Ácido Trans-epoxisuccinil-L-leucil-amido (4-guanidinio) enediamintetraacético

EDANS: Ácido 5-[(2-aminoetil)amino]naftalen-1-sulfónico

EDTA: Ácido Etilen Diamin Tetraacético

Endo H: Endoglicosidasa H

Enzima TEV: Enzima del virus del grabado del tabaco (Tobacco Etch Virus)

DABCYL: Ácido 4-((4-(dimetilamino)fenil)azo)benzoico

GlcNAc: N- AcetilGlucosamina

HPLC: High-Performance Liquid Chromatography (Cromatografía líquida de alto rendimiento)

IMAC: Cromatografía de Afinidad de Iones Inmovilizados

IPTG: Isopropil β -D-1-TioGalactopiranosido

Medio LB: Medio Luria-Bertani

Medio LB/Amp: Medio LB conteniendo Ampicilina

Medio YCB: Medio Yeast Carbon Base (Medio Base sin Carbono para Levaduras)

Medio PDA: Potato Dextrose Agar (agar papa dextrosa)

Medio YP: medio peptona- extracto de levadura

Medio YPGlu/Gal: medio YP con agregado de Glucosa o Galactosa

MCA: 7-amino-4-metoxicumarina

MALDI-TOF MS: Matrix-Assisted Laser Desorption/Ionization-Time Of Flight Mass Spectroscopy (Espectroscopía de Masa Laser de Desorción/Ionización Asistida por Matriz-Analizador de Tiempo de Vuelo)

m/z: Relación masa/carga+1

LC/MS/MS: Liquid Chromatography/Mass Spectroscopy/Mass Spectroscopy (Cromatografía Líquida/Espectrometría de Masas/Espectrometría de Masas)

ON: Overnight (durante toda la noche)

PAs: Proteasas Aspárticas

PCR: Polymerase Chain Reaction (Reacción en Cadena de la Polimerasa)

PDB: Protein Data Bank (Base de datos de proteínas)

PMF: Peptide Mass Finger Print (Huella Peptídica)

PNGasa F: Peptide- N- Glicosidasa (Péptido N-glicosidasa F)

PSI: Plant Specific Insert (Inserto Específico de Plantas)

procirsinaD103A: mutante de sitio activo (se muta el Asp 103 por una Ala)

procirsina Δ PSI: mutante sin el dominio PSI

PVDF: Difloruro de Polivinilideno

RP-HPLC: Reversed phase HPLC (HPLC de Fase Reversa)

RT-PCR: Reverse Transcription- Polymerase Chain Reaction (Reacción en Cadena de la Polimerasa con Transcripción Reversa)

SAPLIPs: Saposin-Like Protein Family (Familia de Proteínas Similares a Saposinas)

SDS-PAGE: Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (Electroforesis Desnaturalizante en Geles de Poliacrilamida)

TA: Temperatura Ambiente

TEMED: N,N,N',N'-tetrametiletlenodiamina

UFC: Unidades Formadoras de Colonias

VC: Volumen de Columna

wt: del inglés *wild-type*, refiere a la proteína cuya secuencia se corresponde con la nativa

ÍNDICE DE CONTENIDOS

Introducción General	2
Peptidasas	2
Clasificación	2
Relevancia	2
Importancia biológica	3
Peptidasas vegetales	4
Proteasas Aspárticas	4
Características generales	5
Organización de la estructura primaria	5
Organización tridimensional	7
Mecanismo catalítico	9
Zimógenos y mecanismos de inactivación	11
Proteasas Aspárticas de Plantas	12
Estructura tridimensional	15
Inserto específico de planta	16
Mecanismos de procesamiento	19
Objetivos	22
Material Vegetal	23
Materiales y Métodos Generales	28
Clonado de preprocirsina	30
<i>Extracción del ARN total</i>	30
<i>Síntesis de la doble cadena híbrida ARN-ADNc</i>	30
<i>Amplificación del ADNc que codifica para la secuencia completa de la preprocirsina</i>	30
<i>Amplificación del plásmido pGEM-preproCV en E. coli</i>	31
Técnicas Generales	32
A. Técnicas de Biología Molecular	32
<u>Técnica nº1:</u> Electroforesis de ácidos nucleicos (ADN y ARN) en gel de agarosa 1% (p/v)	32
<u>Técnica nº2:</u> Reacción en cadena de la polimerasa	32
<u>Técnica nº3:</u> Purificación de ADN desde un gel de agarosa	33


<u>Técnica nº4</u> : Preparación de células bacterianas químicamente competentes	33
<u>Técnica nº5</u> : Clonado de un inserto en el sistema pGEM-T Easy	34
<u>Técnica nº6</u> : Aislamiento del ADN plasmídico (minipreparaciones)	35
<u>Técnica nº7</u> : Digestión con enzimas de restricción	36
<u>Técnica nº8</u> : Subclonado de un inserto en los vectores de expresión pET-23d y pKLAC-1	37
B. Técnicas de purificación y análisis de proteínas	37
<u>Técnica nº9</u> : Cromatografía de exclusión molecular	37
<u>Técnica nº10</u> : Cromatografía de intercambio aniónico	38
<u>Técnica nº11</u> : Cromatografía de afinidad	38
<u>Técnica nº12</u> : Cromatografía líquida de alta performance (HPLC)	39
<u>Técnica nº13</u> : Cuantificación de proteínas	39
<u>Técnica nº14</u> : Electroforesis desnaturizante (SDS-PAGE) discontinua	40
<u>Técnica nº15</u> : Coloración de geles de acrilamida con Coomassie Brilliant Blue R-250	41
<u>Técnica nº16</u> : Tinción de proteínas con Nitrato de Plata	42
<u>Técnica nº17</u> : Transferencia de proteínas a membranas de PVDF	42
<u>Técnica nº18</u> : Inmunodetección o Western blot	43
<u>Técnica nº19</u> : Digestión trípica en geles de poliacrilamida	44



Capítulo Procirsina

Introducción	47
Materiales y Métodos	49
Obtención de procirsina y sus mutantes en <i>Escherichia coli</i>	49
I. Obtención de las construcciones para expresión de procirsina y sus mutantes	49
<i>Construcción pET-procirsina, pET-procirsinaD103A y pET-procirsinaΔPSI</i>	50
II. Análisis filogenético	53
III. Expresión heteróloga de procirsina <i>wild type</i> y sus mutantes	54

<i>Expresión wt-procirina, procirinaD103A y procirinaΔPSI como cuerpos de inclusión</i>	54
<i>Purificación de wt-procirina, procirinaD103A y procirinaΔPSI</i>	55
<i>Detección de la actividad endopeptidásica sobre un sustrato sintético (Screening)</i>	56
<i>Estudio de procesamiento proteolítico de wt-procirina</i>	56
<i>Digestión triptica en geles de poliacrilamida</i>	57
IV. Caracterización bioquímica de wt-procirina recombinante	57
<i>Actividad proteolítica de wt-procirina recombinante sobre diferentes sustratos modelo</i>	57
<i>Determinación del perfil de pH y temperatura</i>	58
<i>Ensayos de inhibición</i>	59
<i>Determinación de parámetros cinéticos</i>	59
<i>Titulación del sitio activo con pepstatina A</i>	59
<i>Determinación de la especificidad de sustrato sobre la cadena β de la insulina oxidada</i>	60
<i>Digestión de κ-caseína</i>	60
V. Actividad coagulante de leche	61
VI. Caracterización bioinformática de wt-procirina	61
Resultados	62
Obtención de procirina y sus mutantes en <i>Escherichia coli</i>	62
I. Clonado de ADNc y análisis de la secuencia	62
II. Análisis filogenético	66
III. Expresión heteróloga de wt-procirina y sus mutantes	69
IV. Caracterización bioquímica de wt-procirina recombinante	75
<i>Actividad proteolítica de wt-procirina recombinante sobre diferentes sustratos modelo</i>	75
<i>Determinación del perfil de pH y temperatura</i>	76
<i>Determinación de la especificidad de sustrato sobre la cadena β de la insulina oxidada</i>	78
<i>Determinación de parámetros cinéticos</i>	79

<i>Ensayos de inhibición</i>	79
V. Clivaje de κ-caseína y actividad coagulante de leche	80
VI. Caracterización bioinformática de wt-procirisina	81
Discusión	87
 Capítulo PSI (Plant Specific Insert)	
Introducción	94
Materiales y Métodos	98
Obtención de PSI en <i>Kluyveromyces lactis</i>	98
I. Obtención de construcciones para expresión de PSI en <i>K. lactis</i>	100
<i>Construcción pKLAC-α-MF-PSI(His)₆ y pKLAC-α-MF-PSIN86S(His)₆</i>	102
<i>Análisis por dinámica molecular de la influencia de la cola de polihistidinas terminal en la estructura tridimensional/conformación del dominio PSI</i>	103
<i>Construcción pKLAC- PSI(His)₆ y pKLAC- PSIN86S(His)₆</i>	104
II. Expresión y purificación de PSI en <i>K. lactis</i>	106
<i>Preparación de células <i>K.lactis</i> GG799 competentes</i>	106
<i>Linealización de las construcciones y electrotransformación de las células GG799</i>	106
<i>Identificación de células con el casete de expresión correctamente integrado y selección de transformantes</i>	107
<i>Expresión de α-MF-PSI(His)₆, α-MF-PSN86SI(His)₆, PSI(His)₆, PSIN86S(His)₆</i>	108
III. Análisis de la glicosilación de los productos expresados en <i>K. lactis</i>	108
<i>Tratamiento de las formas de PSI obtenidas en <i>K. lactis</i> con endoglicosidasas</i>	108
<i>Análisis in silico de los sitios putativos de glicosilación</i>	109

IV. Producción de las proteínas de secreción	110
<i>Purificación de los productos de expresión α-MF-PSI(His)₆ y α-MF-PSIN86S(His)₆, PSI(His)₆, PSIN86S(His)₆</i>	110
<i>Corte in vitro del producto α-MF-PSI(His)₆ con la enzima Kex-2</i>	111
<i>Optimización del corte de la cola de histidinas de la proteína PSIN86S(His)₆ con la enzima TEV</i>	111
<i>Corte de la proteína PSIN86S(His)₆ con TEV</i>	112
<i>Análisis de los productos de expresión por técnicas de proteómica</i>	112
Ensayos antimicrobianos realizados con diferentes formas de PSI expresadas en <i>K. lactis</i>	113
<i>Actividad antibacteriana</i>	113
Test de dilución en agar	113
Test en medio líquido	113
Test de KIRBY-BAUER o de difusión en disco	113
<i>Actividad antifúngica</i>	114
Test de co-crecimiento en placa	114
Test de KIRBY-BAUER o de difusión en disco	114
Resultados	115
Obtención de PSI en <i>Kluyveromyces lactis</i>	115
A. Expresión de proteínas secretadas usando pKLAC1	115
AI. Obtención de los vectores pKLAC-α-MF-PSI(His)₆ y pKLAC-α-MF-PSIN86S(His)₆ para la expresión en <i>K. lactis</i>	115
AII. Expresión de PSI en <i>K. lactis</i> a partir de las construcciones pKLAC-α-MF-PSI(His)₆ y pKLAC-α-MF-PSIN86S(His)₆	118
<i>Selección de clones de <i>K. lactis</i> conteniendo las secuencias del PSI integradas en el genoma como casete de expresión y optimización del tiempo de expresión</i>	118
AIII. Análisis de la glicosilación de los productos expresados en <i>K. lactis</i> a partir de los casetes con las secuencias α-MF-PSI(His)₆ y α-MF-PSIN86S(His)₆	122

<i>Tratamiento de las formas de PSI obtenidas en K.lactis con endoglicosidasas</i>	122
<i>Análisis in silico de los sitios putativos de glicosilación</i>	125
AIV. Producción de las proteínas de secreción α-MF-PSI(His)₆ y α-MF-PSIN86S(His)₆	128
<i>Purificación de los productos de expresión α-MF-PSI(His)₆ y α-MF-PSIN86S(His)₆</i>	128
<i>Corte in vitro del producto α-MF-PSI(His)₆ con la enzima Kex-2</i>	133
<i>Análisis de los productos de expresión de los casetes conteniendo α-MF por técnicas de proteómica</i>	134
B. Expresión de las proteínas PSI(His)₆ y PSIN86S(His)₆ usando una estrategia alternativa en K. lactis	144
<i>Análisis por dinámica molecular de la influencia de la cola de polihistidinas terminal en la estructura tridimensional/conformación del dominio PSI</i>	145
BI. Obtención de los vectores pKLAC-PSI(His)₆ y pKLAC-PSIN86S(His)₆ para la expresión del PSI y su mutante de sitio de glicosilación en K. lactis	148
BII. Expresión de PSI en K. lactis a partir de las construcciones pKLAC-PSI(His)₆ y pKLAC-PSIN86S(His)₆	150
<i>Selección de clones de K. lactis conteniendo las secuencias del PSI integradas en el genoma como casete de expresión y optimización del tiempo de expresión</i>	150
BIII. Análisis de la glicosilación de los productos expresados en K. lactis a partir de los casetes con las secuencias PSI(His)₆ y PSIN86S(His)₆	152
<i>Tratamiento de las formas de PSI obtenidas en K.lactis con endoglicosidasas</i>	152
<i>Purificación de los productos de expresión PSI(His)₆ y PSIN86S(His)₆</i>	154
<i>Corte de la proteína de fusión PSIN86S(His)₆ con TEV</i>	155
<i>Análisis por técnicas de proteómica del producto de expresión del casete PSIN86S(His)₆</i>	158
Ensayos antimicrobianos realizados con diferentes formas de PSI expresadas en K. lactis	162
<i>Actividad antibacteriana</i>	162

<i>Actividad antifúngica</i>	164
Discusión	168
Conclusiones generales y perspectivas futuras	176
Referencias bibliográficas	180
Anexos	
Anexo 1	195
Anexo 2	197

Introducción General



Introducción General

Peptidasas

Clasificación

Las peptidasas (EC 3.4), también denominadas proteasas o proteinasas, son enzimas capaces de hidrolizar enlaces peptídicos. Esta actividad proteolítica puede liberar aminoácidos de los extremos de cadenas polipeptídicas (exopeptidasas) o producir cortes en posiciones internas de las mismas (endopeptidasas). Las exopeptidasas se diferencian en aminopeptidasas y carboxipeptidasas según actúen sobre el extremo N-terminal o C-terminal de su sustrato, respectivamente. Por su parte, la base de datos MEROPS (<http://merops.sanger.ac.uk>) clasifica a las endopeptidasas de acuerdo al mecanismo catalítico en siete grupos: aspárticas, cisteínicas, glutámicas, asparagínicas, treonínicas, serínicas y metalopeptidasas. A su vez, dentro de cada tipo catalítico, las enzimas son agrupadas en familias según la similitud de las secuencias aminoacídicas, y en clanes basados en las relaciones evolutivas y estructurales (Rawlings et al., 2012).

Relevancia

Desde el punto de vista de la industria, las proteasas constituyen el grupo de enzimas con mayor relevancia en el mercado mundial, representando cerca del 60% del total de las enzimas comercializadas. Debido a que la proteólisis cambia las propiedades químicas, físicas, biológicas e inmunológicas de las proteínas, las enzimas proteolíticas cumplen un rol muy importante en la biotecnología (González-Rábade et al., 2011), siendo muy diversos los campos de aplicación: tecnología alimentaria, industrias farmacéuticas y manufacturas de detergente, entre otras (Feijoo-Siota and Villa, 2011).

Importancia biológica

En los organismos vivos, la amplia distribución entre plantas, animales y microorganismos demuestra que las proteasas son esenciales para todos ellos, llevando a cabo importantes roles fisiológicos en numerosos procesos. Las peptidasas fueron inicialmente reconocidas como enzimas proteolíticas del jugo gástrico involucradas en la degradación inespecífica de las proteínas dietarias. Sin embargo, avances posteriores han provisto una nueva visión del mundo proteolítico. Además de la hidrólisis proteica inespecífica, las proteasas también actúan como enzimas de procesamiento que realizan clivajes eficientes, limitados y selectivos de sustratos específicos, iniciando así decisiones irreversibles a nivel post-traduccionales que influyen en muchos procesos biológicos. A modo de ejemplo, en humanos los eventos de procesamiento proteolítico son fundamentales en ovulación, fertilización, desarrollo embrionario, formación ósea, control de la homeostasis, remodelado tisular, extensión neuronal, presentación antigénica, regulación del ciclo celular, activación y migración celular, en procesos inmunes e inflamatorios, cicatrización de heridas, angiogénesis y apoptosis (Puente et al., 2003; Barrett et al., 2004; Muñoz-Pinedo, 2012). Al igual que en el resto de los organismos, las plantas cuentan con un elevado número de proteasas, así en el genoma de *Arabidopsis thaliana* se codifican más de 800 proteasas que están distribuidas en más de 60 familias de acuerdo a su homología de secuencia aminoacídica, las cuales pertenecen a 30 clanes diferentes (Tsiatsiani et al., 2011), cada uno de ellos con un mismo ancestro evolutivo. Las proteasas vegetales cumplen con una gran diversidad de funciones biológicas; además de la degradación de proteínas no funcionales en aminoácidos, la ruptura selectiva de proteínas reguladoras por la vía de la ubiquitina/proteasoma controla aspectos claves del crecimiento, desarrollo y defensa de la planta frente a las señales ambientales. De esta manera, el cuadro emergente es que las proteasas vegetales son reguladores claves de una enorme variedad de procesos biológicos que están involucrados en todos los aspectos del ciclo de vida de las plantas, desde la movilización de proteínas de reserva durante la germinación hasta el inicio de la muerte celular y senescencia programadas (Schaller, 2004; van der Hoorn, 2008; Pesquet, 2012). Algunos procesos particulares de plantas en que las proteasas han mostrado estar involucradas son: formación de la cubierta de semillas y maduración de las mismas, destino de células epidermales, formación del xilema, desarrollo del estoma, biogénesis de cloroplastos, movilización de nutrientes, respuestas a estímulos ambientales y factores de virulencia y senescencia (van der Hoorn, 2008; Tsiatsiani et al., 2011).

Peptidasas vegetales

La diversidad funcional de las peptidasas vegetales se correlaciona con la expresión específica de las mismas en tiempo y espacio y su acumulación en diferentes compartimentos subcelulares (van der Hoorn, 2008).

Durante los últimos años con el número de genomas completamente secuenciados creciendo día a día, quedó claro que las proteasas representan el 5,03% del genoma de *A. thaliana*, 4,58% en el caso de *Oryza sativa* y 3,53% para *Vitis vinífera* (de http://merops.sanger.ac.uk/cgi-bin/compge_index?type=P).

A pesar de la abundancia de proteasas en la biología vegetal y de sus importantes implicancias, sólo unos pocos sustratos naturales han sido identificados. No obstante, con el crecimiento exponencial de datos científicos emergió una nueva disciplina de “ómicas”, la “Proteasa Degradómica” (López-Otín and Overall, 2002; Huesgen and Overall, 2012) en la cual son utilizadas aproximaciones genómicas y proteómicas para identificar proteasas, inhibidores de proteasas y repertorios de sustratos de proteasas (o degradomas) en una variedad de organismos (Tsiatsiani et al., 2011; Wang et al., 2012). En este contexto, el uso reciente de métodos de proteómica tales como la cromatografía diagonal fraccional combinada “COFRADIC” (Gevaert et al., 2003; Staes et al., 2008) y el marcado isotópico del amino terminal de sustratos “TAILS” (Kleifeld et al., 2010), han superado varias dificultades en los métodos convencionales comúnmente utilizados en la proteómica e introducidas por el uso de geles de electroforesis en condiciones desnaturalizantes (SDS-PAGE). Últimamente, la proteólisis ha ganado terreno en el campo de la investigación de plantas y un número creciente de laboratorios están activamente dedicados al estudio del modo de acción, propiedades bioquímicas, función putativa y sustratos naturales de las proteasas (Tsiatsiani et al., 2011).

Proteasas Aspárticas

Históricamente, la primera PA estudiada en los inicios del siglo XIX, fue la pepsina del estómago de animales. En 1874 cuando J. D. Hooker, inspirado directamente por Darwin, detectó la homología de la nepenthesina aislada de la planta carnívora *Nepenthes*, con la pepsina (Blundell et al., 1998; Fruton, 2002) comenzó a estudiarse la familia de la pepsina (EC.3.4.23), familia A1 de la base de datos MEROPS. Durante el último siglo el número de PAs encontradas se incrementó significativamente, y para agosto de 2012, había 4693 miembros

descriptos, agrupados en 16 familias distribuidas a su vez en 6 clanes en la base de datos MEROPS. A lo largo de este tiempo, las proteasas aspárticas fueron encontradas en todo tipo de organismos: virus, bacterias, arqueobacterias, protozoos, plantas y animales.

Los mayores avances en el conocimiento de las PAs se han logrado en las tres últimas décadas (Tang and Wong, 1987; Davies, 1990; Dunn, 2002; Rawlings and Barrett, 2004) con el interés causado al descubrir que las mismas están involucradas en enfermedades humanas, como el cáncer (Benes, 2008), la malaria (Coombs et al., 2001; Klemba and Goldberg, 2002; Weber and Wang, 2010), el SIDA y el Alzheimer (Weber and Wang, 2010).

Características generales

Las PAs están caracterizadas por ser activas en el rango de pH ácido, ser inhibidas por pepstatina A (un hexapéptido producido por *Streptomyces spp*), mostrar especificidad por los enlaces entre residuos grandes e hidrofóbicos y haber conservado una estructura bilobular, con dos dominios estructuralmente similares. Cada uno de estos dominios aporta un residuo ácido aspártico localizado en el fin de un bucle conocido como el “ ψ loop” (Tang and Wong, 1987; Davies, 1990; Rawlings and Barrett, 2004) constituyendo entre ambos el sitio catalítico de unión al sustrato. Las PAs tienen secuencias características en la región de los dos residuos aspárticos catalíticos dentro de un motivo cuyo centro contiene la secuencia Asp-Xbb-Gly, donde Xbb es Thr o Ser, conocido como el motivo “DT/SG” (Tang and Wong, 1987; Davies, 1990; Rawlings and Barrett, 2004).

La homología estructural y secuencial entre los dos dominios de las proteasas aspárticas, condujeron a Tang a sugerir que ambos dominios de PAs podrían haber evolucionado por duplicación génica de una proteína ancestral de aproximadamente 150 amino ácidos teniendo un plegado similar a cada dominio de pepsina (Tang et al., 1978). Esta simetría bilobular permitió proponer modelos de proteasas de retrovirus como homólogas diméricas de pepsina, las cuales fueron confirmadas luego por las estructuras cristalinas de las proteasas del HIV (Blundell et al., 1998).

Organización de la estructura primaria

La mayoría de las PAs de la familia A1 son sintetizadas como zimógenos de simple cadena de alrededor de 35 kDa, que contienen en su extremo N-terminal un péptido señal y un

prosegmento. El péptido señal dirige al precursor hacia el retículo endoplasmático donde es escindido, mientras que el prosegmento de aproximada 50 aminoácidos y rico en residuos básicos, es removido por clivaje proteolítico a pH ácido, resultando en una enzima activa.

Si bien la naturaleza del prosegmento aún no está totalmente esclarecida, esta región podría ser importante en procesos tales como el inicio del correcto plegamiento, estabilidad proteica, bloqueo del sitio catalítico, activación dependiente del pH, y direccionamiento intracelular del zimógeno. Más aún, el rol e importancia del prosegmento para varias PAs depende de sus funciones particulares y del mecanismo de la funcionalidad (Horimoto et al., 2009). Ha sido propuesto que el residuo Lys36p (de acuerdo a la numeración del pepsinógeno porcino, donde p designa al prosegmento) es crítico para la estabilidad y la inactividad del zimógeno de pepsina a pH neutro, y que podría ser requerido para el correcto plegado del precursor (Foltman, 1988). Este residuo se encuentra en una región altamente conservada de los zimógenos de las PAs de vertebrados. Los puentes de hidrógeno, las interacciones electrostáticas e hidrofóbicas formadas entre los residuos del prosegmento y la pepsina madura, incluyendo las interacciones de la tríada Tyr37p, Lys36p, Tyr9p con los dos aspartatos catalíticos, Asp32 y Asp215, podrían resultar fundamentales para estabilizar el pepsinógeno (Richter et al., 1998) al fortalecer la inhibición por parte del propéptido (Francky et al., 2001), así como también para el correcto alineamiento de los residuos del centro activo (Richter et al., 1999).

Algunos zimógenos de PA tienen secuencias internas que son removidas durante el procesamiento, como el inserto PSI de la mayoría de las proteasas típicas de origen vegetal (Mutlu and Gal, 1999; Simões and Faro, 2004) y catepsina D (Faust et al., 1985). Otras, como las napsinas, tienen extensiones en el C-terminal de alrededor de 20 aminoácidos de longitud, de función desconocida (Tatnell et al., 1998), y algunas otras, como la β -secretasa de Alzheimer, tienen una extensión C-terminal con un dominio transmembrana y una cola citosólica de 24 residuos (Hussain et al., 1999; Sinha et al., 1999; Vassar et al., 1999; Yan et al., 1999; Lin et al., 2000). Sin embargo, la extensión C-terminal de la β -secretasa no se escinde durante la maduración, como sucede con las inserciones de la mayoría de las PAs típicas de plantas y catepsina D.

Como ya se ha mencionado, los dos aspartatos catalíticos están ubicados dentro del motivo "DTG" altamente conservado. La secuencia que flanquea el Asp catalítico puede ser descripta en más detalle, por Xaa-Xaa-Asp-Xbb-Gly-Xbb, donde Xaa representa un aminoácido hidrofóbico y Xbb representa un residuo serina o treonina (Rawlings and Barrett, 2004).

Otros residuos aminoacídicos conservados son la Tyr75 y Thr77 (numeración de la pepsina). Estos residuos están localizados en un bucle altamente móvil conocido como “solapa” (“*flap*”), el cual se proyecta hacia afuera y por encima de la hendidura del sitio catalítico, formando parte del subsitio S1 e interactuando con el sustrato de la proteasa (Rawlings and Barrett, 2004).

La mayoría de los miembros de la familia de la pepsina tienen seis residuos cisteína localizados en posiciones conservadas e involucrados en 3 puentes disulfuro, con un patrón también conservado (Rawlings and Barrett, 2004).

Organización tridimensional

Hasta el presente han sido depositadas en la base de datos de estructuras macromoleculares, Protein Data Bank (PDB), 848 estructuras tridimensionales obtenidas mediante cristalografía de rayos X, correspondientes a PAs (<http://www.rcsb.org/pdb/home/home.do> visitada en Julio de 2012). La comparación entre estas estructuras muestra un considerable grado de similitud.

La Fig. 1 representa la estructura cristalina de la pepsina porcina, el miembro representativo de la familia A1. La estructura secundaria en conjunto consiste casi enteramente en hoja- β con muy poco contenido de hélice α . La molécula es bilobada con dos dominios de estructura similar relacionados por un eje de rotación bivalente originalmente observado por Tang y colaboradores en endotiapepsina y penicilopepsina (Tang et al., 1978). Una gran hendidura, de aproximadamente 40 Å de longitud, corre a través de toda la molécula y separa los dos dominios, habiéndose observado la unión a la misma de inhibidores oligopeptídicos (Davies, 1990). Bajo los dos dominios, una hoja β de seis hebras antiparalelas forma el piso de la molécula, al cual cada dominio contribuye con 3 hebras.

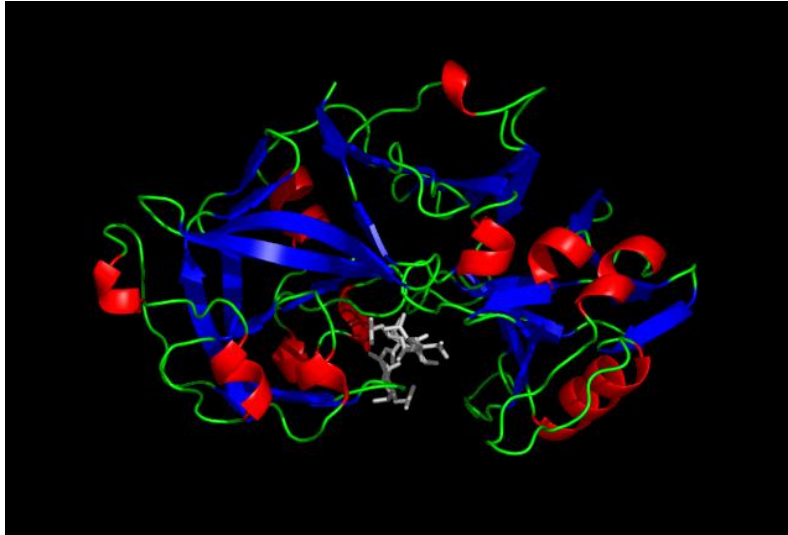


Fig. 1. Estructura cristalina de una PA. Muestra la estructura tridimensional de la pepsina porcina código de PDB: 4PEP (Sieleki et al., 1991).

Los residuos ácido aspártico catalíticos (Asp32 y 215 en la numeración de pepsina) están localizados en los extremos de dos bucles ψ en el medio de la extensa hendidura, donde las estructuras de las PAs están conservadas (Davies, 1990). Un participante esencial en el proceso catalítico es la molécula de agua ligada a los carboxilos de los Asp, la cual actúa como nucleófilo para iniciar la catálisis básica general.

El re-refinamiento de la estructura de rayos X de la apo plasmepsina II de *Plasmodium falciparum*, sugiere que los dos grupos carboxilos de la díada catalítica no son coplanares (Robbins et al., 2009), en notable contraste con la vasta mayoría de estructuras de PAs analizadas. La evidencia de la no coplanaridad de los aspartatos catalíticos fue reafirmada por análisis de simulaciones basadas en dinámica molecular (MD) en entorno acuoso de plasmepsina II, β -secretasa humana, y proteasa de HIV. En el caso de β -secretasa la disposición no coplanar es también prevalente en las simulaciones de la proteasa unida a inhibidores. Cálculos de mecánica cuántica (QM) arrojan más pruebas que sustentan que la disposición no coplanar se encuentra energéticamente favorecida antes de la catálisis en las PAs de origen eucariótico (Fig. 2). Llamativamente, la orientación coplanar de la díada catalítica se observa en las simulaciones de MD de la proteasa de HIV a una temperatura de 100 K pero no a 300 K, lo cual indica que el arreglo no coplanar es la conformación entrópicamente favorable. Este hallazgo sugiere por tanto, que la orientación coplanar en las estructuras cristalográficas de las proteasas apo aspárticas están promovidas por la muy baja temperatura utilizada para la recolección de datos, usualmente alrededor de 100 K (Friedman and Caflisch, 2010).

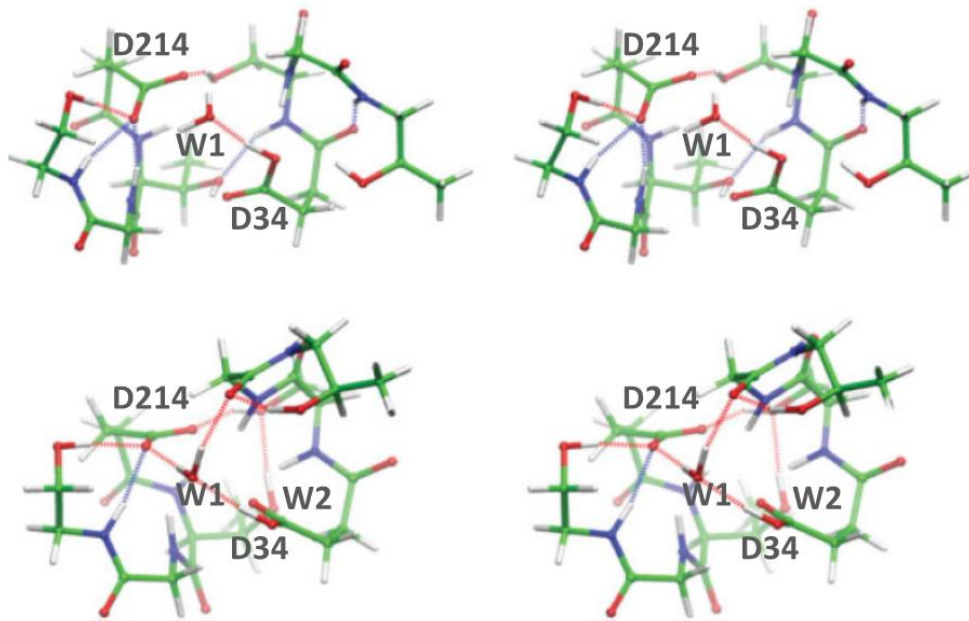


Fig. 2. Optimización de la geometría basada en QM (SCC-DFTB) de un modelo del sitio catalítico en proteasas aspárticas eucarióticas (Asp-Thr-Gly-Ser/Asp-Ser-Gly-Thr) mostrada en estéreo. La conformación no coplanar (arriba) optimizada es más estable que la coplanar (abajo). Figura tomada de Friedman y Casflisch, 2010.

Mecanismo catalítico

Las propuestas actuales para el mecanismo catalítico de las PAs están mayormente basadas en estructuras cristalográficas de complejos formados por las proteasas con inhibidores oligopeptídicos, los que poseen análogos no hidrolizables del enlace peptídico escindible (Coates et al., 2006).

El Asp215 actúa como una base general y está cargado, mientras que el Asp32, actúa como ácido general y está protonado. Una extensa red de puentes de hidrógeno conecta el residuo conservado Tyr75 (numeración de pepsina) del *flap* con el Asp32 del sitio activo, involucrando una segunda molécula de agua; el grupo oxhidrilo del residuo Thr218, localizado a un puente de hidrógeno de distancia del carboxilo del Asp215, protege a este carboxilo de la protonación en medio ácido. La misma cadena de enlaces de hidrógeno ha sido propuesta como una característica esencial de complejos entre enzimas de tipo pepsina y sus inhibidores (Andreeva and Rumsh, 2001). La extensa red de enlaces de hidrógeno hace que el sitio activo sea bastante rígido, a excepción del *flap* que tiene alta movilidad cuando la enzima no tiene sustrato ni inhibidor unido en la hendidura del sitio activo. El *flap* encierra al inhibidor/sustrato

unido en la hendidura, excluyendo el solvente, mientras se torna considerablemente menos móvil (Dunn, 2002).

El mecanismo de acción propuesto para las PAs tipo pepsina (Fig. 3) involucra el primer paso de protonación del oxígeno del grupo carboxilo por el protón compartido por los dos ácidos aspárticos catalíticos. Sigue el ataque nucleofílico por parte de la molécula de agua sobre el carbono carboxílico con transferencia de un protón al Asp32, resultando en la formación de un intermedio tetraédrico.

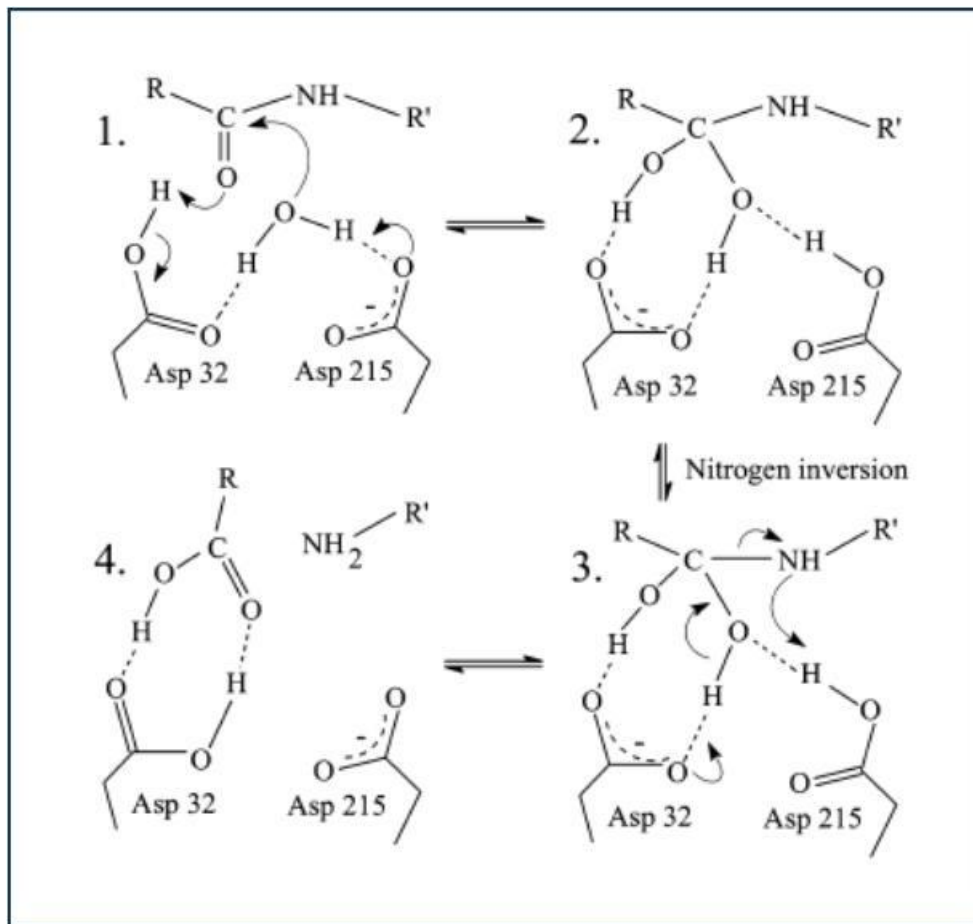


Fig. 3. Mecanismo catalítico propuesto para proteasas aspárticas (Veerapandian et al., 1992). Este mecanismo está basado en la estructura de rayos X de la enzima endotiapepsina unida a un inhibidor. Se propone que la molécula de agua fuertemente unida a los aspartatos en la enzima nativa ataca nucleofílicamente al enlace carbonilo escindible. El intermediario tetraédrico resultante (2) está estabilizado por uniones de hidrógeno al carboxilo del aspartato 32, cargado negativamente. La ruptura del enlace $C-N$ escindible está acompañada por la transferencia de un protón al grupo amino saliente ya sea desde el Asp215 (con inversión del nitrógeno, 3) o desde el seno del solvente. Los puentes de hidrógeno se encuentran indicados por líneas punteadas (figura tomada de Coates et al., 2006).

Los productos son formados a través de la protonación del átomo de nitrógeno, ya sea desde la molécula de solvente o desde los grupos carboxilos catalíticos (James, 2004). La protonación del residuo Asp215 y la carga negativa del Asp32 en el complejo del estado de transición ha sido puesta en evidencia mediante diferentes técnicas, tales como cristalografía de rayos X con resolución atómica de los complejos con inhibidores y difracción de neutrones de cristales. Un estudio similar de la enzima endotiapepsina fúngica nativa estableció que posee un dipéptido Ser-Thr unido al sitio catalítico, probablemente originado por autólisis del N-terminal de la enzima durante el proceso de cristalización, lo cual es consistente con los reportes clásicos. Los estudios llevados a cabo por RMN han confirmado el hallazgo de la barrera baja y los enlaces simples de hidrógeno en los complejos con los análogos de los estados de transición (Coates et al., 2006).

Zimógenos y mecanismos de inactivación

Las PAs son producidas mayormente en la forma de zimógeno para prevenir la degradación proteica indeseada. En 1999, con la resolución de las estructuras cristalinas del pepsinógeno porcino (Sieleki et al., 1991; Hartsuck et al., 1992), de la proplasmepsina II (Khazanovich Bernstein and James, 1999) y de la profitepsina, precursor de una PA de cebada (Kervinen et al., 1999), se describieron tres modos diferentes de inactivación o bloqueo del sitio activo en precursores de PAs: (A) la típica, similar a pepsinógeno, donde los residuos cargados positivamente del propéptido interactúan con los residuos ácido aspártico catalíticos; (B) el mecanismo tipo profitepsina, en el que los residuos con carga positiva que interactúan con los aspartatos catalíticos son un par Lys/Tyr conservado de la secuencia madura de la enzima, y no del propéptido, y (C) el mecanismo tipo proplasmepsina II, donde el propéptido envuelve al dominio C-terminal ocasionando una distorsión en el sitio activo, separando los dos bucles ψ y posicionando los dos Asp catalíticos de modo de prevenir la activación básica general de una molécula de agua nucleofílica. La Fig. 4 muestra las estructuras tridimensionales de los zimógenos de PAs representativos de los tres mecanismos de inactivación diferentes.

La conversión del zimógeno inactivo hacia su forma activa es un proceso complejo. Debido a las diferencias en los mecanismos de activación entre las PAs, no existe un único modelo propuesto. Koelsch y colaboradores plantearon tres mecanismos de activación: (1) autoprosesamiento absoluto del prosegmento (zimógenos gástricos), (2) autoprosesamiento parcialmente asistido por otras enzimas (peptidasas lisosomales y peptidasas vacuolares) y (3) procesamiento completamente asistido (pro-renina) (Horimoto et al., 2009).

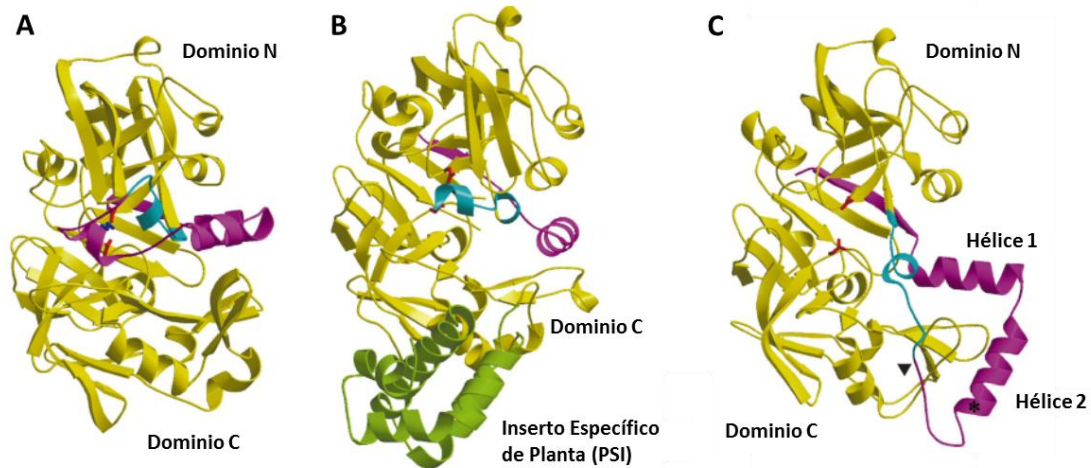


Fig. 4. Estructuras cristalinas de zimógenos de proteasas aspárticas. (A) pepsinógeno A porcino, (B) profitepsina, (C) proplasmepsina II. Las moléculas se muestran con sus dominios N-terminales en el extremo superior y sus dominios C-terminales en el inferior. El prosegmento está indicado en violeta, los primeros 13 residuos de la secuencia madura se muestran en turquesa y el resto de la molécula, en amarillo. Las cadenas laterales de los aspartatos catalíticos y la lisina descrita como inhibitoria en pepsinógeno y profitepsina, están mostrados en rojo y azul, respectivamente. El inserto específico de planta en la profitepsina, se indica en verde. En proplasmepsina II (C), la unión entre el segmento pro y la proteína madura (Gly124p-Ser1) se marca en la figura con un triángulo negro, mientras que el enlace escindido durante la activación (Phe112p-Leu113p) está indicado por un asterisco. Imagen tomada de Khazanovich Bernstein y James, 1999.

A pesar de las diferencias entre las estructuras de los zimógenos de las PAs, varias características comunes emergen, principalmente la ubicación de la hebra β del prosegmento dentro de la hoja β central y los grandes cambios conformacionales en el extremo N-terminal de la proteína madura. Como los prosegmentos de muchos zimógenos de PAs han mostrado ser necesarios para el correcto plegado de la proteasa madura, la hebra β del prosegmento podría actuar como una chaperona intramolecular. La participación de una región del prosegmento en la formación de la hoja β es también una forma muy segura de unir el prosegmento a la porción de la enzima madura y evitar su disociación hasta que el prosegmento sea clivado durante la activación (Khazanovich Bernstein and James, 1999).

Proteasas Aspárticas de Plantas

El Comité de Nomenclatura de la Unión Internacional de Bioquímica y Biología Molecular (NC-IUBMB) ha adoptado el nombre general de *fitepsinas* (EC 3.4.23.40) para todas las peptidasas

aspárticas de origen vegetal (NC-IUBMB, 2006). Sin embargo, en la literatura es poco frecuente encontrar un nombre universal que refiera a estas peptidasas, por el contrario, los nombres tienden a indicar las especies o tejidos de donde cada enzima ha sido purificada. De acuerdo con las base de datos MEROPS (Rawlings et al., 2012), las PAs de plantas están distribuidas entre las familias A1, A2, A3, A11, A28 y A33 del clan AA, la familia A8 del clan AC y las familias A22 y A24 del clan AD, sumando un total de 2261 miembros a julio de 2012.

Las PAs se han identificado en una amplia variedad de especies del reino vegetal (Simões and Faro, 2004; Rawlings et al., 2012) y se han aislado de diferentes tejidos, tales como semillas, hojas, flores, raíces y tubérculos: i) semillas de *Arabidopsis thaliana* (Mutlu and Gal, 1999), arroz (Doi et al., 1980; Asakura et al., 1997), cebada (Runeberg-Roos et al., 1991; Kervinen et al., 1999), trigo (Belozersky, Mikhail A. Sarbakanova and Dunaevsky, 1989; Timotijević et al., 2006; Tamura et al., 2007), cáñamo (St. Angelo et al., 1970), girasol (Park et al., 2000) pepino y calabaza (Wilimowska-Pelc, A Polanowski et al., 1983; Polanowski et al., 1985); ii) hojas de planta de tomate (Rodrigo et al., 1989); iii) hojas y tubérculos de planta de papa (Guevara et al., 1999, 2001); iv) polen de maíz (Radlowski et al., 1996) y v) flores de cardos (Tamer, 1993; Faro et al., 1995; Veríssimo et al., 1996; Domingos et al., 2000; Llorente et al., 2004; Vairo Cavalli, 2005; Sidrach et al., 2006), entre otras.

De acuerdo a los resultados obtenidos en estudios de localización *in situ* las PAs no se encuentran distribuidas al azar dentro de las plantas, sino que las mismas están localizadas muy específicamente en determinados tejidos, dependiendo de la especie (Simões and Faro, 2004). Más aún, en la célula, la actividad proteolítica es llevada a cabo por distintas proteasas direccionadas específicamente a diferentes compartimentos (Pesquet, 2012). Subcelularmente, las PAs de plantas están principalmente localizadas en vacuolas/cuerpos proteicos o en el espacio extracelular (Duarte et al., 2008).

Al igual que sus equivalentes de origen no vegetal, las PAs de plantas están caracterizadas por la presencia de dos residuos ácido aspártico organizados en dos motivos conservados, por lo general Asp-Thr/Ser-Gly (DT/SG), responsables de la actividad catalítica, la cual es máxima a valores de pH ácidos, tienen preferencia de corte por enlaces formados entre aminoácidos hidrofóbicos y grandes y en su mayoría son sensibles a pepstatina A (Mutlu and Gal, 1999; Simões and Faro, 2004). La estructura tridimensional está también conservada, consistiendo en dos dominios tipo barril β que crean una hendidura profunda en cuyo interior se encuentra ubicado el sitio activo (Simões and Faro, 2004).

El elevado número de genes putativos de PAs en plantas (Fig. 5) y la gran diversidad de estas proteasas expresadas tanto por expresión simple como por co-expresión múltiple en la misma planta o tejido, sugieren que existen roles diferentes para cada PA y que éstos podrían estar relacionados con la especificidad, eficiencia catalítica y localización (Mazorra-Manzano et al., 2010). Las PAs de plantas han sido implicadas en procesamiento y/o degradación en diferentes órganos, así como en la senescencia, respuesta al estrés, muerte celular programada y reproducción. La mayoría de las funciones sugeridas se deben a estudios de co-localización de las PAs con sustratos proteicos putativos, de su capacidad de degradar estos sustratos *in vitro* y/o de su expresión específica en ciertos tejidos o bajo ciertas condiciones (Mutlu and Gal, 1999; Simões and Faro, 2004). Sin embargo, sus funciones biológicas no están tan bien asignadas como en el caso de sus equivalentes de origen no vegetal y con certeza sólo se conocen los roles biológicos de dos proteasas vegetales de la familia A1, éstas son CDR1 (Constitutive Disease Resistance 1) identificada en *Arabidopsis* que actúa en la supresión de la enfermedad causada por el patógeno *Pseudomonas syringae* y PCS1 (Promotion of Cell Survive 1) encontrada en *Arabidopsis* y en nucela de cebada que participa en la prevención de la muerte celular durante la gametogénesis y embriogénesis (van der Hoorn, 2008).

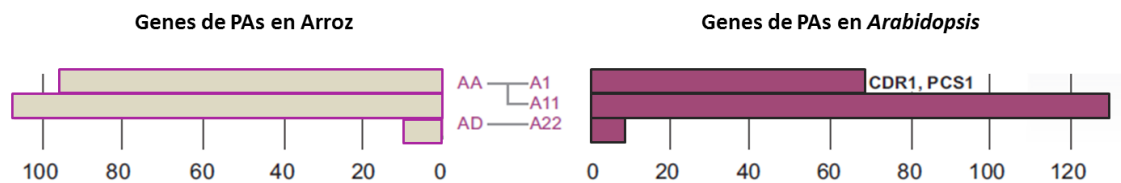


Fig. 5. Distribución de los genes de proteasas aspárticas en arroz (izquierda) y en Arabidopsis (derecha), sobre los diferentes clanes (AA, AD) y familias (A1, A11, A22). Las proteasas cuyos roles biológicos son conocidos a partir de experimentos genéticos, se indican a la derecha.

Como la mayoría de las PAs, las de origen vegetal son también sintetizadas como zimógenos que resultan posteriormente convertidos a enzimas maduras. Las PAs de plantas, a partir del secuenciamiento del genoma completo de *A. thaliana*, se clasificaron en tres categorías dependiendo de la organización putativa de sus dominios y las secuencias de los motivos de sus sitios catalíticos: PAs típicas, PAs de tipo nucelina y PAs atípicas (Faro and Gal, 2005).

La característica más sobresaliente de las PAs típicas de plantas es la presencia de un inserto de alrededor de 100 aminoácidos en el dominio C-terminal, conocido como inserto específico

de planta o PSI (indicado en verde en la Fig. 4 B). Este segmento es generalmente removido durante la maduración de la proteasa.

Las PAs de tipo nucelina codifican proteínas similares a la nucelina, detectada por primera vez en las células nucelares de la cebada y que por sus patrones de expresión así como por su localización específica, se les ha asignado un posible rol en la muerte celular en plantas.

Las PAs atípicas muestran características intermedias entre las secuencias de las PAs típicas y las de tipo nucelina (Faro and Gal, 2005).

Las PAs de la familia A1 poseen una región rica en cisteínas, cuya distribución es también una característica considerada en la clasificación de estas proteasas; mientras que en las PAs típicas, el dominio PSI contiene un “*cluster*” de seis residuos cisteínas, en las aspárticas similares a nucelina y las atípicas, en las cuales el inserto PSI está ausente, el *cluster* de cisteínas también existe en una región localizada entre el primer aspartato catalítico y el residuo Tyr75 (numeración de pepsina) conservado en la porción conocida como *flap* que cubre la hendidura del sitio activo.

En general, las PAs de origen vegetal comparten una alta similitud de secuencia aminoacídica en sus dominios N- y C-terminales (identidad superior al 60%) y alrededor de un 45% de identidad con catepsina D, la PA de origen no-vegetal más cercana. Otra característica distintiva de las PAs de plantas es el segundo motivo catalítico: Asp-Ser-Gly (DSG) presente en todas las PAs de plantas pertenecientes a la familia A1 con la excepción de chlapsina, proveniente de *Chlamydomonas reinhardtii* (accession number: AJ579366) que contiene DTG/DTG (Almeida et al., 2012) similarmente a la vasta mayoría de las PAs de clase A1 de origen no vegetal.

Estructura tridimensional

En relación a las PAs de plantas, sólo dos estructuras cristalinas han sido determinadas (Fig. 6), la estructura de cardosina A madura (código de PDB: 1B5F) (Frazão et al., 1999) y la de profitepsina, precursor de la PA de cebada conteniendo el prosegmento y el dominio PSI (código de PDB: 1QDM) (Kervinen et al., 1999).

Ambas proteasas son polipéptidos de dos cadenas en sus formas maduras y presentan un plegamiento muy similar a aquel encontrado para otras PAs de origen no-vegetal (ver Estructura tridimensional en sección Proteasas Aspárticas). Además de la topología tipo

pepsina común para el cuerpo principal de la profitepsina madura, la caracterización estructural del precursor de la enzima también arrojó nuevos puntos de vista sobre los prosegmentos y el PSI (Kervinen et al., 1999).

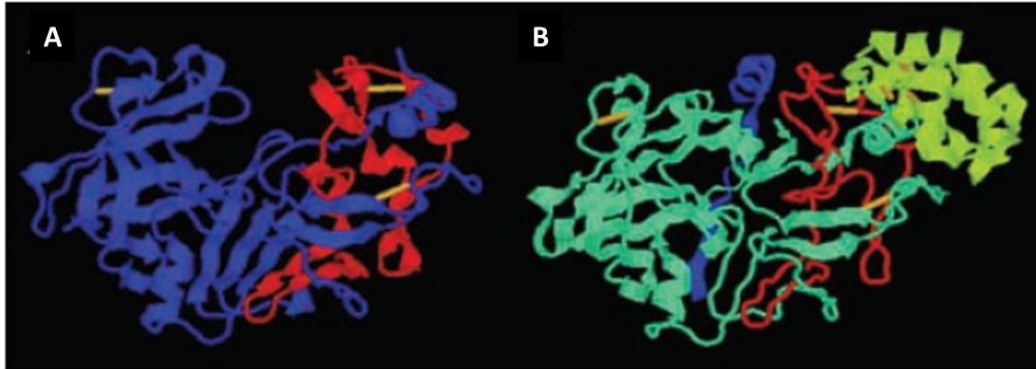


Fig. 6. Representación de cintas de la estructura cristalina de cardosina A y profitepsina. (A) Estructura de cardosina A madura (código de PDB: 1B5F), la cadena pesada se indica en azul y la liviana en rojo; en amarillo se representan los puentes disulfuros. (B) Estructura de profitepsina (código de PDB: 1QDM), en azul se muestra el propéptido del zimógeno, las cadenas correspondientes a la proteína madura se encuentran coloreadas en celeste (cadena pesada) y en rojo (cadena liviana), mientras que el inserto PSI está indicado en verde; en amarillo se representan los puentes disulfuro (Simões and Faro, 2004).

Inserto específico de planta

Las PAs típicas de plantas como se mencionó anteriormente, poseen un segmento interno, el PSI, de alrededor de 100 aminoácidos que es generalmente removido durante la maduración de la proteasa y no está presente en la enzima madura. El PSI comparte una significativa homología de secuencia con las saposinas (Guruprasad et al., 1994), una clase de proteínas involucradas en la degradación lisosomal de esfingolípidos (O'Brien and Kishimoto, 1991; Kishimoto et al., 1992). Esta homología se extiende también a otras proteínas de la familia de proteínas similares a saposina (SAPLIP), dentro de las que se incluyen saposinas A-D, NK-lisina, granulosina, proteína surfactante B, amoebaporos y dominios de esfingomielinasa ácida y aciloxiacil hidrolasa (Munford et al., 1995; Stenger et al., 1998; Vaccaro et al., 1999; Piatigorsky et al., 2001; Ciaffoni et al., 2006). Como otros miembros de esta familia, el PSI contiene seis cisteínas conservadas, varios residuos hidrofóbicos y frecuentemente, un sitio de glicosilación consenso. En el caso particular de la PA de *Chlamydomonas reinhardtii*, además de estas características comunes, el dominio PSI compromete una región adicional de aproximadamente 80 aminoácidos ricos en tripletes de alanina cuya función es aún

desconocida (Almeida et al., 2012). Una característica única del PSI es el intercambio de las porciones N- y C-terminales del dominio tipo saposina, donde la porción C-terminal de una saposina ocupa la posición N-terminal del PSI y está ligada a la porción N-terminal de otra saposina formando el C-terminal del PSI (Fig. 9). Por tanto, el PSI no es realmente una saposina sino una “swaposina” (Munford et al., 1995; Ponting and Russell, 1995; Kervinen et al., 1999).

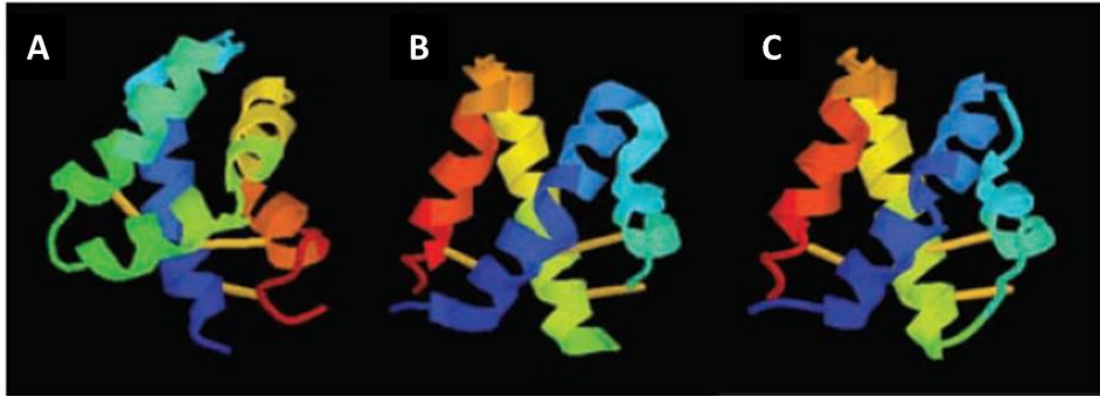


Fig. 7. Plegado tipo saposina. (A) Representación de la estructura de NK-lisina (código de PDB: 1NKL), una proteína similar a saposina (Liepinsh et al., 1997). (B) Dominio PSI de profitepsinas de cebada (código de PDB: 1QDM) (Kervinen et al., 1999). (C) Modelo estructural del dominio PSI de cardosina A, basado en la estructura cristalina de profitepsina. Los dominios N-terminales en las figuras están mostrados en azul y los C-terminales en rojo. Adaptado de Simões and Faro, 2004.

La estructura tridimensional de la NK-lisina fue determinada por RMN (Liepinsh et al., 1997) revelando por primera vez, que el “plegado tipo saposina” compromete cinco α -hélices anfipáticas unidas entre sí mediante tres puentes disulfuro y plegadas en un único dominio globular (Fig. 7).

El “plegado tipo saposina” fue posteriormente confirmado con la determinación de las estructuras tridimensionales del PSI de profitepsina (código de PDB: 1QDM) (Kervinen et al., 1999), granulosina (código de PDB: 1L9L) (Anderson et al., 2003), saposina C (código de PDB: 1M12) (de Alba et al., 2003), saposina B (código de PDB: 1N69) (Ahn et al., 2003), amoebaporo A (código de PDB: 1OF9) (Hecht et al., 2004) y PSI de la PA de papa StAP (código de PDB: 3RF1) (Bryksa et al., 2011).

Poco se sabe acerca de la función del dominio PSI. Se ha sugerido que el mismo está involucrado en el correcto plegamiento del precursor de cyprosina, aislada de *Cynara cardunculus*, y es requerido para el procesamiento y activación de la enzima (White et al.,

1999). De cualquier forma, la función más aceptada del PSI es el direccionamiento de los precursores de las PAs de plantas a la vacuola celular (Guruprasad et al., 1994; Kervinen et al., 1999; Tormakangas et al., 2001), donde las formas maduras de las enzimas son almacenadas. La estructura tridimensional de la profitepsina reveló un sitio putativo de unión a un receptor de direccionamiento comprometiendo al PSI y un área adyacente de la enzima madura (Kervinen et al., 1999), que podría posiblemente, interactuar con receptores de membrana en el Golgi, siendo el complejo posteriormente transportado a la vacuola. El mismo tipo de interacción fue propuesto para el direccionamiento independiente de manosa-6-fosfato de procatepsina D al lisosoma. En este caso, la procatepsina D forma un complejo transitorio con prosaposina y ambas proteínas son co-transportadas al lisosoma (Rijnboutt et al., 1991; Glickman and Kornfeld, 1993; Zhu and Conner, 1994). En la PA de arroz se ha demostrado que el PSI no tiene influencia en la actividad enzimática, aunque aparentemente participa en la velocidad de autólisis de esa enzima (Asakura et al., 2000). Por otro lado, se ha observado que el PSI aumenta la estabilidad y especificidad de sustrato al ser insertado en la pepsina porcina (enzima quimérica) en comparación con la pepsina nativa (Payie et al., 2003). Esta mayor especificidad de corte se observó también en PAs que retienen el PSI en la forma madura, como la proveniente de papa (StAP3) y la obtenida de gluten de trigo (G1AP2), y con una forma recombinante de la PA de *Arabidopsis thaliana* (rAtAP1), que conserva casi la totalidad de su PSI (Mazorra-Manzano et al., 2010). Se cree además que el PSI junto con el sitio de glicosilación de las PAs son importantes para la estabilidad de las mismas (Payie et al., 2003). Un indicio de esto último lo constituye la mayor estabilidad térmica reportada para la proteasa rAtAP A1 (Mazorra-Manzano et al., 2010) respecto a las estabilidades de las PAs de plantas sin el dominio PSI.

De forma similar a todas las proteínas SAPLIPs, el PSI parece ejercer su función a través de interacciones proteína-membrana (Egas et al., 2000; Mendieta et al., 2006; Bryksa et al., 2011). Los efectos sobre membranas caracterizados mediante microscopía de fuerza atómica y dispersión de luz, indicaron que el dominio PSI de la PA de papa, StAP-PSI, posee capacidad de desestabilizar bicapas lipídicas así como también actividad fusogénica. Más aún, se determinó que el mecanismo de dicha interacción con membranas sigue una cinética que se ajusta a la curva de Michaelis-Menten mostrando dependencia con el pH (Bryksa et al., 2011).

Debido a la capacidad de interferir con la permeabilidad de vesículas lipídicas artificiales encontrada para el PSI de cardosina A, se ha sugerido que el mismo está involucrado en mecanismos de defensa contra patógenos o tiene un rol en la autólisis celular (Egas et al., 2000). En concordancia con esta suposición, se han reportado actividades citotóxicas contra

fitopatógenos y patógenos humanos del dominio PSI de la PA de papa, StAP-PSI (Mendieta et al., 2006; Muñoz et al., 2010).

Mecanismos de procesamiento

Las PAs típicas son sintetizadas como preproenzimas de simple cadena que posteriormente se procesan a formas maduras de una o dos cadenas. El procesamiento proteolítico de los zimógenos de PAs es más complejo para las enzimas de origen vegetal dado que además del propéptido, un dominio extra, el PSI, tiene que ser removido. En general, el dominio PSI es escindido del precursor en forma parcial o total, no obstante, existen algunas variaciones en el mecanismo y el orden en que el prosegmento y este inserto son eliminados durante la conversión hacia la forma activa (Simões and Faro, 2004). Se ha sugerido que el procesamiento de las preproenzimas involucra una serie de clivajes proteolíticos secuenciales durante el tráfico hacia el Golgi, lo que da lugar a una mezcla compleja de formas heterogéneas de peptidasas aspárticas maduras dentro de los tejidos de la planta (Duarte et al., 2008).

Procardosina A sufre un procesamiento proteolítico durante la maduración de la flor, en el cual su PSI es removido completamente, probablemente por una PA, luego de lo cual el prosegmento es removido (Ramalho-Santos et al., 1998). Un mecanismo de procesamiento autocatalítico similar ha sido propuesto para cenprosina, PA de *Centaurea calcitrapa* (Domingos et al., 2000) y para oryzasina 1 recombinante, PA de arroz (Asakura et al., 2000).

Otro esquema de procesamiento fue descrito para profitepsina, donde el propéptido es escindido previamente a la remoción incompleta del dominio PSI (Glathe et al., 1998; Tormakangas et al., 2001). Las diferencias observadas entre los sitios de procesamiento *in vivo* e *in vitro* condujeron a sugerir la participación de otras proteinasas y/o exopeptidasas en el mecanismo de activación además del mecanismo de autoprosesamiento (Glathe et al., 1998).

Para la procyprosina producida en *Pichia pastoris* el procesamiento se asemeja al de profitepsina. Éste ocurre por remoción del propéptido y la escisión incompleta del PSI, conduciendo a una enzima madura con las cadenas pesada y liviana conectadas por un puente disulfuro (White et al., 1999). Sin embargo, el procesamiento de la procyprosina parece involucrar la acción de proteasas de la célula huésped y no una autoactivación, dado que ésta fue sólo marginalmente afectada por la presencia de pepstatina en el medio de cultivo de las células de *P. pastoris* (White et al., 1999). También ha sido propuesto un mecanismo de procesamiento similar para una de las PAs expresadas en las semillas de girasol (Park et al.,

2001). En cualquier caso, el procesamiento de las PAs de plantas origina una enzima heterodimérica que carece tanto del prosegmento como del PSI, con una organización de los dominios semejante a la de sus homólogos de origen no-vegetal de la familia A1 de PAs similares a pepsina.

Aunque no se ha identificado ninguna enzima particular como responsable del inicio de los eventos de procesamiento, evidencias obtenidas mediante experimentos *in vitro* sugieren que la maduración completa de las PAs en plantas podría requerir la acción de otra(s) proteasa(s) (White et al., 1999; Domingos et al., 2000; Castanheira et al., 2005; Sampaio et al., 2008), en adición al procesamiento autocatalítico. La diversidad entre las PAs típicas de plantas estaría, por lo tanto, causada en parte por la intervención de diferentes sitios de clivaje proteolítico sobre el mismo precursor, los que potencialmente dan lugar a diferentes isoformas de la proteasa. En este sentido, se ha encontrado microheterogeneidad en ambas subunidades de cardosinas A, E, F, G y H y en cyprosinas (Cordeiro et al., 1992, 1994). La remoción del PSI puede resultar en microheterogeneidad tanto en el extremo C-terminal de la cadena pesada como en el N-terminal de la cadena liviana. Aunque la diversidad y heterogeneidad no sólo se debe a modificaciones postraduccionales, sino también a la naturaleza multigénica de esta familia de enzimas (Pimentel et al., 2007, Sarmiento et al., 2009).

No obstante, aún cuando la diversidad de las PAs en plantas se encuentra bien documentada en la literatura, la correlación entre los patrones de diversidad y sus funciones biológicas está lejos de ser descifrada (Lufrano et al., 2012).

A pesar que el procesamiento de la mayoría de los zimógenos aspárticos de plantas remueven el PSI de manera completa o parcial resultando en proteasas heterodiméricas, un pequeño grupo de las PAs estudiadas conserva el inserto específico de planta aún en su forma madura monomérica. Tal es el caso de las PAs aisladas de tomate (Rodrigo et al., 1989), tabaco (Rodrigo et al., 1991) y papa (Guevara et al., 1999).

Dentro de esta familia de proteasas que ha sido poco estudiada en plantas respecto a otras familias y en relación a las contrapartes de origen no vegetal, el trabajo realizado durante esta tesis aporta un nuevo miembro a la lista de PAs típicas de plantas caracterizadas hasta el momento.

Objetivos



Objetivos

Las proteasas aspárticas típicas de plantas (o tipo pepsina) han sido escasamente estudiadas en comparación con sus contrapartes de origen no vegetal. Una excepción a esta afirmación, lo constituyen las cardosinas (*Cynara cardunculus*) usadas en la manufactura de quesos y la fitepsina (*Hordeum vulgare*), que al día de hoy continúa siendo una de las pocas proteasas de esta clase caracterizada estructuralmente y por tanto, el modelo utilizado en esta área.

Una de las principales razones por las cuales las PAs de plantas no han sido objeto de mayor cantidad de estudios es su escasa abundancia natural. Ésto resulta especialmente cierto en el caso de sus precursores, los cuales contienen el llamativo dominio PSI (inserto específico de plantas), único entre las PAs.

En un intento por identificar nuevas fuentes vegetales de PAs típicas, extractos de flores de varias especies pertenecientes a la familia Asteraceae que crecen en Argentina fueron probados en ensayos de coagulación de leche. Una de las especies que mostró esta capacidad coagulante sugiriendo por tanto la presencia de PAs, fue *Cirsium vulgare*—comúnmente conocido como cardo negro y considerada una maleza en los campos de nuestra región. Fundamentalmente debido a su amplia disponibilidad y a la ausencia de estudios reportados sobre sus PAs, se decidió emplear esta especie de la familia Asteraceae como material vegetal.

En este contexto, el presente trabajo intenta generar un aporte al conocimiento de las PAs de origen vegetal mediante el estudio de procirsina, una proenzima de *C. vulgare* perteneciente a la familia A1, y del dominio PSI contenido en su secuencia.

Para llevar adelante este propósito, se plantearon los siguientes objetivos generales y los objetivos específicos que contribuyen a lograr cada objetivo general:



Clonar y secuenciar los ADNc de PAs típicas de *Cirsium vulgare*

- a- Aislar el ARN total de flores frescas de *C. vulgare*
- b- Clonar los ADNc que codifican para los precursores de PAs típicas (preproenzimas)
- c- Obtener la secuencia nucleotídica y de aminoácidos de estos precursores

- d- Realizar un estudio comparativo con los precursores de otras proteasas aspárticas conocidas
- e- Generar un modelo tridimensional de cirsina



Producir y caracterizar enzimáticamente la procirsina recombinante

- a- Construir el vector de expresión para procirsina y sus mutantes (de sitio activo y sin el dominio PSI) para el sistema de elección (*Escherichia coli*)
- b- Expresar las diferentes formas de procirsina en el sistema heterólogo *E. coli*
- c- Purificar la proteína recombinante y sus mutantes
- d- Caracterizar las propiedades enzimáticas de las proteínas expresadas



Obtener el dominio PSI del precursor de cirsina en forma recombinante

- a- Construir el vector de expresión para el PSI y el mutante del sitio de glicosilación para el sistema escogido (*Kluyveromyces lactis*)
- b- Expresar ambas proteínas en el sistema *K. lactis* y analizar la influencia de la glicosilación en la secreción del PSI
- c- Estudiar las actividades antibacteriana y antifúngica del dominio PSI.

Material Vegetal



Material Vegetal

Cirsium vulgare (Savi) Tenore (Asteraceae)



Cirsium vulgare se conoce en Argentina con los nombres comunes de cardo negro (Argentina; Marzocca, 1976) y en otros lugares como bull thistle, spear thistle y common thistle.

Es una hierba bianual, erecta, de 1-2 m de alto. El tallo por lo general es ramificado en la parte superior, aracnoideo-lanoso, con alas dentado-espinosas.

Las hojas no poseen un pecíolo bien definido, las basales formando una roseta, hasta 30 cm de largas, las superiores son alternas, abrazan el tallo y profundamente pinatipartidas a bipinatipartidas con los segmentos triangulares, terminando en una espina de hasta 1 cm.

La inflorescencia está formada por cabezuelas solitarias o poco agrupadas en los extremos de las ramas. Las flores/cabezuelas son de 3-6 cm de diámetro, presentan involucro

campanulado a ovoide, sus brácteas se disponen en varias series, linear-subuladas, las más largas hasta 4 cm de largo, atenuadas en el ápice hacia una espina amarillenta de 2 a 2,5 mm de largo, aracnoideo-tomentosas y con frecuencia más o menos reflejas. El receptáculo es plano; las flores de 150 a 250 presentan corolas de 2,5 a 3,5 cm de largo, de color lila, los segmentos del limbo de 5 a 8 mm de largo, un poco más cortos a dos veces más largos que la garganta; anteras de 8 a 11 mm de largo.

La característica especial es que se trata de una planta muy espinosa (de Rzedowski, Graciela C. Rzedowski, 2001). Se la encuentra en climas templados, suelos fértiles y húmedos (Marzocca et al., 1976). Puede invadir pastizales naturales. Sus rosetas pueden cubrir superficies considerables, y desplazar vegetación nativa por lo que puede tener un efecto marcado sobre la biodiversidad y los ecosistemas y es considerada maleza invasiva.

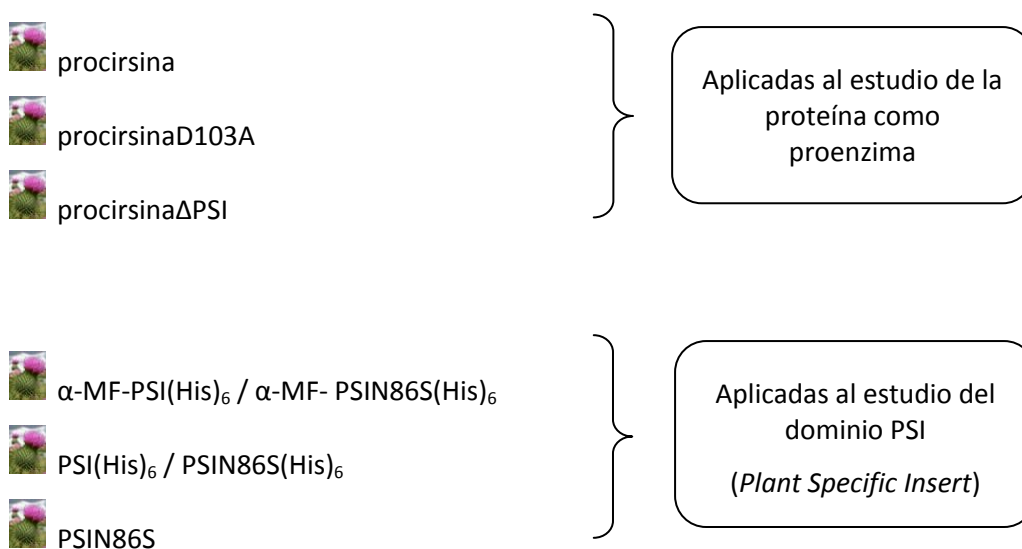
Una muestra recolectada en las inmediaciones de la ciudad de La Plata ha sido depositada en el herbario LPE de la Facultad de Ciencias Exactas de la Universidad Nacional de La Plata, Argentina con el voucher LPE 1160.

Materiales y Métodos Generales



Materiales y Métodos Generales

A fin de llevar a cabo los estudios que conforman los objetivos de la presente tesis doctoral, se obtuvieron las siguientes proteínas recombinantes como material de partida:



Todas estas proteínas derivan del precursor inicial de la enzima, al cual se denominó preprocirsina. Por lo cual, primeramente se clonó la secuencia codificante para la preprocirsina en el vector pGEM-T Easy y a partir de este plásmido (pGEM-preproCV) se generaron las construcciones utilizadas para la expresión de las proteínas recombinantes.

Estas proteínas pueden ser agrupadas en dos categorías de acuerdo al sistema utilizado para su expresión (procariota o eucariota) y la fracción desde la cual las mismas fueron recuperadas (fracción intracelular insoluble o fracción extracelular), como se muestra en la Fig. 8.



Fig. 8. Esquema de las etapas seguidas en la obtención de las proteínas expresadas en el desarrollo de este trabajo, clasificadas según el sistema de expresión y la fracción de la cual se recuperaron.

En esta sección se describe el clonado de los precursores de proteasas aspárticas de flores de *Cirsium vulgare* (preprocirsina) y en forma general, aquellas técnicas generales empleadas durante la realización de la presente Tesis Doctoral.

Clonado de preprocirsina



Extracción del ARN total

A partir de inflorescencias inmaduras recolectadas en las inmediaciones de la ciudad de La Plata durante las estaciones de primavera-verano, se aisló el ARN total empleando el kit comercial Invisorb Spin Plant RNA Mini Kit (Invitek), que hace uso de las propiedades de unión selectiva de las membranas a base de sílice. Aproximadamente 100 mg de pimpollos fueron triturados con nitrógeno líquido en un mortero y el polvo resultante se procesó siguiendo las instrucciones proporcionadas por el fabricante. El ARN eluido de las microcolumnas que contienen las membranas de sílice, se cuantificó por absorbancia a 260 nm usando un espectrofotómetro UV-visible de micro-volumen (Nanodrop 1000, Thermo) y su integridad se evaluó en gel de agarosa 1% (p/v) como se describe en la *Técnica nº1* del apartado de Técnicas Generales.



Síntesis de la doble cadena híbrida ARN-ADNc

Se sintetizó la primera cadena de ADN complementaria al ARNm por reacción de retrotranscripción mediante un kit comercial (1st Strand cDNA Synthesis Kit for RT-PCR (AMV), Roche) utilizando los cebadores oligo-p(dT)₁₅ y como molde, aproximadamente 1,2 µg del ARN total previamente obtenido y desnaturalizado a 65 °C durante 15 min seguido de una incubación de 5 min en hielo.

La reacción se llevó a cabo en un volumen final de 20 µl. La composición de la mezcla de reacción y los parámetros de termociclado empleados en la síntesis del híbrido ARN-ADNc, fueron los indicados en el kit arriba mencionado (catálogo nº: 11483188001).



Amplificación del ADNc que codifica para la secuencia completa de la preprocirsina

A fin de obtener el amplicón de aproximadamente 1500 pb que codifica para la preproenzima, de la cual se desconocía por completo su secuencia, se utilizaron cebadores específicamente diseñados para los extremos 5' y 3' altamente conservados de las secuencias codificantes para las proteasas aspárticas de plantas (Tabla 1).

Cebador	Secuencia (5´->3´)	Tm (°C) ^(*)
preproCV_F	ATGGGTACCTCAATCAAAGCA	54
preproCV_R	TCAAGCTGCTTCTGCAAATCC	54

TABLA 1. Secuencias de los cebadores utilizados en la amplificación de la secuencia entera de preprocirsina para la obtención de la construcción pGEM-preproCV.

^(*) Tm= 4(C+G)+2(A+T)

La reacción de amplificación se realizó como se indica en la *Técnica n°2*, con una alícuota del ADNc sintetizado como molde y 1 U de enzima Taq Polymerase (Amersham). Se incluyeron 35 pasos en el segmento II del programa de termociclado con 45 seg de desnaturalización, 45 seg de hibridación a 54 °C y 2 min de extensión a 72 °C.

El peso molecular esperado se verificó mediante electroforesis de agarosa 1% (p/v) según se indica en la *Técnica n°1*.



Amplificación del plásmido pGEM-preproCV en E.coli

El producto de amplificación obtenido mediante PCR se purificó a partir del gel de agarosa para su posterior ligación al vector de clonación pGEM-T Easy (Promega) como se indica en las *Técnicas n°3 y n°5*, dando origen al plásmido pGEM-preproCV.

Posteriormente, este plásmido se empleó para transformar una alícuota de células TOP10F[´]químicamente competentes preparadas de acuerdo a la *Técnica n°4*. El protocolo de transformación utilizado se describe en la *Técnica n°5*. Después de la propagación de las células transformadas en una placa con medio Luria-Bertani (LB) conteniendo ampicilina, se realizó la selección de clones positivos mediante análisis de restricción con la enzima *StyI* (New England Biolabs) según lo detallado en la *Técnica n°7*.

Finalmente, varios clones positivos se eligieron para aislar su ADN plasmídico por medio de kits comerciales (*Técnica n°6*) y determinar la secuencia de preprocirsina contenida en estos plásmidos de forma automática (Stab Vida, Portugal/Macrogen Korea, Amsterdam).

Las secuencias obtenidas se sometieron a análisis de alineamiento múltiple y a búsqueda de dominios conservados empleando las herramientas BLAST y CD-Search (Marchler-Bauer et al.,

2011) del sitio de internet del NCBI (National Center for Biotechnology Information, <http://www.ncbi.nlm.nih.gov>).

Técnicas Generales

A. Técnicas de biología molecular

Técnica nº1: Electroforesis de ácidos nucleicos (ADN y ARN) en gel de agarosa 1% (p/v)

Para el análisis de las muestras de ácidos nucleicos se utilizaron geles horizontales de agarosa 1% (p/v) en buffer TAE (Tris-HCl 40 mM, EDTA 1 mM, Ácido Acético Glacial 0,11%, pH 8,0) con agregado de bromuro de etidio (10 mg/ml) para la visualización de las bandas. Las muestras de ácidos nucleicos se aplicaron con un quinto de volumen de buffer de muestra (6x, Fermentas), usando marcadores de peso molecular de 1 kb (O'Gene Ruler 1kb DNA Ladder, ready-to-use 0,1 µg/µl, Fermentas). Las corridas electroforéticas se llevaron a cabo con buffer TAE a voltaje constante (90 volts) en un sistema Sub-Cell GT Cell o Mini-Sub Cell GT Cell (Bio-Rad). Para la detección por fluorescencia y el registro de los geles se utilizó un sistema de captura de imagen automático (Molecular Imager Gel Doc XR System, Bio-Rad).

32

Técnica nº2: Reacción en cadena de la polimerasa

La amplificación de regiones determinadas de ADNc se llevó a cabo mediante la técnica de reacción en cadena de la polimerasa (PCR) bajo condiciones estándar (Mullis, 1990) en un termociclador My Cycler thermalcycler (BioRad). La mezcla de reacción que se utilizó se describe en la Tabla 2, mientras que los parámetros utilizados en el programa de termociclado se detallan a continuación en la Tabla 3.

En las reacciones de mutagénesis que se llevaron a cabo usando el kit comercial Quick Change Site-Directed Mutagenesis Kit (Stratagene), el programa de termociclado en todos los casos se adaptó a las indicaciones dadas por el fabricante (catálogo nº: 200518).

Componente	Volumen (μ l)
Buffer de reacción (10X)	5,0
Mezcla dNTPs (10mM)	2,0
Cebador F (100 μ M)	0,5
Cebador R (100 μ M)	0,5
ADN Polimerasa	csp 0,5-2,5 U ⁽¹⁾
Agua MilliQ	csp 50 μ l
ADN molde	csp 10-60 ng

Tabla 2. Composición de la mezcla de la reacción en cadena de la polimerasa (PCR).

⁽¹⁾ Dependiendo de la polimerasa empleada.

Segmento	Paso	Nº de repeticiones	Temperatura (°C)	Tiempo
I	Desnaturalización inicial	1	95	5 min
	Desnaturalización (D)		94	t _D
II	Hibridación (H)	30-35	T _H	t _H
	Extensión (E)		68/72 ⁽¹⁾	t _E
III	Extensión final	1	68/72 ⁽¹⁾	10 min
IV	Mantenimiento	1	4	∞

Tabla 3. Programa de termociclado para la amplificación de secuencias de ADN por PCR.

⁽¹⁾ Dependiendo de la polimerasa usada en la reacción.

Técnica n°3: Purificación de ADN desde un gel de agarosa

Luego de desarrollada la electroforesis, la banda de ADN del peso molecular que se pretendía aislar se cortó del gel de agarosa, se colocó en un tubo de 1,5 ml y se purificó mediante el uso de kits comerciales (Invisorb Spin DNA Extraction Kit, Invitex/ High Pure PCR Product Purification Kit, Roche) basados en la unión selectiva de ADN de más de 100 pb a columnas rellenas de lana de fibra de vidrio. Se siguió el procedimiento indicado por el fabricante, estimándose la concentración del ADN eluido por absorbancia a 260 nm.

Técnica n°4: Preparación de células bacterianas químicamente competentes

En el primer día del procedimiento se traspasaron las bacterias a hacer competentes con la ayuda de un palillo estéril desde un stock de células conservado a -80 °C a 3 ml de medio LB.

Las bacterias se incubaron durante toda la noche a 37 °C bajo agitación continua; a la mañana siguiente 1ml del crecimiento se diluyó a 100 ml de medio fresco y se volvió a incubar hasta que el inóculo alcanzó un valor de DO_{600nm} entre 0,3 y 0,4. Alcanzado ese valor, las células se lavaron primero en 5 ml de $CaCl_2$ 50 mM y luego se resuspendieron en 1 ml de esta solución, adicionándose por último glicerol esterilizado en una concentración final de 37% (v/v). Las bacterias se conservaron a -80 °C en alícuotas de 100 μ l. Las diferentes células bacterianas empleadas en el desarrollo de la presente tesis se indican en la Tabla 4.

Cepa	Genotipo Relevante	Aplicación
TOP10F' (Invitrogen)	$F'\{lacI^q Tn10 (Tet^R)\} mcrA \Delta(mrr-hsdRMS-mcrBC) \Phi80lacZ\Delta M15 \Delta lacX74 recA1 araD139 \Delta(ara-leu)7697 galU galK rpsL endA1 nupG$	Clonado de plásmidos (amplificación)
XL1Blue (Stratagene)	$recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacI^q Z\Delta M15 Tn10 (Tet^r)]$	
BL21Star (Invitrogen)	$F- ompT hsdS_B (rB-mB-) gal dcm rne131 (DE3)$	Expresión de proteínas

Tabla 4. Cepas de *E. coli* utilizadas en el clonado y expresión de proteínas.

Técnica n°5: Clonado de un inserto en el sistema pGEM-T Easy

Cada una de las secuencias de ADNc que codifican las proteínas a expresar fue primero amplificada en la cepa TOP10F' de *E.coli* mediante clonación en el sistema pGEM-T Easy (Promega). El vector de clonado pGEM-T Easy se encuentra linealizado, con una timina 3' terminal en ambos extremos del sitio de inserción previniendo la recircularización del vector y generando una unión compatible con cualquier producto de PCR obtenido por empleo de polimerasa Taq. Esta característica da versatilidad a la vez que aumenta la eficiencia de ligación de los productos de PCR al vector.

El proceso de clonado en este sistema se llevó a cabo siguiendo las indicaciones provistas por el fabricante (Part# TM042). La ligación del producto de PCR purificado (inserto) al vector se realizó por incubación a 4 °C durante toda la noche. Posteriormente la mezcla resultante se introdujo en células TOP10F' o XL1-Blue, cuya competencia química se produjo en el laboratorio de la manera descrita en la *Técnica n°4*. El protocolo empleado para llevar a cabo esta transformación se encuentra detallado en el manual del sistema pGEM-T Easy. En forma

resumida, la mezcla de ligación se adicionó a una alícuota de células competentes y luego de ser incubada en hielo por 30 min, la misma se sometió a un shock térmico a 42 °C durante 45 segundos, seguido de una breve incubación en hielo. Las células se llevaron a 37 °C en medio LB por el lapso de 1,5 hs con agitación continua y transcurrido este tiempo se hicieron crecer *overnight* a 37 °C sobre placa de agar LB conteniendo IPTG (0.5 mM), X-Gal (80 µg/ml) y Ampicilina (100 µg/ml).

Debido a que el sitio de inserción en el vector pGEM-T Easy interrumpe el marco de lectura abierto de la enzima β-galactosidasa, los clones que incorporaron el vector conteniendo el inserto a clonar, pueden ser identificados como colonias blancas resistentes a la ampicilina en la placa de crecimiento. Cuando el gen de la β-galactosidasa no es interrumpido, las colonias resultan de color azul.

En forma complementaria, la presencia del inserto completo que se pretende clonar se confirmó en 8 a 12 colonias blancas de la placa de transformantes a través de un análisis de restricción (ver *Técnica nº7*).

Finalmente, el ADN del uno de los clones que resultaron positivos en el análisis de restricción se sometió a secuenciación automática (Stab Vida, Portugal/Macrogen Korea, Amsterdam) para el posterior subclonado del inserto en el vector de expresión elegido según *Técnica nº8*.

Técnica nº6: Aislamiento del ADN plasmídico (minipreparaciones)

El aislamiento de ADN plasmídico en forma de minipreparaciones con elevado grado de pureza y calidad para los procedimientos de biología molecular, se realizó mediante el uso del kit comercial High Pure Plasmid Isolation Kit (Roche). Este kit combina la liberación del ADN plasmídico por el método de lisis alcalina de las bacterias (Sambrook et al., 1989), con la unión selectiva del mismo a columnas de fibra de vidrio en las condiciones usadas en el procedimiento.

Para la obtención de las minipreparaciones se partió de un cultivo bacteriano proveniente de una única colonia de transformantes, de volumen y DO_{600nm} adecuados a las indicaciones del fabricante. Las células se lisaron empleando los buffers del kit y el ADN se extrajo y purificó en las columnas de fibra de vidrio siguiendo las instrucciones proporcionadas. El ADN eluido fue posteriormente cuantificado por absorbancia a 260 nm en un equipo Nanodrop 1000 (Thermo).

En aquellas aplicaciones en que la pureza de la minipreparación no constituyó un factor relevante, se empleó el método de lisis alcalina sin modificaciones (Sambrook et al., 1989), que permite obtener un mayor rendimiento en relación al uso de kits comerciales.

Técnica n°7: Digestión con enzimas de restricción

Dado que las enzimas de restricción reconocen secuencias nucleotídicas específicas, se utilizó la digestión con estas enzimas con el fin de generar extremos cohesivos o romos sobre moléculas de ADN que posteriormente se ligaron entre sí.

También se sometió a digestión por enzimas de restricción a aquellos plásmidos sobre los que se pretendía deducir la presencia/ausencia de una determinada secuencia mediante la escisión del fragmento de interés o por generación del perfil de bandas electroforéticas resultante (análisis de restricción).

En forma general, la digestión de plásmidos aislados como minipreparaciones de ADN (*Técnica n°6*) con una o más enzimas de restricción fue llevada a cabo durante 3-4hs a 37 °C en la mezcla que se detalla en la Tabla 5.

Componente	Volumen (μl)
Buffer de reacción compatible con la/s enzima/s usada/s, NEB (10x) ^(*)	2,0
Albúmina Sérica Bovina (BSA, 10 mg/ml) ^(*)	0,2
Ribonucleasa	2,0
Enzima de restricción (NEB)	variable ^(**)
ADN plasmídico	variable ^(**)
Agua MilliQ	csp 20

Tabla 5. Composición de la mezcla utilizada en la digestión de plásmidos por enzimas de restricción.

^(*) El buffer apropiado así como el requerimiento de BSA dependen de la/s enzima/s de restricción usada/s; ^(**) La cantidad de ADN a digerir y en consecuencia la cantidad de enzima usada, se adecuaron a los requerimientos de las aplicaciones posteriores.

Una vez transcurrido el tiempo estimado de reacción, el perfil de bandas generado así como el grado de digestión del ADN fueron analizados en un gel de agarosa 1% (p/v) según *Técnica n°1*.

Técnica n°8: Subclonado de un inserto en los vectores de expresión pET-23d y pKLAC1

El ADN a subclonar (inserto) y el vector escogido fueron digeridos con las mismas enzimas de restricción (*Técnica n°7*), purificados como se indica en la *Técnica n°3* y colocados en un tubo de 1,5 ml en relación molar vector:inserto de 1:3. La ligación se llevó a cabo empleando la enzima T4 DNA Ligase (New England Biolabs) y la composición del medio de reacción especificada por el fabricante de la misma (catálogo M0202). La mezcla de ligación fue incubada a 16 °C durante toda la noche y luego conservada a -20 °C hasta ser utilizada en la transformación de células químicamente competentes TOP10F' o XL1-Blue. La competencia química de las bacterias fue producida según lo descrito en la *Técnica n°4* y la transformación realizada de la misma manera que en el paso precedente de clonado en el sistema pGEM-T Easy, detallado en la *Técnica n°5*.

Dado que a diferencia del vector de clonación pGEM-T Easy, ninguno de los dos vectores de expresión utilizados presenta identificación de transformantes por discriminación de colonias blancas/azules sobre placas indicadoras, la selección de positivos se realizó directamente por análisis con enzimas de restricción de minipreparaciones de ADN (ver *Técnica n°7*) aisladas de 6 a 12 colonias escogidas al azar de la placa de transformantes.

Una vez más, la secuencia nucleotídica del plásmido uno de los clones que resultaron positivos en el análisis de restricción se confirmó mediante secuenciación automática (Stab Vida, Portugal/Macrogen Korea, Amsterdam) para su empleo en la expresión de la proteína codificada por la construcción generada.

B. Técnicas de purificación y análisis de proteínas

Técnica n°9: Cromatografía de exclusión molecular

Esta técnica se utilizó como un primer paso de purificación de las proteínas recombinantes obtenidas a fin de separar sus formas agregadas y péptidos de menor masa molecular generados por expresión parcial o por degradación de las mismas. Se empleó una columna HiLoad 26/60 Superdex S200 prep grade (GE Healthcare Bio-Sciences AB), cuyo gel está compuesto por una matriz de agarosa de alto entrecruzamiento unida covalentemente a

cadena de dextranos responsables de las propiedades de separación de este medio sobre proteínas globulares en un rango de 10 a 600 kDa.

La columna se conectó a un sistema de media presión BioLogic DuoFlow (BioRad) o Äkta FPLC (GE Healthcare Life Sciences) y se pre-equilibró con 2,5 volúmenes de columna (VC) del buffer de corrida empleado en cada caso. Se aplicaron 10 ml de muestra, previamente filtrada a través de una membrana de nitrocelulosa de 0,22 µm a la columna, a una velocidad de flujo de 2 ml/min. La elución se realizó empleando 2 VC del mismo buffer de preequilibrio a la misma velocidad de flujo. Se registró la absorbancia del eluato a 280 nm colectando en forma automática fracciones de 10 ml.

Técnica nº10: Cromatografía de intercambio aniónico

Esta técnica cromatográfica que permite separar los polipéptidos según sus propiedades de carga superficial se empleó para la purificación de wt-procirina y sus mutantes. Se utilizó una columna Mono Q 5/50 GL (GE Healthcare Bio-Sciences AB) pre-empaquetada con MonoBeads. Las pequeñas partículas perfectamente esféricas de poliestireno/divinil benceno hacen que este relleno tenga alta resolución, a esta ventaja se suma la elevada selectividad dada por el grupo intercambiador de aniones.

La columna se adosó a un sistema FPLC y la cromatografía se desarrolló a una velocidad de flujo constante de 0,75 ml/min en las siguientes etapas:

- i. equilibrio de la columna con 5 VC del buffer de equilibrio Tris-HCl 20 mM de pH 8,0
- ii. siembra de la muestra
- iii. lavado con 5 VC del buffer de equilibrio para eliminar las proteínas no retenidas
- iv. elución de las proteínas mediante un gradiente lineal (0,0 - 0,5 M) de NaCl en el buffer de equilibrio manteniendo en forma manual los porcentajes de NaCl en los que se detectó aumento de la absorbancia a 280 nm.

Técnica nº11: Cromatografía de afinidad

La cromatografía de afinidad se basa en la interacción reversible entre la proteína que se pretende purificar y un ligando unido a la matriz cromatográfica. En la presente tesis, se aplicó este principio a la purificación de las proteínas recombinantes marcadas con 6 residuos de

histidinas, las cuales forman complejos con iones metálicos de transición a través de sus anillos imidazólicos. Por tanto, se empleó para esta separación una columna His Trap HP (GE Healthcare) pre-empaquetada con Ni Sepharose High Performance conectada a un sistema FPLC.

La denominada cromatografía de afinidad por iones metálicos inmovilizados (IMAC), se llevó a cabo mediante las siguientes etapas:

- i) equilibrio de la columna con 5 VC del buffer de unión al ligando (fosfato de sodio 25 mM de pH 7,5,conteniendo imidazol 10 mM y NaCl 500 mM) a velocidad de flujo de 1 ml/min
- ii) aplicación de la muestra a velocidad de flujo de 0,5 ml/min
- iii) lavado de la columna con 10 VC del buffer de unión a 1 ml/min fin de barrer las sustancias no ligadas y proteínas unidas de manera inespecífica al soporte de la resina
- iv) elución de la proteína deseada por competición con imidazol mediante un gradiente escalonado de concentración creciente (50 mM; 100 mM y 500 mM) en el buffer de unión a flujo de 1 ml/min
- v) lavado y regeneración de la columna según las indicaciones del fabricante.

Técnica nº12: Cromatografía líquida de alta performance (HPLC)

Para llevar a cabo la separación de péptidos con alta resolución se utilizó una columna C18 (KROMASIL 100 C18 250 4,6 mm) conectada a un sistema Prominence (Shimadzu Corporation, Tokyo, Japan) y pre-equilibrada con TFA 0,1% (v/v) (Fluka). Se aplicaron 100 a 300 µl de muestra y mediante un gradiente lineal de hidrofobicidad (0-80%) con acetonitrilo (Sigma-Aldrich) en TFA 0,1% (v/v) se realizó la elución a 1 ml/min. La absorbancia del eluato se monitorió en forma continua a 220 nm.

Técnica nº13: Cuantificación de proteínas

La cantidad de proteínas totales presente en una muestra se determinó usando el kit comercial de Coomassie Protein Assay Reagent (Thermo Scientific) basado en el método colorimétrico de Bradford (1976). En este método las proteínas se unen al colorante Coomassie G-250 provocando un corrimiento del espectro de la forma rojiza/marrón del colorante (absorbancia máxima a 465 nm) hacia la forma azul del mismo (absorbancia máxima

a 610 nm), siendo máxima la diferencia entre ambas formas del colorante a 595 nm, longitud de onda a la cual se hace la lectura.

Los ensayos de cuantificación se llevaron a cabo en microplacas siguiendo las instrucciones del fabricante. La absorbancia de las muestras a 595 nm se determinó en un espectrofotómetro modelo Victor³V (Perkin Elmer). Luego, la cantidad de proteína se calculó utilizando una recta de calibración construida con concentraciones crecientes de albúmina sérica bovina (0; 1; 5; 10; 15; 20; 25 µg/ml) el mismo día de la determinación.

Técnica n°14: Electroforesis desnaturalizante (SDS-PAGE) discontinua

Preparación de los geles

En la electroforesis discontinua se arman dos tramos de gel de composición ligeramente diferente que se polimerizan en forma secuencial: *gel de resolución* en la porción inferior (Tabla 5) y *gel de stacking o apilamiento* (Tabla 6), donde se aplican las muestras, en la porción superior. Para el armado de los geles y el desarrollo de la electroforesis se usó un equipo Mini-Protean III (Bio-Rad).

Componente	Volumen
Agua	2,4 ml
Buffer Tris-HCl 1,5 M, pH 8,8	2,3 ml
Solución Acrilamida/Bisacrilamida 40% (37,5:1)	2,2 ml
Solución SDS 20%	70 µl
Persulfato de amonio 10%	70 µl
TEMED	7,0 µl

Tabla 5. Composición del Gel de Resolución (12,5 %).

Componente	Volumen
Agua	1,9 ml
Buffer Tris-HCl 0,625 M, pH 6,5	750 µl
Solución Acrilamida/Bisacrilamida 40% (37,5:1)	300 µl
Solución SDS 20%	30 µl
Persulfato de amonio 10%	30 µl
TEMED	3,0 µl

Tabla 6. Composición del Gel de Apilamiento (4 %).

Preparación y aplicación de las muestras

Las muestras a analizar por SDS-PAGE se trataron previamente a 90 °C durante 10 min en solución desnaturante conteniendo β -mercaptoetanol o DTT (Tabla 7). El volumen aplicado en el gel varió entre 20 μ l y 65 μ l según la concentración de proteínas en las muestras.

Para la electroforesis en condiciones no reductoras se emplearon las mismas soluciones desnaturantes sin el agregado de los agentes reductores, β -mercaptoetanol y DTT respectivamente.

Como estándar se utilizaron 3 μ l de la mezcla de marcadores de amplio rango de masa molecular Precision Plus Protein Unstained Standards (BioRad) cuando el gel fue teñido con Coomassie Brilliant Blue R-250 o con nitrato de plata y Precision Plus Protein All Blue Standards (BioRad), en los casos en que las proteínas del gel se transfirieron a membrana de PVDF (*Técnica nº17*).

La electroforesis se llevó a cabo a voltaje constante de 150 V en el buffer de corrida (Bicina 0,1 M, Tris Base 0,1 M, SDS 0,1%) durante 45 min.

Solución desnaturante (2 \times)		Solución desnaturante (6 \times)	
Componente	Cantidad	Componente	Cantidad
Urea	24 g	Tris-HCl 0,5 M, 0,4 % SDS, pH 6,8	7 ml
Tris-HCl 1 M, pH 7,4	12,5 ml	Glicerol	3,0 ml
Solución SDS 20%	10 ml	SDS	1 g
Azul de Bromofenol	2 mg	DTT	0,93 g
Agua	c.s.p 50 ml	Agua	c.s.p 10 ml

Tabla 7. Composición de las soluciones desnaturantes (2 \times) y (6 \times) empleadas en la preparación de las muestras analizadas por SDS-PAGE.

Técnica nº15: Coloración de geles de acrilamida con Coomassie Brilliant Blue R-250

Finalizada la corrida electroforética, el gel se sumergió en solución colorante con agitación orbital suave hasta coloración azul oscuro. Luego se transfirió a la solución decolorante realizándose sucesivos lavados hasta que el fondo del gel se destiñó completamente. Las composiciones de las soluciones empleadas se detallan en la Tabla 8.

Solución Colorante		Solución Decolorante	
Componente	Cantidad	Componente	Cantidad
Coomassie Brilliant Blue R-250	2 g	Ácido acético glacial	50 ml
Ácido acético glacial	100 ml	Metanol	250 ml
Metanol	500 ml	Agua	700 ml
Agua	400 ml		

Tabla 8. Composición de las soluciones empleadas para la tinción de proteínas con Coomassie Blue en geles de poliacrilamida.

Técnica n°16: Tinción de proteínas con nitrato de plata

La técnica usada para teñir con plata proteínas separadas en un gel SDS-PAGE está basada en el procedimiento de O'Connell y Stults (O'Connell and Stults, 1997) y se encuentra resumida en la siguiente tabla:

Pasos	Solución utilizada	Tiempo
Fijación	Etanol 30%, ácido acético 10% (v/v) en agua	3 veces de 30 min
Lavado I	Etanol 20 % (v/v) en agua	10 min
Lavado II	Agua destilada	10 min
Sensibilización	Tiosulfato de sodio (0,2 g/l)	1 min
Lavado III	Agua destilada	2 veces de 20 seg
Coloración	Nitrato de plata (2,0 g/l)	30 min*
Lavado IV	Agua destilada	5-10 seg
Revelado	Formaldehído (0,7 ml/l), Carbonato de sodio anhidro (30 g/l), Tiosulfato de sodio (10 mg/l)	hasta intensidad deseada
Finalización	Tris base (50 g/l), ácido acético 2,5% (v/v)	1 min

Tabla 9. Composición de las soluciones empleadas para la tinción con plata de proteínas en geles de poliacrilamida.

(*) Durante este paso el gel debe estar protegido de la luz bajo agitación orbital suave.

Técnica n°17: Transferencia de proteínas a membranas de PVDF

La transferencia de proteínas separadas en un gel de poliacrilamida se realizó empleando membranas de difluoruro de polivinilideno, PVDF Western Blotting Membranes (Roche). Estas

membranas presentan ventajas frente a las de nitrocelulosa, ya que retienen mayor cantidad de proteínas, son compatibles con la mayoría de los sistemas de detección, mecánicamente son más estables y poseen una gran resistencia química a los reactivos utilizados para la caracterización estructural de las proteínas.

De forma breve, finalizada la electroforesis desnaturalizante (ver *Técnica nº14*), el gel conteniendo las proteínas y la membrana de PVDF previamente activada en metanol, se equilibraron en el buffer de transferencia (Tris base 25 mM, Glicina 192 mM, 10-20% (v/v) Metanol, pH 8,3) durante 5 minutos. Luego se procedió al armado del casete de la manera esquematizada en la Fig. 9 y al desarrollo de la transferencia en el buffer mencionado por aplicación de un voltaje constante de 40 V durante toda la noche en cámara fría. Para el armado del sandwich y la transferencia se utilizó el sistema de electrotransferencia Trans-Blot cell (Bio-Rad).

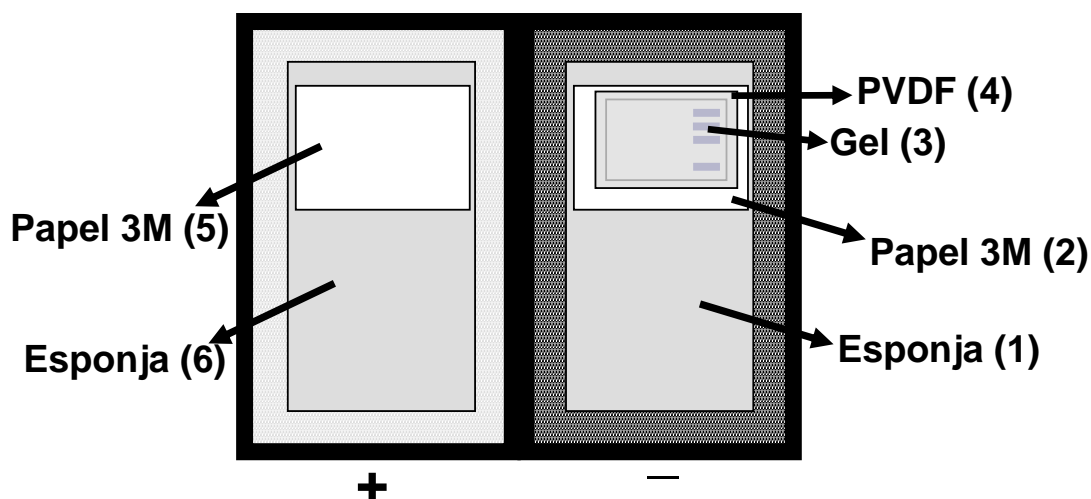


Fig. 9. Armado del casete para transferencia de proteínas a membrana de PVDF.

Una vez transferidas las proteínas, la membrana fue sometida a inmunodetección con anticuerpos específicos según la *Técnica nº18*.

Técnica nº18: Inmunodetección o Western blot

A fin de identificar las proteínas recombinantes expresadas con una cola de polihistidinas en sus extremos C-terminales y transferidas a membranas de PVDF (ver *Técnica nº17*), se utilizaron anticuerpo comercial anti-His obtenidos en ratón (GenScript). La detección

inmunoquímica se realizó empleando anticuerpo secundario de cabra anti-ratón (IgG+IgM) conjugado a fosfatasa alcalina y el reactivo revelante ECF Substrate for Western Blotting (GE Healthcare).

En forma resumida, el procedimiento utilizado se describe a continuación:

La membrana sobre la que se transfirieron las proteínas se bloqueó durante 1 h con una solución de leche descremada 5% (p/v) en buffer TBS-T (Tris 20 mM de pH 7,6 conteniendo NaCl 137 mM y 0,1% (v/v) de Tween-20). Transcurrido el tiempo de bloqueo, la membrana se incubó 1 h con una dilución adecuada del anticuerpo primario preparada en TBS-T con 0,5% (p/v) de leche magra. Se realizaron 5 lavados con la solución en que se diluyó el anticuerpo y a continuación se incubó la membrana durante 1 h con el anticuerpo secundario a la concentración aconsejada por el fabricante, en solución de leche descremada 0,5% (p/v) en buffer TBS-T. Se repitieron los lavados con buffer TBS-T.

Finalmente, la membrana fue revelada con el reactivo ECF siguiendo las instrucciones de fábrica. La señal quimiofluorescente, se adquirió utilizando un escaner Molecular Imager FX (BioRad).

Técnica nº19: Digestión trípica en geles de poliacrilamida

Para obtener la huella peptídica (Peptide Mass Fingerprint, PMF) por espectrometría de masa MALDI-TOF (PMF MALDI-TOF MS) de una proteína se hizo una digestión trípica *in situ* de las proteínas separadas mediante SDS-PAGE siguiendo el protocolo que se describe a continuación:

Las muestras conteniendo los polipéptidos a ser digeridos se sometieron a SDS-PAGE siguiendo la *Técnica nº18* y se colorearon con una solución de Coomassie coloidal comercial, Instant Blue (Expedion), durante 2 hs con agitación continua. El exceso de colorante se retiró mediante lavados sucesivos con agua de calidad MilliQ. Las bandas seleccionadas se recortaron con bisturí y almacenaron en tubos eppendorf a -20 °C hasta ser procesadas. Se realizó la reducción y carbamidometilación de las muestras con DTT 10 mM y con iodoacetamida 50 mM; seguidamente se digirieron con 0,3 µg de tripsina a 37 °C durante toda la noche. Los péptidos se disolvieron en TFA 0,1% (v/v) y analizaron por MALDI-TOF MS utilizando un espectrómetro de masas Ultraflex MALDI-TOF/TOF (Bruker Daltonics) equipado con un láser de nitrógeno de 337 nm en el Servicio de Proteómica y Bioinformática (SepBio) del Instituto de Biotecnología y de Biomedicina (IBB) de la Universidad Autónoma de Barcelona. Se utilizó

como matriz HCCA (α -ciano-4-hidroxicinámico) y los siguientes calibrantes peptídicos para la calibración externa: bradiquinina [1–7 (757,39916)], angiotensina II (1.046,5418), angiotensina (1.296,6848), sustancia P (1.347,7354), bombesina (1.619,8223), sustrato de renina (1.758,93261), hormona adrenocorticotropina ACTH clip [1–17 (2,093.0862)], hormona adrenocorticotropina ACTH clip [18–39 (2.465,1983)] y somatostatina 28 (3.147,4710). Los resultados se analizaron mediante el software SequenceEditor 3.2 (Bruker Daltonics, Biotools 3.2) por comparación de la digestión triptica teórica de las secuencias deducidas a partir de los ADNc con los espectros de masas obtenidos en forma experimental a fin de confirmar la identidad de la proteína analizada.



Procirsina



Introducción—Procirsina

Desde tiempos remotos, extractos acuosos de flores de la tribu Cardueae (familia *Asteraceae*), como ser *Cynara cardunculus*, *Cynara humilis* y *Cynara scolymus* han sido utilizados en la obtención artesanal de quesos a partir de leche ovina y/o caprina en la Península Ibérica (Roseiro et al., 2003; Reis and Malcata, 2011). Hace ya unas décadas, se ha responsabilizado a las PAs presentes en estos extractos por la capacidad coagulante de leche observada en ellos (Heimgartner et al., 1990; Cordeiro et al., 1994). Inicialmente, fue esta propiedad la que impulsó el estudio de proteasas aspárticas aisladas de varias flores de la familia *Asteraceae*, con el objetivo principal de encontrar potenciales cuajos alternativos.

Sin embargo, a medida que se fue avanzando en el estudio de las PAs de plantas, otras características propias de este grupo comenzaron a tomar relevancia. Un aspecto particularmente saliente resultó ser la presencia del dominio PSI en el precursor, su implicancia y la subsecuente complejidad generada en el procesamiento hacia formas maduras de estas proteasas.

A la fecha, las PAs provenientes de flores de cardos estudiadas son: cardosinas, principalmente A, B, E, F, G y H de *C. Cardunculus* (Faro et al., 1995; Veríssimo et al., 1996; Sarmiento et al., 2009); cyprosinas 1, 2 y 3, también de *C. Cardunculus* (Cordeiro et al., 1994; White et al., 1999); cenprosinas de *Centaurea calcitrapa* (Domingos et al., 2000); proteasa aspártica de *C. humilis* similar a cardosina A (Esteves, 1995); cynarasas A, B y C (Llorente et al., 2004; Sidrach et al., 2006) de *C. scolymus*; silpepsina de *Silybum marianum* (Vairo Cavalli, 2005); PAs presentes en *Onopordum turcicum* y *acanthium* (Tamer, 1993; Brutti et al., 2012), en *Arctium minus* (Cimino et al., 2010), *Carduus acanthoides* y *Carduus thoermeri* (Lufrano et al., 2007).

En este capítulo se presenta el clonado del precursor de una proteasa aspártica de flores de *Cirsium vulgare*, a la que se denominó 'cirsina', y el estudio comparativo de los elementos encontrados en la secuencia aminoacídica deducida a partir de su ADNc, con las secuencias de PAs típicas de plantas más ampliamente estudiadas hasta la fecha.

Por otro lado, se caracteriza la actividad mostrada por la forma zimogénica de cirsina obtenida en el sistema de expresión *E. coli* y se estudia su procesamiento *in vitro*. A través de un análisis

filogenético y un modelado *in silico* de procirsina, se busca proporcionar información que permita hipotetizar sobre las causas que expliquen el hallazgo de un zimógeno activo, aún cuando el mismo presenta en su secuencia los residuos Lys11/Tyr13, previamente descritos como claves en la inhibición de la actividad de los precursores de las PAs típicas de plantas.

Materiales y Métodos—Procirsina

Obtención de procirsina y sus mutantes en *Escherichia coli*

Escherichia coli es una bacteria que puede crecer rápidamente en un medio mínimo que contenga una fuente de carbono y sales que proporcionen nitrógeno, fósforo y algunas trazas de metales. Su genética molecular, fisiología y sistema de expresión han sido ampliamente caracterizados. Respecto a otros métodos de expresión, *E. coli* es un sistema de bajo costo y fácil de obtener en cultivos con una alta densidad celular.

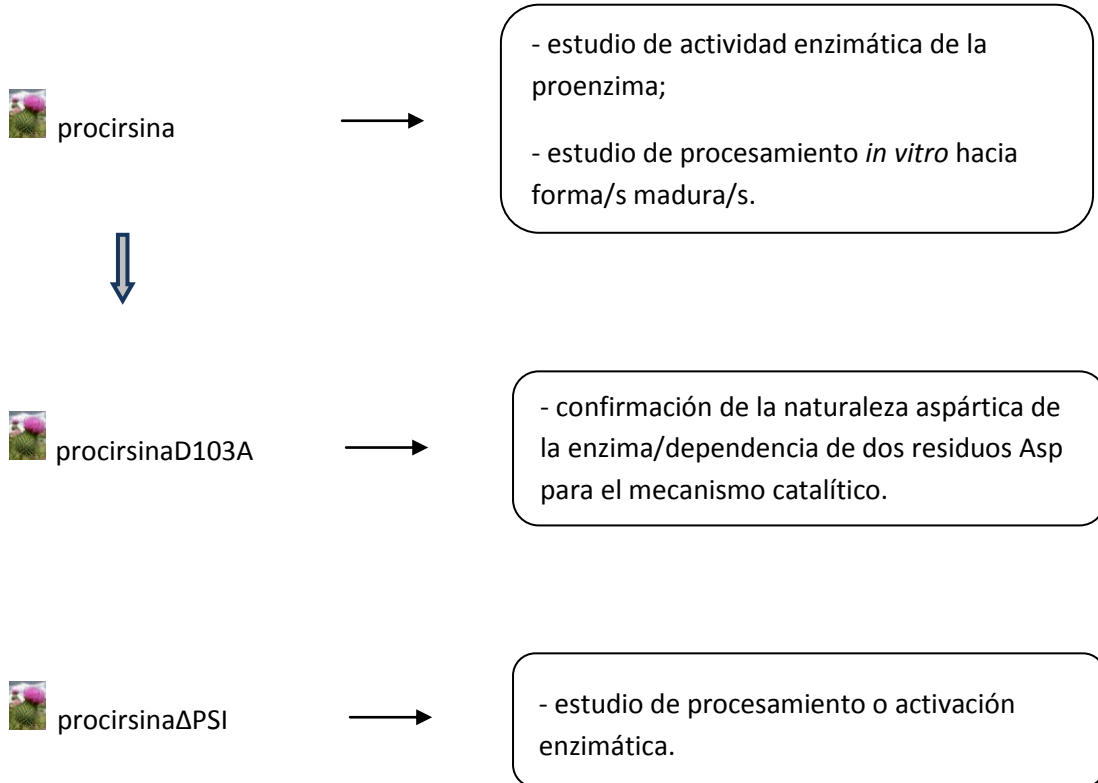
En esta tesis se utilizó la cepa de *E. coli* BL21 Star™ (DE3), donde la designación DE3 indica que la misma contiene el lisogen DE3 que transporta el gen para la T7 RNA polimerasa bajo el control del promotor *lacUV5*, inducible mediante IPTG. La cepa también transporta un gen *rne* mutado (*rne131*) que codifica para una enzima RNasa trunca incapaz de degradar el ARNm, aumentando su estabilidad y por tanto, los niveles de expresión de los genes heterólogos. Por otro lado, esta cepa es deficiente tanto en la proteasa *lon* (*E. coli* B/r) como en la proteasa de la membrana externa, OmpT, reduciéndose de este modo la degradación de las proteínas heterólogas expresadas.

De este modo, la cepa BL21 Star resulta adecuada para la expresión de proteínas recombinantes con altos rendimientos a partir de cualquier vector de expresión basado en el sistema T7 que contenga el gen heterólogo de interés.

I. Obtención de las construcciones para expresión de procirsina y mutantes en *E. coli*

A partir del plásmido pGEM-preproCV obtenido de la manera detallada en *Materiales y Métodos Generales*, se generó la construcción _pET-procirsinaconteniendo la secuencia que codifica para la proenzima sin el péptido señal. Esta construcción sirvió luego como molde en las mutagénesis que dieron lugar a las construcciones pET-procirsinaD103A. y

pET-procirsina Δ PSI. De esta forma se generó el material que permitió expresar el conjunto de proteínas necesario para abordar el estudio de procirsina.



 *Construcción pET-procirsina, pET-procirsinaD103A y pET-procirsina Δ PSI*

La estrategia planteada para armar la construcción pET-procirsina consistió en amplificar primero el ADNc codificante de cirsina en su forma pro-, sin su péptido señal (pre-), en el vector pGEM-T Easy para su posterior subclonado entre los sitios *NheI* y *XhoI* del vector de expresión pET-23d.

Se diseñaron cebadores específicos de acuerdo a los datos obtenidos en la secuenciación automática del clon 8 de la construcción pGEM-preproCV. El inicio del prosegmento se identificó mediante predicción a través del software SignalP 4.0 Server (<http://expasy.org/tools/>) y comparación con otras proteasas aspárticas de origen vegetal ya estudiadas. Las secuencias de reconocimiento para las enzimas de restricción *NheI* y *XhoI* se incorporaron a los cebadores F y R (del inglés, *forward* y *reverse*), respectivamente. En el cebador R además se agregó un codón de fin de traducción a continuación del triplete que

codifica para el último aminoácido de procirsina. Las secuencias de los cebadores se muestran en la Tabla 11.

Cebador	Secuencia (5'→3')	Tm (°C) ^(*)
Nhe_proCV8_F	<u>GCTAGCTCCAATGATGGATTAATTAGAG</u>	58
Xho_proCV8_stop_R	<u>CTCGAGTCAAGCTGCTTCTGCAAATCC</u>	62

TABLA 11. Secuencias de los cebadores utilizados en la obtención de la construcción pET-procirsina. Los sitios de restricción y el codón de terminación de la traducción incorporados se encuentran subrayados por línea punteada y continua, respectivamente.

(*) $T_m = 4(C+G)+2(A+T)$

La amplificación de la secuencia de interés por PCR se llevó a cabo según la mezcla de reacción y las condiciones descritas en la *Técnica n°2* de *Materiales y Métodos Generales*, utilizando 1 U de enzima Taq Polimerasa (GE Healthcare Life Sciences) y el clon 8 de la construcción pGEM-preproCV como ADN molde. El segmento II del programa de termociclado consistió en 35 ciclos: 45 seg de desnaturalización a 94 °C; 45 seg de hibridación a 56 °C; 2 min de extensión a 72 °C.

El producto de amplificación del tamaño esperado se purificó desde un gel de agarosa 1% (p/v) para ser clonado en el vector pGEM-T Easy (ver *Técnicas n°3* y *n°5*), obteniéndose el plásmido pGEM-procirsina. La selección de clones positivos a partir de 6 colonias transformadas se realizó mediante análisis de restricción con las enzimas *NheI* y *XhoI*.

Luego una alícuota de una minipreparación de pGEM-procirsina (ver *Técnica n°6*) proveniente del clon 2 (Fig. 10) y otra del vector pET-23d fueron sometidos a una doble digestión *NheI-XhoI* durante 3,5 hs (ver *Técnica n°7*) a fin de generar extremos cohesivos en ambas moléculas de ADN. El fragmento correspondiente al ADNc de procirsina escindido a partir del plásmido pGEM-procirsina se subclonó entonces en el vector pET-23d siguiendo el procedimiento detallado en la *Técnica n°8*. La presencia del plásmido pET-procirsina se verificó nuevamente por doble digestión *NheI-XhoI* en 6 colonias de la placa de transformantes.

Las construcciones conteniendo las secuencias del mutante de sitio activo de procirsina (pET-procirsinaD103A) y de la procirsina sin el dominio PSI (pET-procirsinaΔPSI) se obtuvieron a partir del clon 2 del plásmido pET-procirsina generado (Fig. 10). Se empleó para ello el kit

comercial Quick Change Site-Directed Mutagenesis (Stratagene) y los pares de cebadores mostrados en la Tabla 12.

Cebador	Secuencia (5' → 3')
D103A_proCV_F	G TTCAC GTGATTTTT GCG ACCGGATCCTCTAATCTATGGGTG
D103A_proCV_R	CACCCATAGATTAGAGGATCCGGT CGC AAAAATCACAGTGAAC
ΔPSI_proCV_F	CTAAGGGAGGTGGCGGTGGCATGGGAGAATCAGCAGTAGATTG
ΔPSI_proCV_R	CCCATGCCACCGCCACCTCCCTTAGCACCAATTGCATG

TABLA 12. Secuencias de los cebadores utilizados en la obtención de construcciones en pET-23d conteniendo la mutación puntual D103A y la delección del dominio PSI. Las mutaciones introducidas se encuentran indicadas mediante letra negra y línea continua.

La mezcla de reacción usada se describe en la *Técnica n°2* con 5 U de enzima *Pfu* Turbo Polimerase (Stratagene o Agilent). La amplificación se llevó a cabo de acuerdo a los parámetros de termociclado indicados en el manual Quick Change Site-Directed Mutagenesis Kit (catálogo n°: 200518) utilizado.

Finalizada la reacción de PCR, el producto resultante se trató con 10 U de *DpnI* y se transformó según el protocolo detallado en el manual del kit, en células TOP10F' o XL1-Blue competentes (*Técnica n°4*). La presencia de la mutación correspondiente en un clon elegido al azar de la placa de transformantes se confirmó directamente por secuenciamiento automático.

En la Fig. 10, a modo de resumen se presenta un esquema de la construcción de los vectores pET-23d conteniendo las secuencias para la expresión de las proteínas procirsina *wild-type* y sus mutantes, D103A y ΔPSI.

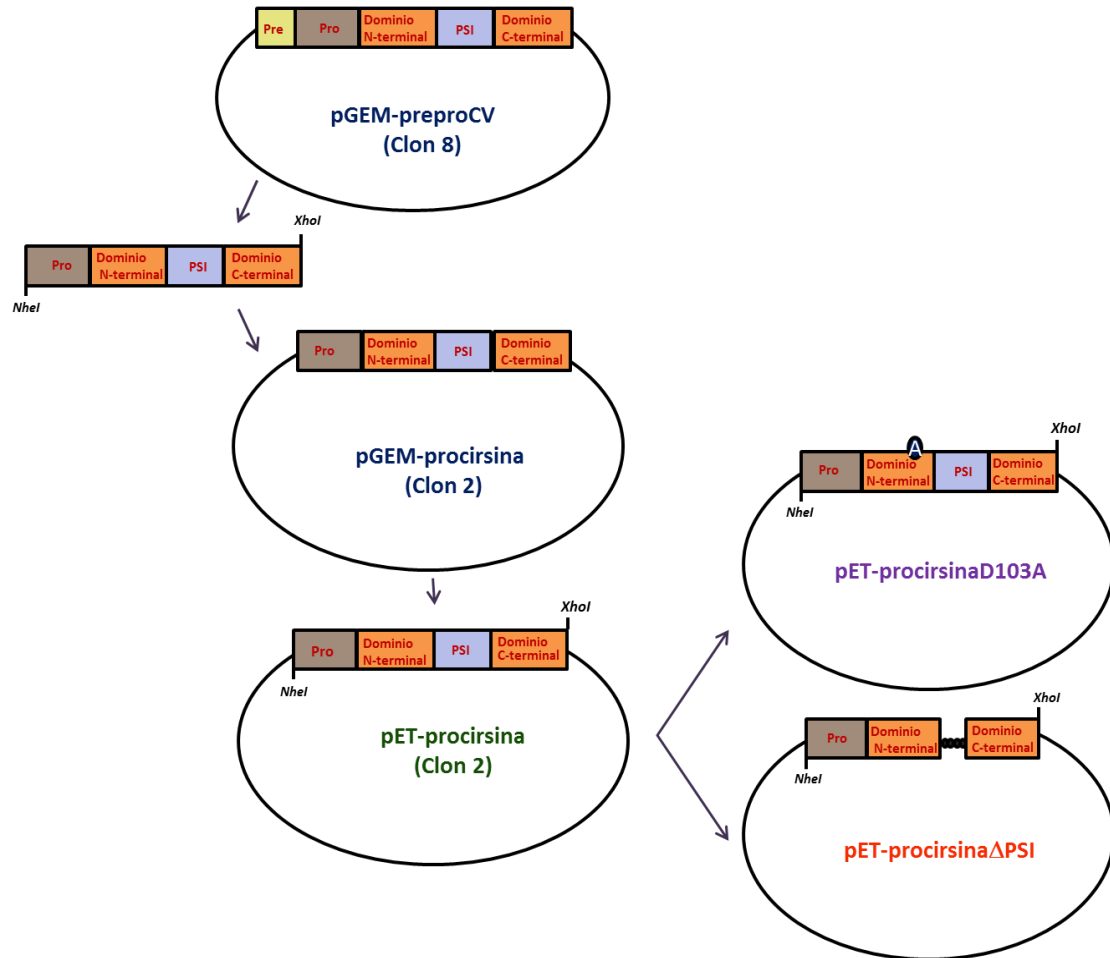


Fig. 10. Esquema de la construcción de los vectores de expresión para procirsina *wild-type* y sus mutantes en *E. coli* (pET23d).

II. Análisis filogenético

Se llevaron a cabo búsquedas de similitud de secuencia con el programa BLAST a fin de identificar las proteínas homólogas más cercanas a procirsina. Todas aquellas secuencias con *E-values* por debajo de 1×10^{-150} (120 proteínas) se sometieron al programa de alineamientos múltiples ClustalX y se realizó la estimación filogenética empleando el programa PHYML (Guindon et al., 2009), que hace uso del método de reconstrucción filogenética por máxima verosimilitud. Esta estimación se realizó usando un modelo de evolución de proteínas JTT + F con una distribución gamma discreta, basado en el intercambio de aminoácidos estimado a partir de matrices de bases de datos de las alineaciones de proteínas, que incorporan una frecuencia promedio de aminoácidos. Debido a este ajuste simultáneo de la topología y longitud de las ramas, sólo unas pocas iteraciones son suficientes para llegar a un punto de

topología óptimo. Como prueba de confiabilidad se realizó un Bootstrap no paramétrico de 500 replicantes seguido por un árbol consenso por la regla de la mayoría.

III. Expresión heteróloga de procirsina *wild-type* y sus mutantes

En todos los casos, la proteína de interés se expresó como cuerpos de inclusión (agregados insolubles e inactivos) recuperables en alta concentración y en forma relativamente pura, con menor susceptibilidad a la proteólisis bacteriana. Para tal fin se escogió el vector de expresión pET-23d (Novagen), apropiado por presentar altos niveles de expresión. Por otro lado se colocó un codón de fin de traducción inmediatamente después del inserto de modo que las proteínas se expresaron con sus extremos N- y C-terminales nativos en lugar de fusionadas a una cola de polihistidinas del vector para prevenir posibles interferencias durante el proceso de “refolding”.



Expresión de wt-procirsina, procirsinaD103A y -procirsinaΔPSI como cuerpos de inclusión

A fin de expresar la proteína procirsina y sus mutantes, una alícuota de células de la cepa BL21 Star de *E. coli* químicamente competentes, se transformó con el plásmido conteniendo la secuencia de la proteína a expresar (pET-procirsina, pET-procirsinaD103A o pET-procirsinaΔPSI), siguiendo el protocolo usado en la etapa de clonado en pGEM-T Easy (*Técnica n°5*) y plaqueando sobre medio sólido LB/Amp. Al día siguiente, se generó un preinóculo de 20 ml por inoculación de una colonia de la placa crecida a 37 °C en medio líquido LB/Amp (1 µg/ml) e incubando a 37 °C con agitación orbital suave. Transcurrida la noche, el preinóculo se transfirió a 1 litro de medio LB conteniendo 30 µg/ml de ampicilina en un frasco tipo Fernbach, prolongándose el crecimiento hasta alcanzar una DO_{600nm} de 0,6 a 0,7. En ese momento se indujo la expresión adicionando IPTG en una concentración final de 0,5 mM. Luego de 3 hs a 37 °C, las células se cosecharon por centrifugación del medio de cultivo a $9000 \times g$ durante 15 min y se resuspendieron en buffer TN (Tris-HCl 50 mM, NaCl 50 mM, pH 7,5). La disrupción celular se realizó por agregado de aproximadamente 7000 U/ml de Lisozima (Sigma) a la suspensión y un ciclo de congelado/descongelado. Posteriormente, el lisado se incubó en hielo durante 3 hs con DNasa I (12 µg/ml, Sigma) y $MgCl_2$ (20 mM). La mezcla resultante se volcó sobre 1 litro de buffer TN y se dejó durante toda la noche a 4 °C con

agitación. El sedimento colectado por un paso de centrifugación de 20 min a $10000 \times g$ y 4°C , se lavó una segunda vez en buffer TNT (Tris-HCl 50 mM, NaCl 50 mM, pH 7,5 con agregado de 1% de Triton X-100). La proteína expresada en forma de cuerpos de inclusión se recuperó por centrifugación a $9000 \times g$ y 4°C por 20 min y finalmente, se desnaturalizó durante 24 hs a 4°C en buffer Tris 0,1M–glicina 1 mM (pH 10,5), conteniendo urea 8 M, EDTA 1 mM, y β -mercaptoetanol 100 mM. Los cuerpos de inclusión desnaturalizados se centrifugaron una vez más en las mismas condiciones y al sobrenadante obtenido se le adicionó DTT 10 mM, glutatona reducida 1 mM, glutatona oxidada 0,1 mM y β -mercaptoetanol 100 mM. A continuación se realizó una dilución rápida (20 veces) de la solución de proteína desnaturalizada en Tris 20 mM y luego de ajustar el pH al valor de 8,0 se llevó a 4°C para permitir el plegado de la proteína.



Purificación de wt-procirsina, procirsinaD103A y procirsina Δ PSI

Transcurridos 42 días, la solución de proteína se concentró aproximadamente 100 veces mediante ultrafiltración de flujo tangencial en un sistema Pellicon 2 Mini Ultrafiltration Module (Millipore), y una celda de ultrafiltración (Amicon) con membrana de límite nominal de peso molecular de 10 kDa en ambos pasos. Con el objetivo de eliminar agregados de alto peso molecular e impurezas insolubles, la muestra se ultracentrifugó a $140000 \times g$ por 20 min a 4°C y el sobrenadante conteniendo la proteína recombinante se sometió a un proceso de purificación en dos etapas utilizando un sistema cromatográfico de media presión:

- 1) *cromatografía de exclusión molecular* a fin de separar la proteína plegada de formas agregadas y de péptidos de menor masa molecular generados por expresión parcial y/o degradación de la proteína, y de otros contaminantes. En este paso, se utilizó una columna HiLoad 26/60 Superdex S200 prep grade y buffer Tris-HCl 20 mM de pH 8,0 (ver *Técnica n°9*).
- 2) *cromatografía de intercambio iónico* mediante la cual se separaron diferentes formas de la proteína recombinante por sus propiedades de carga superficial. La cromatografía se realizó según lo indicado en la *Técnica n°10*, empleando buffer Tris-HCl 20 mM de pH 8,0 y el mismo buffer con agregado de NaCl en concentración 1 M para generar el gradiente salino.



Detección de la actividad endopeptidásica sobre un sustrato sintético (Screening)

Con el objeto de identificar las fracciones cromatográficas que contuvieran la enzima aspártica procirsina, se determinó en las mismas la actividad endopeptidásica a lo largo de la purificación. Se utilizó el sustrato sintético MCA-Lys-Lys-Pro-Ala-Glu-Phe-Phe-Ala-Leu(DNP) (GenScript), donde MCA es el grupo 7-amino-4-metoxicumaril y DNP es el 2,4-dinitrofenilo. El péptido contenido en este sustrato ha mostrado ser muy específico para PAs típicas (Verísimo et al., 1996), las cuales cortan preferentemente enlaces en los que participan aminoácidos aromáticos.

Luego de ser eluidas de cada una de las columnas cromatográficas, las fracciones se analizaron por SDS-PAGE en condiciones reductoras y no reductoras (según *Técnica nº14*). Por otro lado, la cantidad de proteína contenida en cada fracción se estimó por absorbancia a 280 nm en un espectrofotómetro Nanodrop (Thermo) y seguidamente se incubó con el sustrato sintético en una relación proteína:sustrato de 1:1 en masa. El ensayo se llevó a cabo a 37 °C, durante toda la noche en buffer ácido acético-acetato de sodio 50 mM, NaCl 0,1 M (pH 4,0) en un volumen de reacción de 500 µl. La separación de los fragmentos resultantes de la digestión del sustrato se realizó mediante HPLC de fase reversa tal como se detalla en la *Técnica nº12*.

A fin de confirmar la naturaleza aspártica de la enzima estudiada, este proceso de *screening* de actividad endopeptidásica se repitió con la forma mutante de sitio activo, procirsina D103A.



Estudio de procesamiento proteolítico de wt-procirsina

Muestras de wt-procirsina recombinante purificadas a pH 8,0 como se indicó en el ítem preanterior, se incubaron ON a 37 °C en relación 1:1 (v/v) con los siguientes buffers de activación: ácido cítrico-citrato de sodio 0,1 M (pH 3,0 y 4,0); ácido acético-acetato de sodio 0,1 M (pH 5,0 y 6,0) y Tris-HCl 0,1 M (pH 7,0).

Después de la incubación, las muestras se analizaron mediante SDS-PAGE en condiciones reductoras tal como se describe en las *Técnicas nº14 y nº15* y se ensayaron frente al sustrato fluorogénico de PAs típicas a pH 4,0 y 37 °C de la manera descrita previamente.



Digestión triptica en geles de poliacrilamida

Se obtuvo la huella peptídica (PMF) de la banda principal de 50 kDa proveniente de la expresión del mutante de sitio activo, procirsina D103A, por MALDI-TOF MS como se explica en la *Técnica n°19*.

IV. Caracterización bioquímica de wt-procirsina recombinante



Actividad proteolítica de wt-procirsina recombinante sobre diferentes sustratos modelo

La actividad proteolítica de wt-procirsina recombinante purificada se testeó a pH 4,0 y 37 °C, contra varios sustratos sintéticos fluorogénicos considerados sustratos modelos para la actividad de diferentes PAs. La velocidad de hidrólisis se monitoreó por incremento en la intensidad de fluorescencia ($\lambda_{\text{excitación}}/\lambda_{\text{emisión}}$ péptidos fluorogénicos MCA/DNP: 328/393 nm; $\lambda_{\text{excitación}}/\lambda_{\text{emisión}}$ péptidos fluorogénicos EDANS/DABCYL: 335/490 nm) por un período de 600 seg (Fig. 11) en un espectrofluorómetro Jobin Yvon FluoroMax-3 Spectrofluorometer (Horiba). Los sustratos testeados en concentración 2 μM en la mezcla de reacción y relación proteína:sustrato entre 0,9:1 y 1,8:1 (v/v), se enumeran en la Tabla 13.

Sustrato	Secuencia	Buffer de reacción
sustrato de PAs típicas	MCA-KKPAEFFALK(DNP)	HOAc / NaOAc 50 mM, NaCl 0,1 M
sustrato para CDR1	MCA-KLHPEVLVLEK(DNP)	HOAc / NaOAc 50 mM, NaCl 0,1 M, DMSO 8% (v/v)
sustrato 1 de PA atípicas	MCA-KKLADVVALEKK(DNP)	HOAc / NaOAc 50 mM, NaCl 0,1 M, DMSO 8% (v/v)
sustrato para BACE 1	MCA-KSEVNLDAEFK(DNP)	HOAc / NaOAc 50 mM, NaCl 0,1 M, DMSO 8% (v/v)
sustrato 1 de la proteasa de HIV	RE(EDANS)SQNYPIVQK(DABCYL)R	HOAc / NaOAc 50 mM, NaCl 0,1 M, DMSO 9,6% (v/v);
sustrato 1 de Renina	RE(EDANS)IHPFHLVIHTK(DABCYL)R	HOAc / NaOAc 50 mM, NaCl 0,1 M, DMSO 9,6% (v/v)

Tabla 13. Sustratos sintéticos evaluados en la determinación de actividad proteolítica de procirsina.

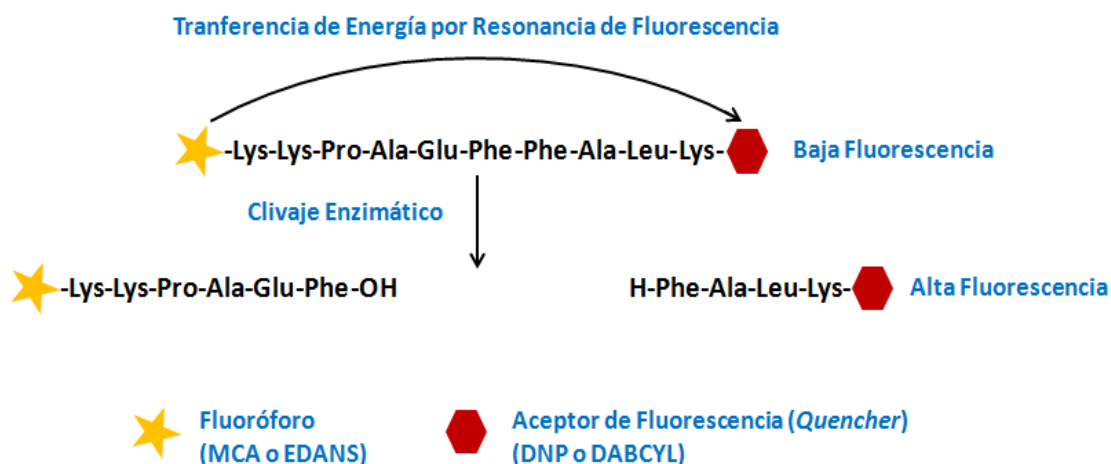


Fig. 11. Fundamento de la medida de actividad proteolítica con los sustratos empleados.



Determinación del perfil de pH y temperatura

El perfil de pH característico de la actividad endopeptidásica de las fracciones purificadas de wt-procirsina se determinó utilizando el sustrato fluorogénico para PAs típicas MCA-KKPAEFFALK(DNP) en concentración 2 μ M en la mezcla de reacción.

El rango de pH ensayado fue 2,25–7,0 en los siguientes buffers: ácido cítrico-citrato de sodio 50 mM (pH 2,25 y 2,5); ácido acético-acetato de sodio 50 mM (pH 3,5; 3,75; 4,0; 4,5; 5,0; 6,0); Tris-HCl 50 mM (pH 7,0), conteniendo en todos los casos NaCl 100 mM. El aumento en la intensidad de fluorescencia producido por la hidrólisis del sustrato se monitoreó espectrofluorométricamente ($\lambda_{\text{excitación}}/\lambda_{\text{emisión}}$: 328/393 nm) durante 10 min.

El efecto de la temperatura sobre la actividad proteolítica de wt-procirsina recombinante se estudió mediante la preincubación de la enzima en buffer ácido acético-acetato de sodio 50 mM, NaCl 100 mM (pH 4,0) a diferentes temperaturas entre 10 y 65 °C por el lapso de 10 min. Transcurrido el tiempo de incubación se agregó el sustrato fluorogénico, se mezcló por inversión rápida y se siguió la intensidad de fluorescencia emitida a 393 nm durante 10 min por excitación del fluoróforo a 328 nm.



Ensayos de inhibición

Con el fin de establecer el mecanismo proteolítico de las enzimas contenidas en las diferentes preparaciones enzimáticas se emplearon inhibidores específicos de grupo (Dunn, 2001) sobre la actividad endopeptidásica empleando el sustrato fluorogénico para PAs típicas.

Las muestras de proteínas se preincubaron durante 6 min a temperatura ambiente con cada uno de los siguientes inhibidores: pepstatina A (1 μM , inhibidor de aspartilpeptidasas), E-64 (10 μM , inhibidor de cisteinpeptidasas), Pefabloc (1 mM, inhibidores de serinpeptidasas), EDTA (5 μM , inhibidor de metalopeptidasas), amastatina (10 μM , inhibidor de aminopeptidasas) y DTT (1 mM, reductor de puentes disulfuro). Finalizado el tiempo de incubación, se agregó el sustrato, en proporción 3 veces mayor a la masa de proteínas, y se determinó la actividad endopeptidásica de la manera indicada en el inciso preanterior. En cada caso se incluyó como control el solvente en que el inhibidor se encontraba disuelto.



Determinación de parámetros cinéticos

Para los estudios cinéticos a 37 °C se usó el péptido sintético de PAs típicas ya descrito, en un rango de concentraciones entre 0,5 y 5 μM en buffer ácido acético-acetato de sodio 50 mM de pH 4,0 conteniendo NaCl 100 mM. La concentración de enzima utilizada en el ensayo se determinó por titulación del sitio activo de wt-procirsina purificada con pepstatina A. Los ensayos se llevaron a cabo con el mismo espectrofluorómetro y el mismo programa que en los incisos anteriores, durante 600 seg.

La relación entre la variación de fluorescencia y la concentración de sustrato peptídico hidrolizado en la reacción se determinó por medición del cambio total de fluorescencia ocurrido en la hidrólisis completa del mismo por parte de la enzima pepsina.

Los parámetros cinéticos V_{MAX} y K_{M} se calcularon a partir del gráfico Lineweaver-Burk obtenido con el módulo de cinética del programa Sigma Plot 10.0 para Windows (SPSS Inc.).



Titulación del sitio activo con pepstatina A

Para determinar la proporción de moléculas de enzima activas, se empleó una modificación del método descrito por Salvesen y Nagase (Salvesen and Nagase, 2001). Muestras de

wt-procirsina purificada se incubaron durante 6 min a 37 °C en presencia de cantidades crecientes del inhibidor específico de aspartilpeptidasas típicas pepstatina A (0 a 0,5 pmoles), con el que forman un complejo en relación 1:1. Luego, se determinó la actividad residual sobre el sustrato de PAs típicas a pH 4,0 como se indicó previamente y los valores obtenidos se graficaron en función de los pmoles de pepstatina A. La abscisa al origen de este gráfico indicó el número de picomoles de enzima cuando la actividad se inhibió completamente.



Determinación de la especificidad de sustrato sobre la cadena β de la insulina oxidada

La cadena β de la insulina oxidada (Sigma) en una concentración de 1 mg/ml se incubó con wt-procirsina recombinante purificada (relación enzima:sustrato 1:100) en buffer ácido acético-acetato de sodio 0,1 mM de pH 4,0. Transcurrida la digestión ON a 37 °C, el sobrenadante de la mezcla de reacción se centrifugó durante 6 min a 20000 × *g* y los fragmentos de digestión se separaron por cromatografía en fase reversa como se describe en la *Técnica nº12*. Los dos péptidos mayoritarios aislados se sometieron a identificación por LC/MS/MS en un equipo 4000 QTRAP (Applied Biosystems) del Servicio de la Unidad de Proteómica del Centro de Neurociencias y Biología Celular de la Universidad de Coimbra.



Digestión de κ-caseína

La κ-caseína (0,3 mg/ml) se incubó en buffer ácido acético-acetato de sodio 50 mM (pH 5,5) con wt-procirsina en una relación en masa proteína:sustrato de 1:250. La reacción de digestión procedió a 37 °C durante 18 hs y el producto resultante se analizó luego por SDS-PAGE y tinción con Coomassie como se indica en las *Técnicas nº14 y nº15*. Se incluyó también un control sin agregado de wt-procirsina en la mezcla de reacción.

V. Actividad coagulante de leche

La mezcla de reacción se llevó a cabo agregando una muestra de wt-procirsina purificada en un volumen 5 veces mayor de leche descremada (Molico) al 12% (p/v), preparada en CaCl₂ 10 mM de pH 6,5 y termostalizada a 37 °C. Se continuó la incubación a esa temperatura hasta observar la formación de un coágulo macroscópico. Como controles negativos de coagulación

se emplearon buffer de purificación y wt-procirsina preincubada 10 min con pepstatina A (1 μ M) en las condiciones de trabajo.

VI. Caracterización Bioinformática de wt-procirsina

Para estimar moldes estructurales putativos que permitan predecir la estructura de procirsina, se usaron métodos de asignación de plegamiento tales como FFas03 (Jaroszewski et al., 2005) y HHPred (Soding et al., 2006). Usando los moldes detectados, se construyeron modelos estructurales mediante modelado por homología con el programa Modeller (Sali and Blundell, 1993). Se identificaron los mejores modelos usando la función objetivo del Modeller y la calidad de los mismos se asignó usando potencial DOPE, en el programa Modeller así como también en el Prosall (Wiederstein and Sippl, 2007). Los puentes disulfuro se consideraron explícitamente en el proceso de modelado. La carga de la región del prosegmento en función del pH se estimó mediante el programa Protein Calculator v3.3 (<http://www.scripps.edu/~cdputnam/protcalc.html>).

Resultados—Procirsina

Obtención de procirsina y sus mutantes en *Escherichia coli*

A partir del ADNc que codifica para la cirsina completa, en su forma de preprocirsina (número de acceso GenBank: JN703462), contenido en el vector pGEM-T Easy, se subclonó la región correspondiente a la procirsina, sin el péptido señal putativo (primeros 25 residuos), en el vector pET-23d. Empleando esta construcción, se obtuvieron posteriormente el mutante del sitio activo y el mutante sin el dominio PSI. Cada una de las construcciones obtenidas según lo indicado en *Materiales y Métodos—Procirsina*, se expresaron en *Escherichia coli*.

I. Clonado de ADNc y análisis de la secuencia

Con el fin de identificar y caracterizar posteriormente las proteasas aspárticas típicas de *Cirsium vulgare*, se comenzó por sintetizar el ADNc de preprocirsina usando ARN total aislado de pimpollos. Luego, el híbrido ADNc-ARN se usó como molde en la reacción de PCR con cebadores diseñados específicamente para regiones conservadas de los extremos 5' y 3' del marco de lectura abierto (*ORF*) de PAs de plantas. El producto amplificado obtenido, de 1530 pb, se clonó y varios clones conteniendo el inserto deseado se secuenciaron en forma automática (Fig. 12). Este análisis resultó en la identificación de al menos cuatro secuencias de aminoácidos diferentes (Anexo 1) exhibiendo altos niveles de identidad de secuencia entre ellas y con otras secuencias de PAs típicas de plantas disponibles en bases de datos públicas, sugiriendo claramente la presencia de varias PAs en flores de *C. vulgare*. Una de estas secuencias de ADNc se seleccionó para ser posteriormente caracterizada. Su secuencia nucleotídica y la secuencia aminoacídica que de ella se deduce se muestran en la Fig. 13 (número de acceso GenBank: JN703462).

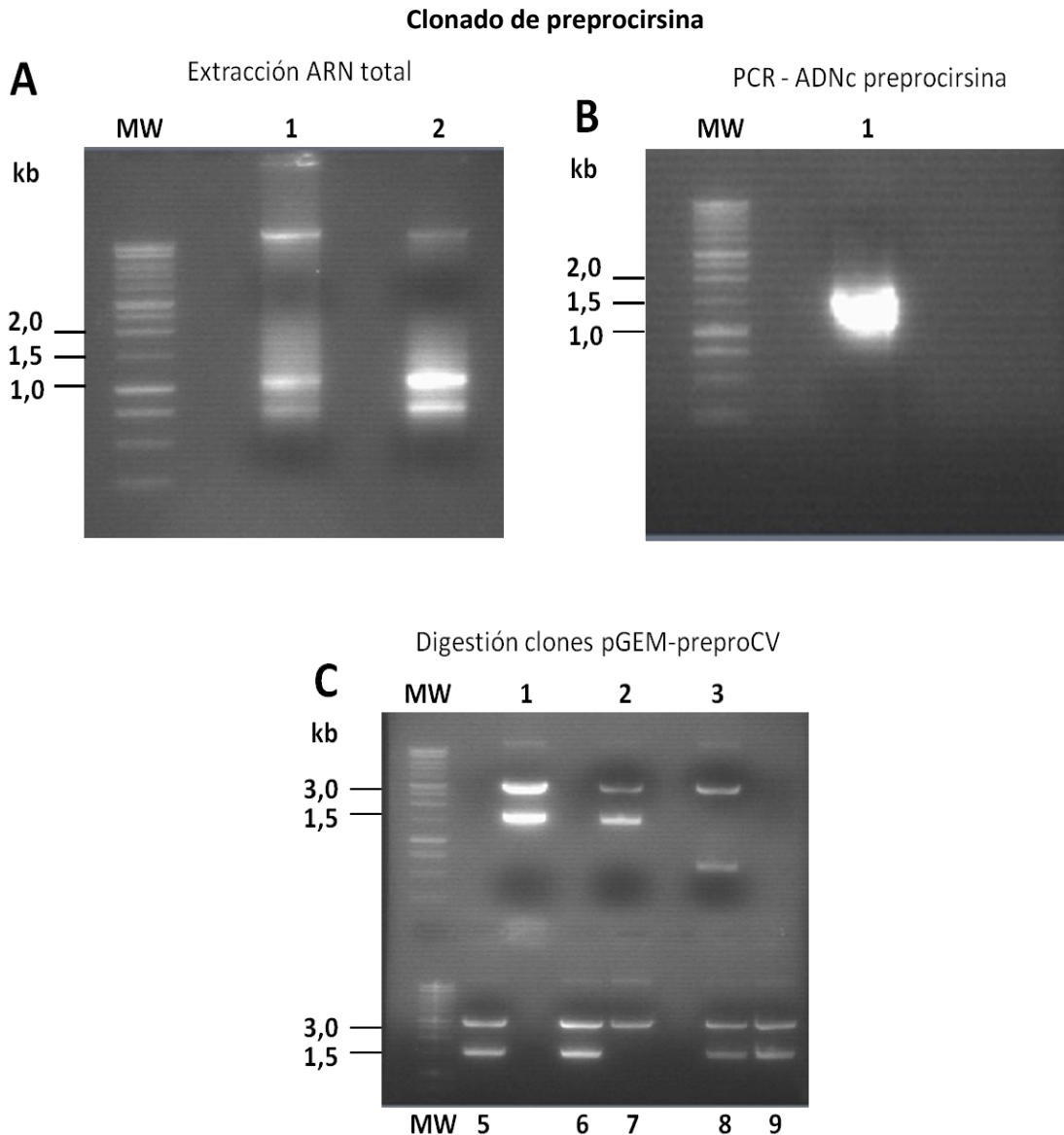


Fig. 12. Amplificación y clonado de la secuencia correspondiente al marco de lectura abierto de preprocirsina. La primera cadena de ADNc se sintetizó a partir del ARN total aislado de flores inmaduras de *C. vulgare* y se utilizó como molde en la reacción de amplificación (PCR) con cebadores específicamente diseñados para los extremos 5' y 3' altamente conservados en las secuencias codificantes de PAs. El producto de amplificación obtenido se clonó y secuenció en forma automática. **(A)** Análisis de la extracción de ARN total por electroforesis en gel de agarosa (1%): *Calle MW*, marcadores de masas moleculares de ADN de 1 kb; *Calles 1 y 2*, muestras del ARN total de *C. vulgare* aislado. **(B)** Análisis electroforético en gel de agarosa (1%) del producto amplificado mediante RT-PCR: *Calle MW*, marcadores de masas moleculares de ADN 1 kb; *Calle 1*, ADNc completo de 1,5 kb. **(C)** Análisis por restricción de clones transformados con el plásmido pGEM-preproCV: *Calles MW*, marcadores de masas moleculares de ADN de 1 kb; *Calles 1 a 9*, plásmido pGEM-preproCV aislado de los clones 1 a 9 respectivamente, y digeridos con la enzima de restricción *Styl*.

Este producto proteico, llamado preprocirsina, codifica para una preproenzima de 509 residuos con un péptido señal hidrofóbico de 24 aminoácidos predicho utilizando el programa SignalP 4.0 (Petersen et al., 2011), un prosegmento de 44 residuos, y un polipéptido de 441 aminoácidos de largo interrumpido por 105 residuos correspondientes al dominio PSI. La preproenzima tiene una masa molecular predicha de 55 kDa con un pI teórico de 5,3. La secuencia de aminoácidos deducida del ADNc, muestra la estructura primaria característica de las proteasas aspárticas típicas de plantas, además de las dos tríadas catalíticas altamente conservadas (DTG 103-105 y DTG 290-292) y un residuo tirosina en la posición 147 (los cuales se corresponden a los motivos DTG 32-34, DTG 215-217 y el residuo Y75, respectivamente, usando la numeración de la pepsina como referencia).

Asimismo, fueron predichos dos sitios de N-glicosilación conservados en varias PAs típicas de plantas (Domingos et al., 2000; Vieira et al., 2001) mediante el programa NetNGlyc (Blom et al., 2004) en los residuos N139 (motivo NGT) y N400 (motivo NET), este último en el dominio PSI.

Cuando se comparó la secuencia aminoacídica de preprocirsina con la de otros precursores de PAs de plantas, ésta mostró un elevado grado de similitud con cyprosina B (CAA57510) proveniente de *Cynara cardunculus* (identidad: 95%, positivos: 98%), cenprosina (CAA70340) de *Centaurea calcitrapa* (identidad: 94%, positivos: 97%), la PA de *Helianthus annuus* (BAA76870, identidad: 83%, positivos: 92%), cardosina A (AJ132884) y cardosina B (AJ237674) de *C. cardunculus* (identidad: 75%, positivos: 83% and 85%, respectivamente) y fitepsina (X56136) de *Hordeum vulgare* (identidad: 71%, positivos: 84%) perteneciendo todas estas peptidasas a la familia A1.

La elevada similitud de secuencia entre el precursor de cirisina y el de cyprosina se muestra en la Fig. 14, en donde se puede observar que la mayoría de las diferencias de aminoácidos entre ambas secuencias se hallan en los residuos que conforman el péptido señal, el prosegmento y el dominio PSI; mientras que únicamente 5 residuos localizados en las secuencias correspondientes a los dominios N- y C- terminales de las proteínas maduras, son diferentes. Notablemente, de manera mayoritaria, estas diferencias corresponden a sustituciones conservativas sugiriendo que cirisina y cyprosina están estrechamente relacionadas.

Secuencia de preprocirsina

```

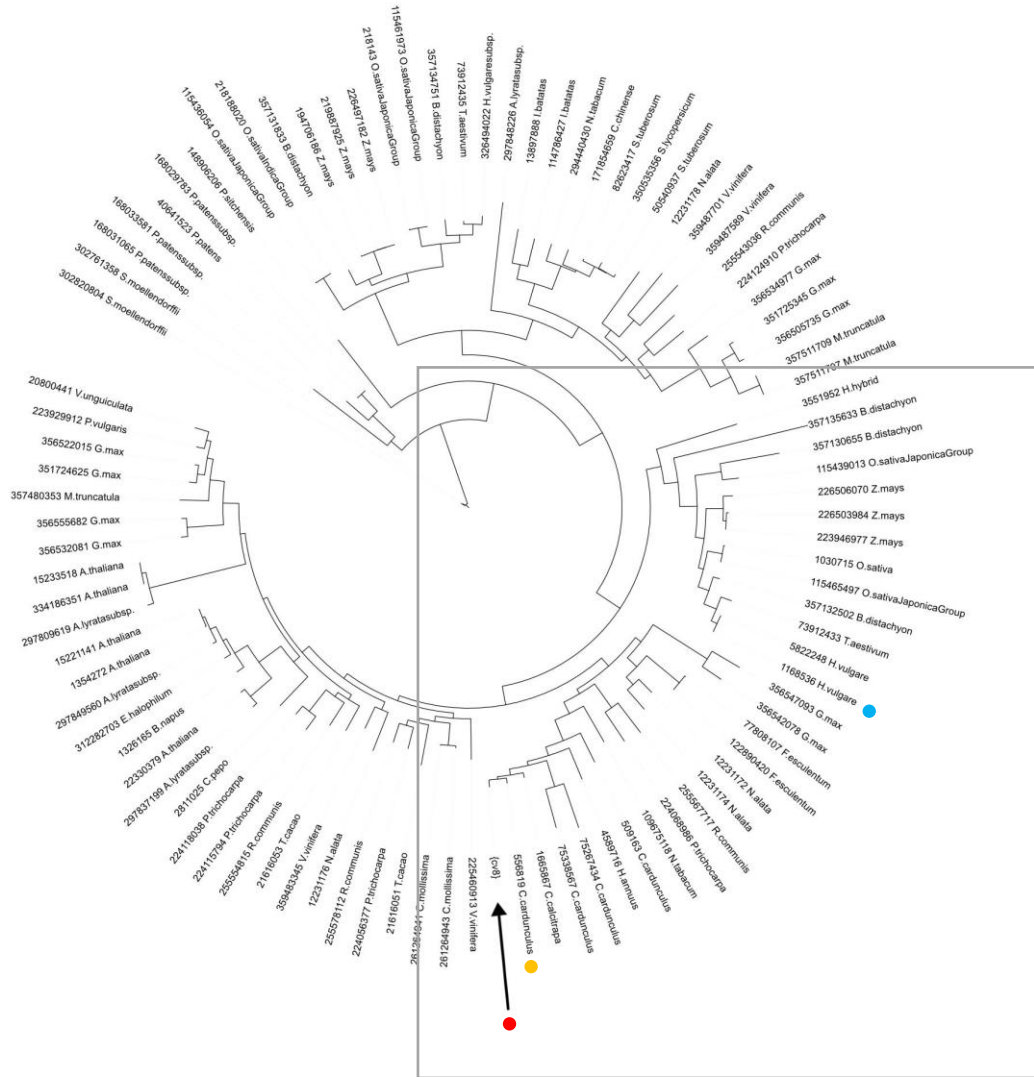
M G T S I K A S L L A L F L L F L L S P T A
ATGGGTAOCTCAATCAAAGCAAAGCCTGCTTGCTTGT TTT TGCTTTTTCCTTCATCACCTACCGCA
I S V S N D G L I R V G L K K R K V D Q I N
ATTTGGTCTCCAATGATGGATTAATTAGAGTTGGACTTAAAAAGAGGAAGGTGGACCAAATTAAC
Q L S G H G A S M E G K A R K D F G F G G T
CAACTTAGTGGACATGGGGCGTCTATGGAAGGAAAGGCTAGAAAAGATT TTGGCTTTGGTGGTACT
L R D S D S D I I A L K N Y M D A Q Y Y G E
CTGAGGGACTCGGACAGTGACATTATTGCCTAAAGAACTACATGGATGCTCAGTACTATGGTGAG
I G I G A P P Q K F T V I F D T G S S N L W
ATTGGTATTGGAGCTCCACCTCAAAGTCACTGTGATTT TTGACA CCGGAAGTTCTAATCTATGG
V P S A K C Y F S V A C L F H S K Y K S S H
GTGCCCTTCTGCAAAGT GCTACTTTT CAGTAGCTT GCCTTT TTTACTCAAAGTACAAGT CGAGCCAT
S S T Y K K N G T S A A I Q Y G T G S I S G
TCAAAGTCACTCAAAGAAAATGGCACATCTGCTGCTATTCAATATGGAAGTGGATCAATCTCTGGT
F V S Q D S V K L G D L V V K E Q D F I E A
TTTGTAGCCAAGACTCTGTCAAAC TGGTGATCTTGTGTAAAGAGCAGGATT TTATAGAGGCA
T K E P G I T F L A A K F D G I L G L G F Q
ACCAAAGAGCCTGGCATCACTT TCTTGGCAGCCAAGT TTGATGGTATACTTTGGCCTTGGAT TTCAG
E I S V G K S V P V W Y N M V N Q G L V Q E
GAGATCTCTGT TGGGAATCTGTTCTGCTGTTACAACATGGTTAATCAAGGTCCTGTTCAGAA
P V F S F W F N R N A N E E E G G E L V F G
CCTGTGT TTT CCT TTT GGT TTAATCGCAATGCTAATGAGGAAGAAGGGGTGAAC TCGTGT TTGGT
G V D P N H F K G K H T Y V P V T X K G Y W
GGTGTGACCCTAATCATT TTAAGGTAAGCACACATATGTCCCTGTGACCGANAAGGGCTATTGG
Q F N M G D V L I E D K T T G F C S D G C A
CAGT TTAATATGGGTGACGTTCTTATTGAAGATAAAACTACCGGAT TTT GTT CTGATGGTTGTGCA
A I A D S G T S L L A G P T A I I T E I N H
GCAATTCGGACTCTGGAACTCTT TGT TGGCAGGTC CAAAGGCTATTAATTACTGAAATCAATCAT
A N G A K G V M S Q Q C K T L V S Q Y G K S
GCAATGGTGCTAAGGGAGTCA TGAGCCAGCAAT GCAAAACAT TGGTTAGTCAGTATGGAAGAGT
I I E M L L S E A Q P D K I C S Q M K L C T
ATAATCGAGATGCTCCTGTCTGAGGCGCAACCTGATAAAAATCTGTTCTCAAATGAAGT TATGCACT
F D G A R D V S S I I E S V V D K N N G K S
TTTGTGGTGCTCTGTGATGTTAGTTCAA TAA TTGAGAGTGTGGTTGACAAGAACAACGGCAAGTCT
S G G A N D E M C T F C E M A V V W M Q N Q
TCTGGTGGCGCAAACGATGAGATGTGTAOCTTCTGTGAAA TGGCAGTCGTTT GGTGCAAAACCAA
I K R N E T E D N I I N Y V N E L C D R L P
ATCAAACGAAA CGAGACTGAAGATAACA TAA TCAACTATGTCAAACGAGCTGTGTGATCGCTTACCC
S P M G E S A V D C N S L S S M P N I A F T
AGTCCAATGGGAGAAT CAGCAGTAGATTGCAACAGTCTTT CCTCCATGCCAATATTGCOCTTCACA
I G G K V F F E L C P E Q Y I L K I G E G E A
ATTGGTGGCAAAGTTT TCGAGCTCTGCCAGAACAGTACATCTCAA AAA TCGGTGAGGGAGAAGCA
A Q C I S G F T A M D V A P P R G P L W I L
GCACAATGCATCAGTGGAT TCACTGCTATGGATGTGGCCCCTCCTCGTGGACCTCTATGGATCTTG
G D V F M G R Y H T V F D Y G K S R V G F A
GGAGATGTCT TTA TGGGTCGAT ACCACACGGTGT TCGATTATGGCAAAGT CACGAGTTGGAT TTGCA
E A A -
GAAGCAGCTTGA

```

Fig. 13. ADNc que codifica para la forma completa de cirsina (preprocirsina) y su secuencia aminoacídica derivada. Secuencias nucleotídica y aminoacídica traducidas de preprocirsina, se encuentran indicadas las regiones correspondientes a: péptido señal N-terminal (púrpura); prosegmento (celeste) e inserto específico de planta (naranja). Los sitios putativos de glicosilación se especifican mediante un círculo y las tríadas catalíticas se muestran en rojo y recuadradas. La secuencia del ADNc de preprocirsina ha sido incorporada en la base de datos GenBank bajo el número de acceso JN703462.

Análisis filogenético de procirsina

A



B

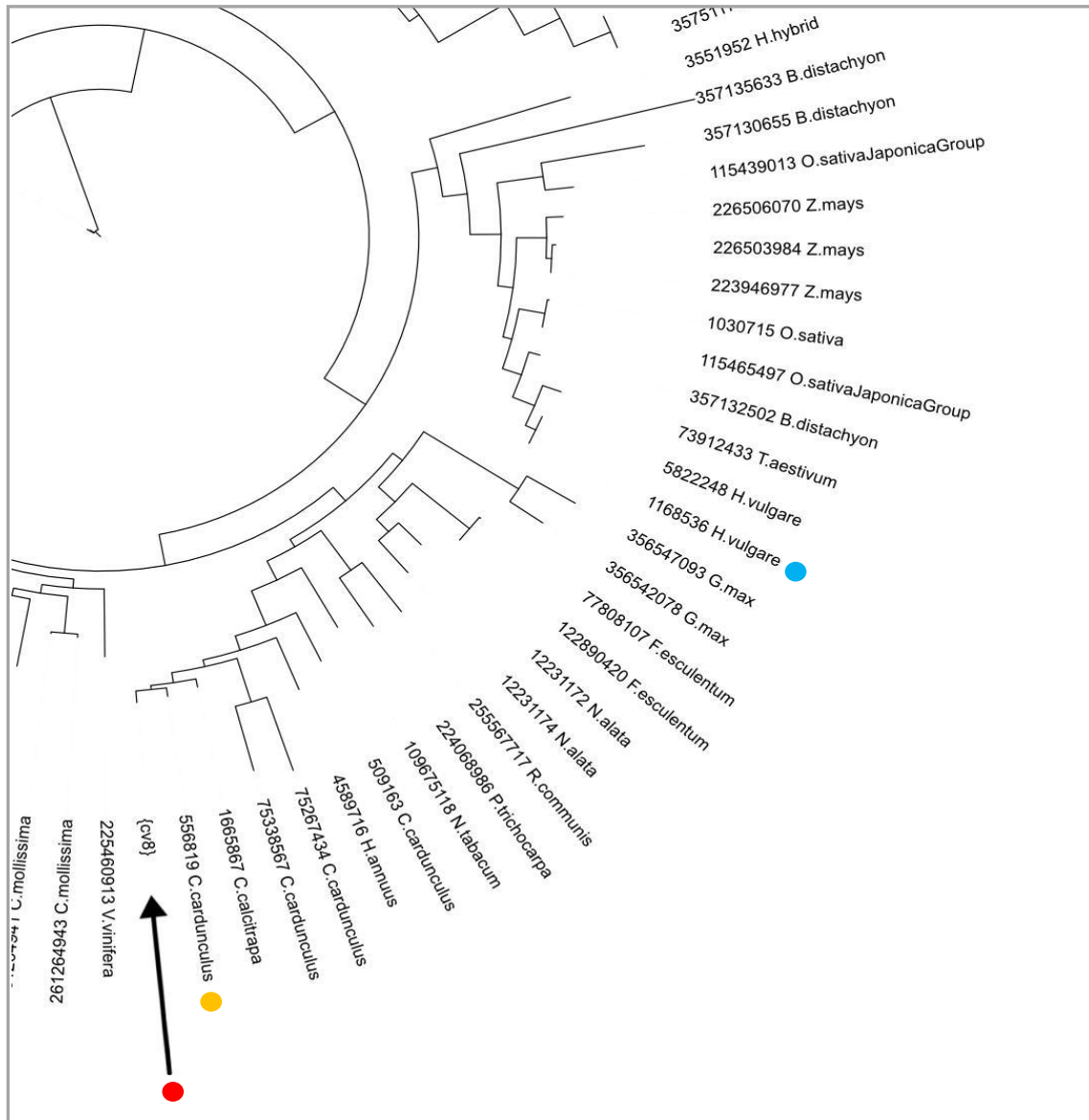


Fig. 15. Análisis filogenético de proteínas homólogas cercanas a procirsina. El árbol filogenético se obtuvo por cálculo de probabilidad máxima usando el programa PhyML tal y como se detalla en *Materiales y Métodos—Procirsina*. Las diferentes proteínas en el árbol están indicadas con el correspondiente número de identificación del NCBI (National Center for Biotechnology Information), seguido por los nombres del género y de la especie. Procirsina está indicada con un punto rojo y procypsosina B con un punto amarillo, mientras que la proteína con estructura cristalográfica conocida (código PDB: 1qdm) correspondiente a *H. vulgare*, se muestra mediante un punto azul.

III. Expresión heteróloga de wt-procirsina y sus mutantes

A fin de caracterizar esta nueva PA de *C. vulgare*, el fragmento de ADNc que codifica para la procirsina se amplificó y subclonó en el vector de expresión pET-23d. La secuencia correspondiente al clon número 4 de pET-procirsina, que por análisis de restricción mostró contener un inserto del tamaño esperado para el ADNc de procirsina (Fig. 16, A), se confirmó mediante secuenciación automática. Posteriormente se utilizó este plásmido para producir la proenzima recombinante en un sistema de expresión basado en *E. coli* (cepa BL21 Star (DE3)). Una proteína de alrededor de 50 kDa correspondiente a la procirsina recombinante se recuperó de la fracción insoluble en forma de cuerpos de inclusión (Fig. 16, B). Los cuerpos de inclusión se lavaron en dos pasos y, después de desnaturalizarse en urea 8 M, el *refolding* se indujo mediante un paso de dilución rápida seguido del ajuste lento del pH hasta un valor final de 8,0.

Obtención y expresión de pET-procirsina

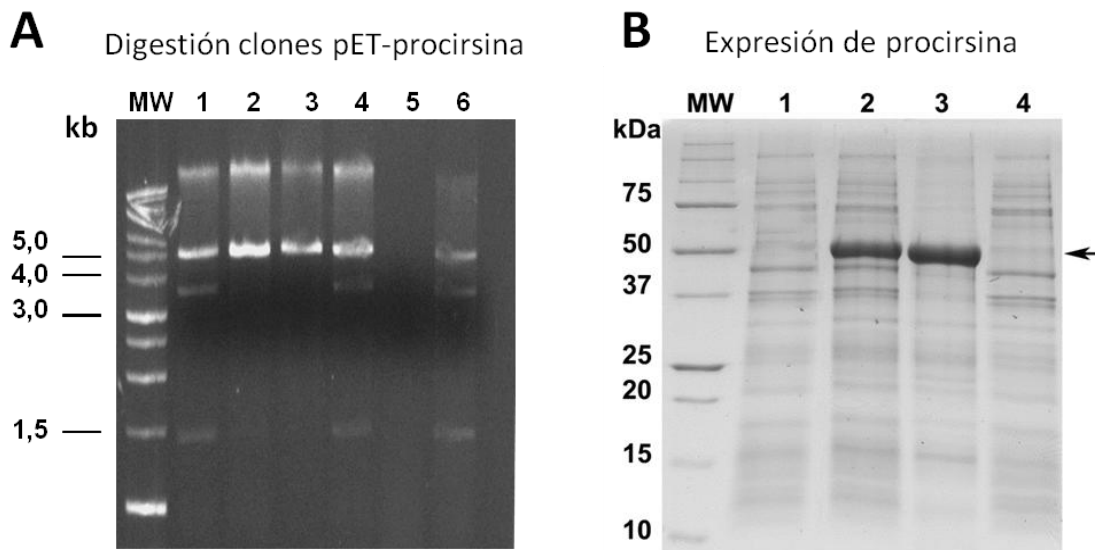


Fig. 16 Obtención de la construcción pET-procirsina y análisis de la expresión en *E. coli*. El plásmido conteniendo la secuencia de ADNc que codifica para procirsina sin el péptido señal putativo (los primeros 25 residuos aminoacídicos) se expresó en la cepa BL21Star (DE3) de *E. coli* en forma de cuerpos de inclusión. **(A)** Análisis de restricción de los clones transformados con el plásmido **pET-procirsina**: *Calles MW*, marcadores de masas moleculares de ADN de 1 kb; *Calles 1 a 6*, plásmidos aislados de los clones 1 a 6 respectivamente, sometidos a doble digestión con las enzimas de restricción *XhoI* y *NheI*. **(B)** Se colectaron muestras a lo largo de la expresión y se analizaron por SDS-PAGE: *Calle MW*, marcadores de masas moleculares; *Calle 1*, extractos proteicos de *E. coli* colectados inmediatamente antes de inducir la expresión; *Calle 2*, extractos proteicos de *E. coli* colectados después de 3 horas de expresión; *Calle 3*, fracción proteica insoluble; *Calle 4*, fracción proteica soluble. El gel se coloreó con Coomassie Brilliant Blue.

Después del *refolding*, los agregados solubles de alto peso molecular se removieron mediante cromatografía de exclusión molecular en una columna de filtración en gel HiLoad 26/60 Superdex 200 (Fig. 17 A).

Purificación de procirsina

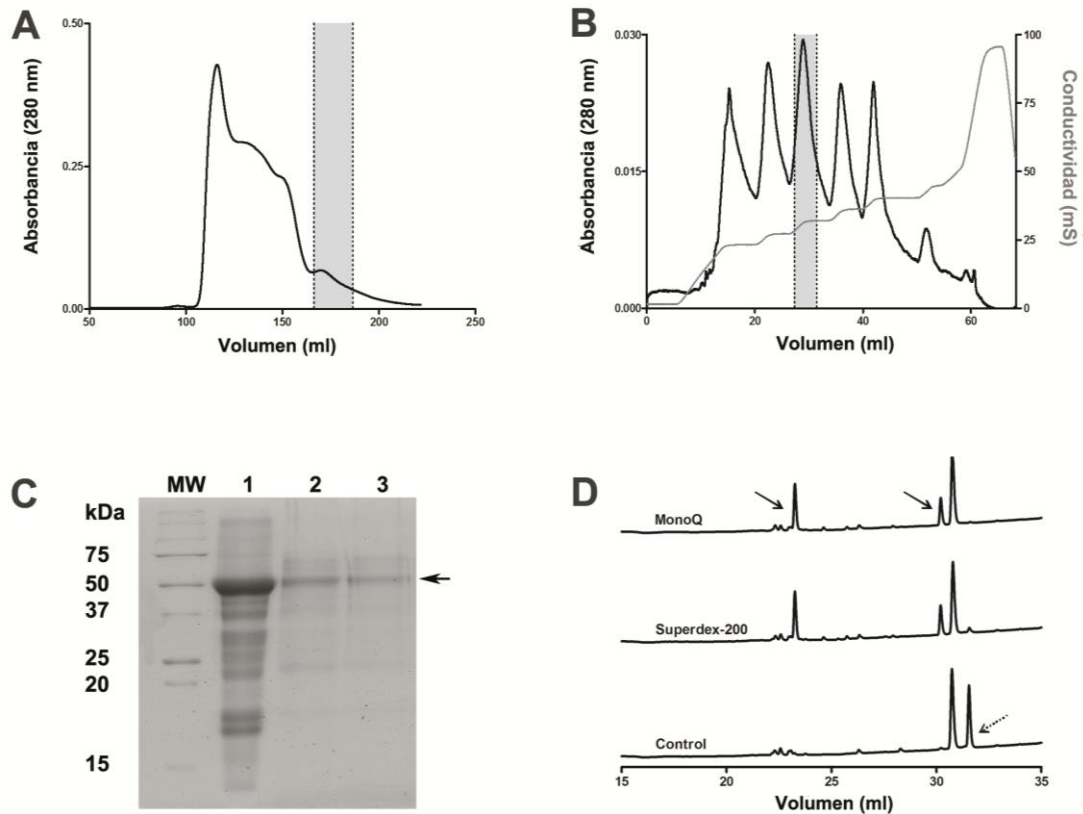


Fig. 17. Purificación de wt-procirsina recombinante. La procirsina *wild-type* (wt-procirsina) se expresó en forma de cuerpos de inclusión en *E. coli* y se sometió a *refolding* por dilución rápida. La proteína recombinante se purificó por cromatografía de exclusión molecular seguida de intercambio iónico como se describió en *Materiales y Métodos—Procirsina*. Las fracciones que mostraron actividad frente al sustrato fluorogénico de PAs típicas están indicadas en las figuras A y B por el sombreado entre líneas punteadas. **(A)** Cromatografía de exclusión molecular de wt-procirsina usando una columna HiLoad 26/60 Superdex 200. **(B)** El eluato de la columna Superdex 200 (S-200) se juntó en un único pool de fracciones activas y se recromatografió en una columna de intercambio iónico Mono Q. **(C)** Las fracciones colectadas en los diferentes pasos del procedimiento de purificación de wt-procirsina se analizaron por SDS-PAGE y se tiñeron con Coomassie Brilliant Blue: *Calle MW*, marcadores de masas moleculares; *Calle 1*, eluato proveniente de S-200 aplicado a la columna Mono Q (fracciones juntas que se encuentran resaltadas entre líneas de puntos en B); *Calle 2*, eluato de la columna Mono Q (fracciones reunidas que se indican entre líneas de puntos en A). **(D)** Perfiles en RP-HPLC de la actividad frente al sustrato de PAs típicas. Se ensayó la actividad proteolítica de los eluatos de wt-procirsina (fracciones reunidas de S-200 y de Mono Q) mediante digestión del sustrato MCA-Lys-Lys-Pro-Ala-Glu-Phe-Phe-Ala-Leu-Lys(DNP) y los fragmentos resultantes se separaron por RP-HPLC. Los productos de digestión se señalan con flecha sólida. El control corresponde al péptido sintético incubado en ausencia de enzima.

La actividad de las fracciones colectadas se evaluó frente al sustrato de PAs típicas MCA-Lys-Lys-Pro-Ala-Glu-Phe-Phe-Ala-Leu-Lys(DNP) a pH 4,0 y los productos de digestión correspondientes se separaron por RP-HPLC. La versión cromogénica equivalente de este sustrato ha sido usado para ensayar la actividad de una amplia variedad de PAs típicas (Dunn and Hung, 2000; Castanheira et al., 2005) dado que se mostró que las enzimas de tipo pepsina cortan preferencialmente enlaces peptídicos que comprometen aminoácidos aromáticos o hidrofóbicos en las posiciones P1 y P1'. La Fig. 17 D muestra un perfil de elución representativo obtenido por RP-HPLC empleando una columna C18 para las fracciones activas de procirsina recombinante (resaltadas con línea de puntos en la Fig. 17 A). Las fracciones activas se juntaron y purificaron por cromatografía de intercambio iónico en una columna Mono Q (Fig. 17 B). Se ensayó la actividad proteolítica bajo condiciones similares por RP-HPLC (Fig. 17 D) y las fracciones activas correspondientes al segundo paso de purificación (región resaltada con líneas de puntos en la Fig. 17 B) se juntaron para su análisis por SDS-PAGE confirmando la presencia de una proteína de masa molecular esperada ca. 50 kDa (Fig. 17 C, Línea 3).

Con el objetivo de confirmar el rol en la actividad proteolítica de los residuos aspárticos catalíticos predichos a partir de la secuencia, se reemplazó el residuo Asp103 potencialmente activo por un residuo Ala y luego la proenzima mutante se expresó, purificó y testeó bajo las mismas condiciones descritas para la proteína recombinante *wild-type* (Fig. 18). Se analizaron las fracciones purificadas por SDS-PAGE mostrando un patrón proteico similar al observado para wt-procirsina (Fig. 18 C). Los ensayos de actividad confirmaron la ausencia de la misma frente al sustrato de PAs típicas para el mutante de sitio activo (Fig. 18 D), como se propuso mediante el análisis de la secuencia aminoacídica deducida a partir de la secuencia nucleotídica.

Purificación de procirsinaD103A

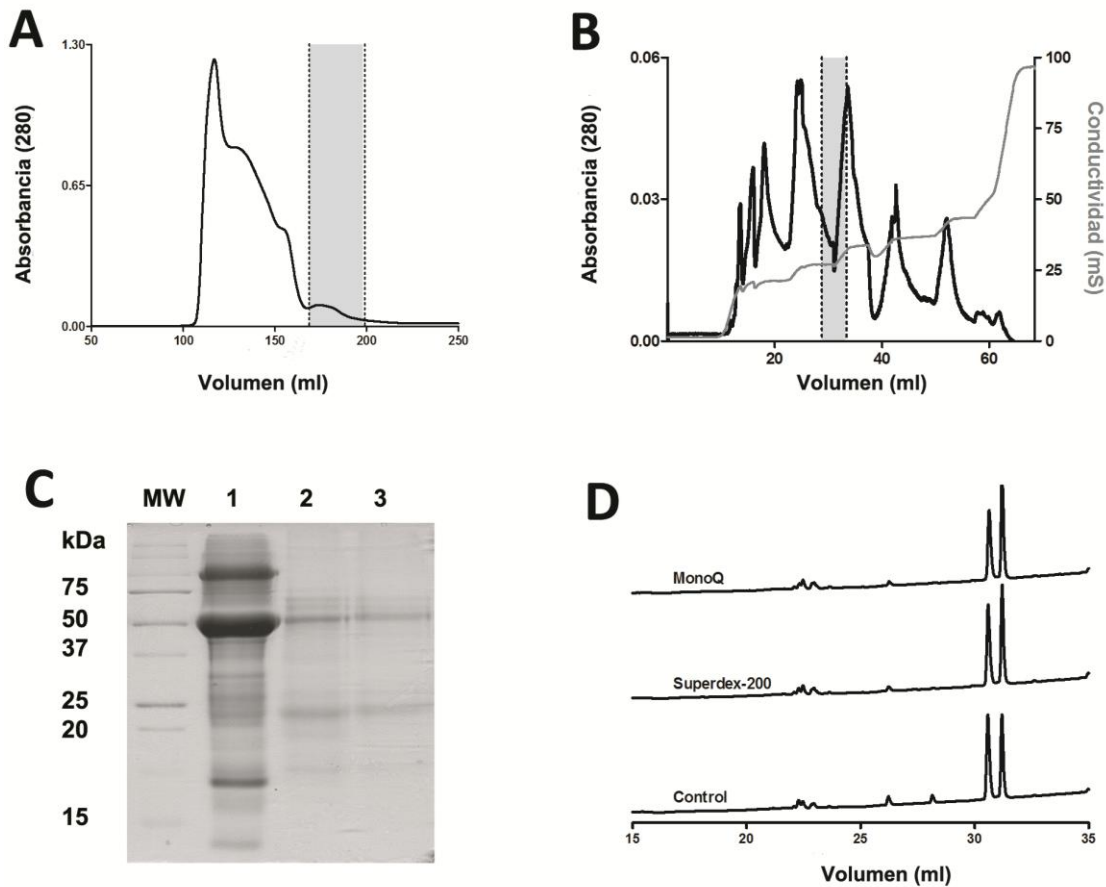


Fig. 18. Purificación de procirsinaD103A recombinante. El mutante de sitio activo procirsinaD103 se expresó en forma cuerpos de inclusión en *E. coli*, luego se sometió a *refolding* por dilución rápida y purificación siguiendo el mismo procedimiento que para wt-procirsina. **(A)** Cromatografía de exclusión molecular de procirsinaD103A usando una columna HiLoad 26/60 Superdex 200. **(B)** Las fracciones eluidas de la columna Superdex 200 (S-200) se juntó en un único pool de fracciones activas y se recromatografiaron en una columna de intercambio iónico Mono Q. Las fracciones eluidas de cada columna que mostraron actividad frente al sustrato fluorogénico de PAs típicas están indicadas en las figuras A y B por la zona sombreada entre líneas punteadas. **(C)** Análisis mediante SDS-PAGE de las fracciones colectadas durante diferentes pasos de purificación de procirsinaD103A. El gel se tiñó con Coomassie Brillante Blue: *Calle 1*, eluato de columna S-200 aplicado a Mono Q; *Calle 2*, eluato proveniente de la columna Mono Q. **(D)** Perfiles de elución de RP-HPLC obtenido en los ensayos de *screening* de actividad usando fracciones purificadas de procirsinaD103A recombinante. La actividad proteolítica de los eluatos (Superdex 200 y MonoQ) se ensayó por digestión del sustrato de PAs típicas MCA-Lys-Lys-Pro-Ala-Glu-Phe-Phe-Ala-Leu-Lys(DNP) y los fragmentos resultantes se separaron por RP-HPLC. EL control corresponde al péptido incubado en ausencia de la enzima.

La identidad de la proteína mutante de sitio activo se confirmó por MALDI-TOF-MS (Fig. 19 y Tabla 14). Por otro lado, no se encontró coincidencias cuando el perfil peptídico de la proteína analizada se comparó con bibliotecas teóricas de péptidos generadas a partir de las secuencias de la base de datos MASCOT.

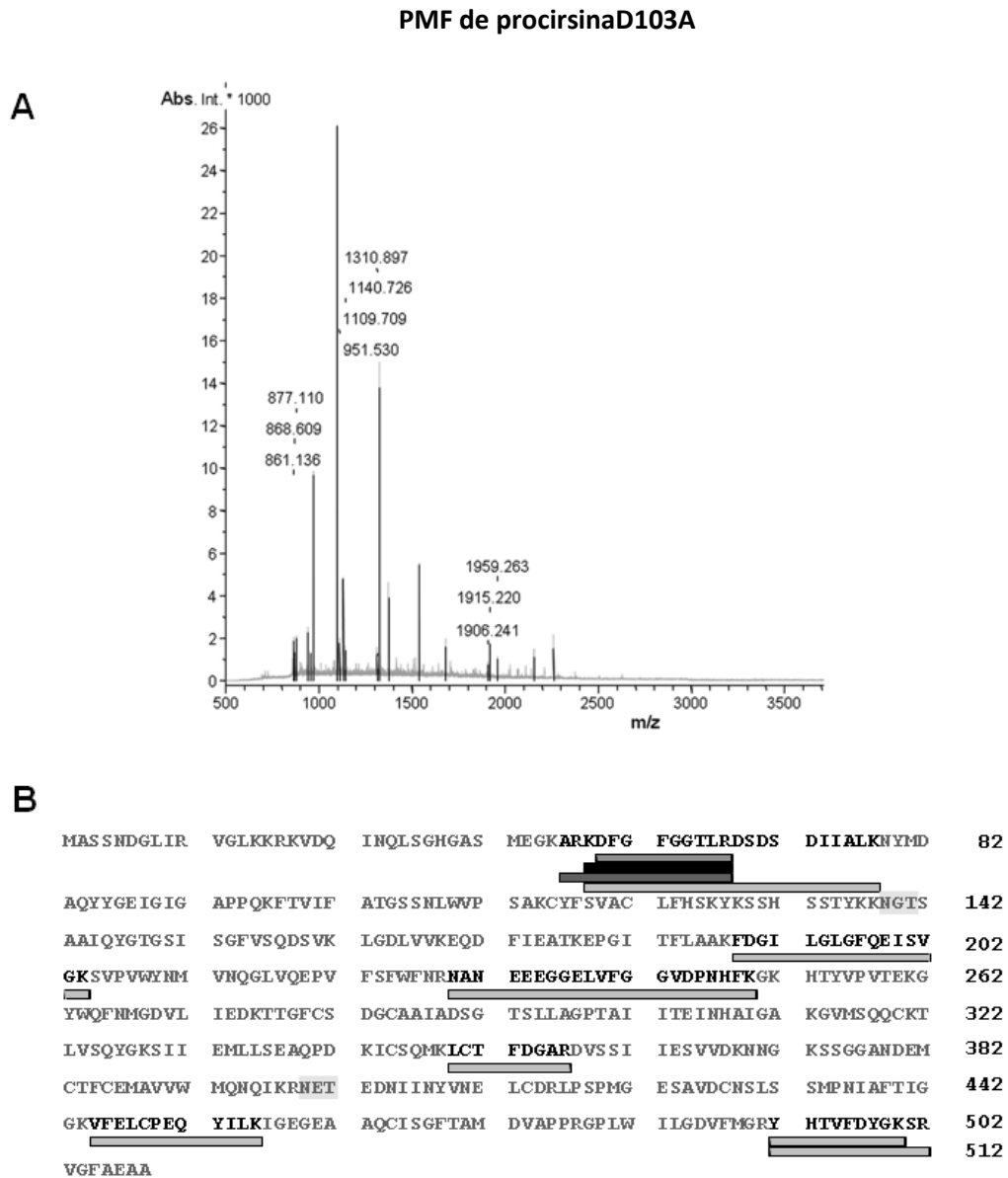


Fig. 19. Análisis por espectroscopía de masas MALDI-TOF de las digestiones tríplicas de procirsinaD103A. (A). Las masas moleculares de los péptidos observados que coinciden con aquellas obtenidas por digestión tríplica teórica se indican sobre los picos correspondientes del espectro. **(B)** Secuencia de procirsinaD103A, las barras debajo de la misma representan el solapamiento de las secuencias de los péptidos identificados. La escala de grises se correlaciona con la intensidad de los picos. La secuencia del motivo putativo de glicosilación se encuentra sombreada.

m/z	Intensidad
861,136	1652,49
868,609	1132,21
877,110	1780,87
939,524	2280,02
951,530	1291,95
969,579	9673,36
1097,714	27308,89
1109,709	1758,12
1129,678	4784,67
1140,726	1443,07
1310,897	1080,26
1324,904	13790,25
1372,873	3904,85
1539,024	5443,40
1680,153	1576,51
1906,241	759,72
1915,220	1520,40
1959,263	807,37
2155,432	1128,00
2259,371	1498,03

Tabla 14. Análisis por espectroscopía de masas MALDI-TOF de las digestiones tripticas de procirsinaD103A. El residuo catalítico putativo Asp103 fue reemplazado por un residuo alanina en la secuencia aminoacídica de procirsina; la proenzima mutante fue expresada en *E. coli* y purificada bajo las mismas condiciones empleadas para wt-procirsina recombinante.

Para determinar si el precursor de cirsina recombinante era capaz de sufrir un procesamiento de autoactivación *in vitro* hacia formas intermediarias y/o formas maduras, como se ha descrito para otras PAs típicas de plantas (Glathe et al., 1998; Domingos et al., 2000; Castanheira et al., 2005), wt-procirsina recombinante purificada se incubó con diferentes buffers abarcando el rango de pH entre 3,0 y 7,0. No se observó ningún patrón de conversión definido en el análisis por SDS-PAGE a ningún valor de pH testado (Fig. 20) ni se detectó aumento significativo de la actividad proteolítica como resultado de esta incubación. De este modo, no pudo asignarse claramente ningún intermediario ni forma madura para la procirsina recombinante mediante estos ensayos de activación.

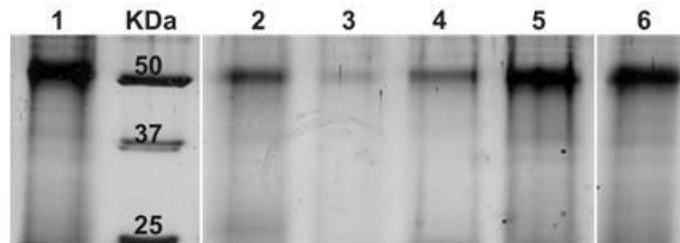
Procesamiento *in vitro* de wt-procirsina

Fig. 20. Efecto del pH sobre el procesamiento proteolítico de wt-procirsina recombinante. Wt-procirsina recombinante purificada se incubó durante toda la noche a 37 °C en un rango de pH desde pH 3,0 a pH 7,0 (ácido cítrico-citrato de sodio 100 mM, pH 3,0 y 4,0; ácido acético-acetato de sodio 100 mM, pH 5,0 y 6,0; Tris-HCl 0,05M, pH 7,0). Las fracciones solubles se analizaron mediante SDS-PAGE: *Calle 1*, wt-procirsina recombinante (control sin incubación); *Calles 2 a 6*, wt-procirsina recombinante incubada a pH 3,0; 4,0; 5,0; 6,0 y 7,0 respectivamente. Wt-procirsina mostró autodegradación bajo condiciones francamente ácidas pero ningún patrón definido de conversión fue observado.

A fin de evaluar en mayor profundidad el rol del dominio PSI en el procesamiento proteolítico y en la actividad de la procirsina, se diseñó una tercera construcción en la cual se eliminó la secuencia entera codificante para el PSI (región entre Val315 y Pro420). Esta construcción (Δ PSI-procirsina) se expresó en *E. coli* en forma de cuerpos de inclusión, se replegó y purificó siguiendo el mismo procedimiento que el descrito para la wt-procirsina recombinante. Sin embargo, el producto proteico purificado mostró ser inactivo contra el sustrato sintético MCA-Lys-Lys-Pro-Ala-Glu-Phe-Phe-Ala-Leu-Lys(DNP).

IV. Caracterización bioquímica de wt-procirsina recombinante



Actividad proteolítica de wt-procirsina recombinante sobre diferentes sustratos modelo

Dado los resultados obtenidos con las diferentes construcciones de procirsina recombinante testeadas, todos los estudios de caracterización se focalizaron exclusivamente en wt-procirsina. Las fracciones activas purificadas de wt-procirsina recombinante se juntaron en un único *pool* y se ensayó su capacidad de cortar varios sustratos sintéticos modelos. Sólo el sustrato fluorogénico MCA-Lys-Lys-Pro-Ala-Glu-Phe-Phe-Ala-Leu-Lys(DNP), de PAs típicas, se clivó significativamente en estos ensayos. El sustrato diseñado para PAs descritas como atípicas MCA-Lys-Leu-His-Pro-Glu-Val-Leu-Phe-Val-Leu-Glu-Lys(DNP) (Simões et al., 2007) sólo se cortó pobremente, mientras que un segundo sustrato para este grupo de PAs,

MCA-Lys-Lys-Leu-Ala-Asp-Val-Val-Asn-Ala-Leu-Glu-Lys-Lys(DNP) (Simões I), no resultó clivado por wt-procirsina recombinante. Tampoco se observó hidrólisis con los siguientes sustratos fluorogénicos: sustrato de BACE 1 [MCA-Lys-Ser-Glu-Val-Asn-Leu-Asp-Ala-Glu-Phe-Lys(DNP)], sustrato 1 de renina [Arg-Glu(EDANS)Ile-His-Pro-Phe-His-Leu-Val-Ile-His-Thr-Lys(DABCYL)Arg] y sustrato 1 de la proteasa de HIV [Arg-Glu(EDANS)Ser-Gln-Asn-Tyr-Pro-Ile-Val-Asn-Lys(DABCYL)Arg]. Este patrón de especificidad de corte es consistente con la alta similitud encontrada entre la secuencia aminoacídica de procirsina y la de otras PAs típicas. La Tabla 15 resume los resultados obtenidos con los diferentes sustratos empleados.

Sustrato	Secuencia	Eficiencia de corte
sustrato de PAs típicas	MCA-KKPAEFFALK(DNP)	cortado
sustrato para CDR1	MCA-KLHPEVLVLEK(DNP)	pobrementemente cortado
sustrato 1 de PA atípicas	MCA-KKLADVVALEKK(DNP)	no cortado
sustrato para BACE 1	MCA-KSEVNLDAEFK(DNP)	no cortado
sustrato 1 de la proteasa de HIV	RE(EDANS)SQNYPIVQK(DABCYL)R	no cortado
sustrato 1 de Renina	RE(EDANS)IHPFHLVIHTK(DABCYL)R	no cortado

Tabla 15. Actividad proteolítica de wt-procirsina recombinante sobre diferentes sustratos modelo.



Determinación del perfil de pH y temperatura

El efecto del pH sobre la actividad proteolítica de la wt-procirsina recombinante se estudió usando el mismo sustrato—MCA-Lys-Lys-Pro-Ala-Glu-Phe-Phe-Ala-Leu-Lys(DNP)—en un rango de valores de pH entre pH 2,25 y pH 7,0 (Fig. 21 A). La actividad máxima se observó a pH 4,0 lo cual es consistente con el pH ácido óptimo reportado para otras PAs típicas (Kervinen et al., 1993; Ramalho-Santos et al., 1998; White et al., 1999; Castanheira et al., 2005; Mazorra-Manzano and Yada, 2008; Sampaio et al., 2008). La mitad de la actividad máxima registrada en el perfil de pH se observó a pH 3,19 y a pH 4,45.

Se determinó además la dependencia de la hidrólisis con la temperatura. Como se muestra en la Fig.21 B, la actividad máxima se detectó entre los 30 °C y 37 °C a pH 4,0. Para temperaturas

por encima de 37 °C hubo un descenso en la actividad enzimática, reteniéndose a los 65 °C sólo un 10% de la actividad de la wt-procirsina recombinante.

Efecto del pH y de la temperatura en la hidrólisis del sustrato MCA-KKPAEFFALK(DNP)

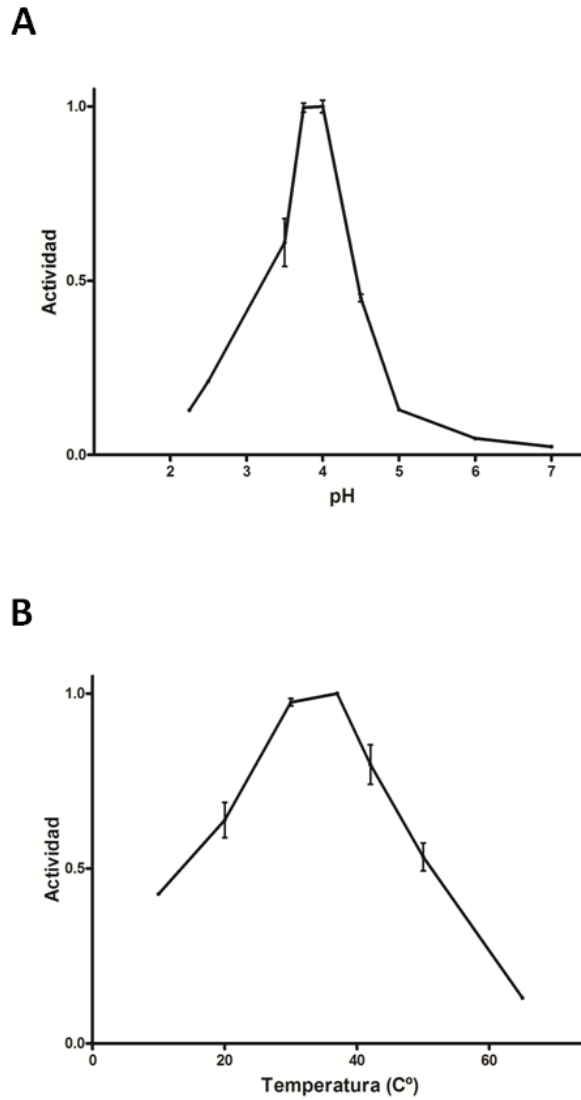


Fig. 21. Efecto del pH y de la temperatura sobre la actividad proteolítica de wt-procirsina recombinante. Wt-procirsina recombinante purificada se ensayó frente al péptido sintético fluorogénico MCA-Lys-Lys-Pro-Ala-Glu-Phe-Phe-Ala-Leu-Lys(DNP). **(A)** Los estudios de actividad a diferentes valores de pH se realizaron por incubación de wt-procirsina a 37 °C con buffers entre pH 2,25 y pH 7,0 conteniendo NaCl 100 mM (ácido cítrico-citrato de sodio 50 mM, pH 2,25 y 2,5; ácido acético-acetato de sodio 50 mM, pH 3,5 a 6,0; Tris-HCl 50 mM, pH 7,0). **(B)** Estudios de actividad a diferentes temperaturas. Wt-procirsina purificada recombinante se incubó a temperaturas entre 10 °C y 65 °C en buffer ácido acético-acetato de sodio 50 mM de pH 4,0 con NaCl 100 mM.



Determinación de la especificidad de sustrato sobre la cadena β de la insulina oxidada

Por otro lado, la cadena β de la insulina oxidada se usó como sustrato para evaluar la especificidad de procirsina recombinante debido a su alto contenido de aminoácidos hidrofóbicos (Kervinen et al., 1993; Bleux et al., 1998; Park et al., 2000; Payie et al., 2003; Athauda et al., 2004; Guevara et al., 2004; Simões et al., 2007). La hidrólisis de la cadena β de la insulina oxidada por incubación con wt-procirsina recombinante se siguió mediante RP-HPLC (Fig. 22). Los productos de digestión predominantes se analizaron por LC/MS/MS revelando que el sitio de clivaje más susceptible era Leu15-Tyr16.

Hidrólisis de la cadena β de insulina oxidada

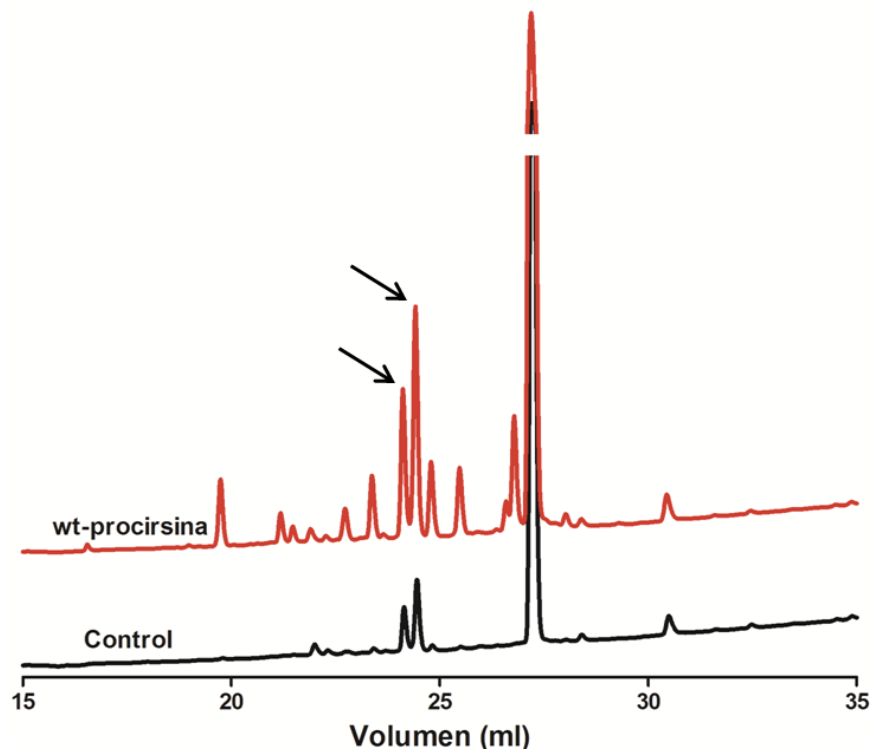


Fig. 22. Digestión de la cadena β de la insulina oxidada a pH 4,0. La especificidad de wt-procirsina recombinante se evaluó por hidrólisis de la cadena β de la insulina oxidada a pH 4,0 y 37 °C durante toda la noche. Los picos correspondientes a los productos mayoritarios de digestión (indicados en la figura mediante flechas negras), se recolectaron y se identificaron sus secuencias. El control se realizó por incubación de la cadena β de la insulina en ausencia de enzima.



Determinación de parámetros cinéticos

Se determinaron los parámetros cinéticos para la hidrólisis del sustrato MCA-Lys-Lys-Pro-Ala-Glu-Phe-Phe-Ala-Leu-Lys(DNP). Los datos cinéticos se describieron bien por la clásica curva hiperbólica consistente con el modelo de Michaelis-Menten (Fig. 23). Los valores determinados para el clivaje a pH 4,0 han sido $K_m = 7,13 \pm 1,30 \mu\text{M}$ y $\kappa_{\text{cat}} = 4,24 \pm 0,05 \text{ s}^{-1}$. La eficiencia catalítica para wt-procirsina recombinante resultó por tanto igual a $594 \text{ mM}^{-1} \text{ s}^{-1}$.

Cinética enzimática de la hidrólisis del sustrato MCA-KKPAEFFALK(DNP)

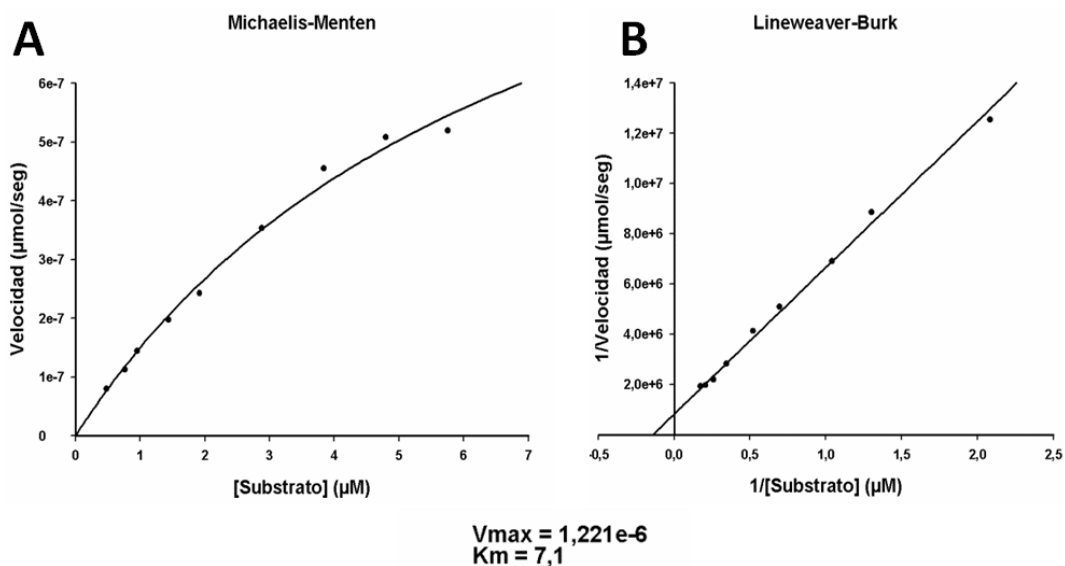


Fig. 23. Determinación de los parámetros cinéticos $V_{\text{máx}}$ y K_m de la hidrólisis del sustrato de PAs típicas por parte de wt-procirsina recombinante. Los ensayos de actividad para los estudios cinéticos se llevaron a cabo a pH 4,0 y 37 °C, usando como sustrato el derivado fluorogénico del péptido sintético MCA-Lys-Lys-Pro-Ala-Glu-Phe-Phe-Ala-Leu-Lys(DNP). **(A)** Representación de la velocidad inicial de la hidrólisis enzimática a diferentes concentraciones de sustrato. **(B)** Gráfico de Lineweaver-Burk (doble recíproca).



Ensayos de inhibición

Como se esperaba, la incubación con el inhibidor específico de PAs pepstatina A produjo una fuerte inhibición (98%) de la actividad de procirsina (Tabla 16) y está en concordancia con la ausencia de actividad observada para la proteína mutante de sitio activo procirsinaD103A.

También se evaluó el efecto de inhibidores específicos de otras clases de proteinasas y, de manera inesperada, Pefabloc y EDTA resultaron capaces de inhibir parcialmente la actividad de wt-procirsina recombinante. En tanto que la incubación con DTT así como con E-64 y amastatina no tuvieron mayor efecto sobre la actividad.

Inhibidor	Concentración (mM)	Actividad Residual (%) ^a
Pepstatina A	0,001	2
EDTA	0,005	61
E-64	0,01	88
Pefabloc	1	47
Amastatina	0,01	80
DTT	1	74

Tabla 16. Actividad de wt-procirsina recombinante luego de la incubación con inhibidores específicos de proteasas.

^(a) Las reacciones se llevaron a cabo con el sustrato sintético MCA-KKPAEFFALK(DNP) en concentración 192 μ M en buffer ácido acético–acetato de sodio (50 mM) de pH 4,0 a 37 °C, después de 6 min de incubación a temperatura ambiente de la enzima con cada inhibidor.

V. Clivaje de κ -caseína y actividad coagulante de leche

Una buena enzima coagulante de leche se caracteriza por una alta especificidad de corte del κ -caseína macropéptido, lo cual dispara la desestabilización de las micelas de caseína y, por tanto, induce la coagulación de la leche (Hyslop, 2003; Jacob et al., 2011). Como los extractos de flores de *C. vulgare* mostraron actividad coagulante de leche y dada la similitud entre procirsina y otras PAs típicas de plantas reconocidas como coagulantes de leche, en particular con la cyprosina de *C. cardunculus*, se ensayó el potencial de la wt-procirsina recombinante como agente coagulante de leche. El corte de la κ -caseína bovina purificada por parte de la enzima recombinante se evaluó mediante análisis por SDS-PAGE. Se observó un único producto de corte bien definido de 15 kDa, demostrando la alta especificidad de la procirsina recombinante hacia κ -caseína (Fig. 24 A). Un patrón de digestión similar sobre la κ -caseína ha sido obtenido por acción de la quimosina así como de otras PAs de plantas (Timotijević et al., 2006) y corresponde al corte del enlace peptídico Phe105-Met106, el cual es responsable de desencadenar la coagulación de la leche.

La actividad coagulante de leche de procirsina se confirmó posteriormente mediante la incubación de la enzima con una solución de leche magra a 37 °C durante toda la noche y la observación visual del coágulo formado. Como se señala en la Fig. 24 B, procirsina recombinante resultó capaz de coagular la leche, en tanto, esta actividad se inhibió totalmente en presencia de pepstatina A.

Hidrólisis de κ -caseína y actividad coagulante de leche

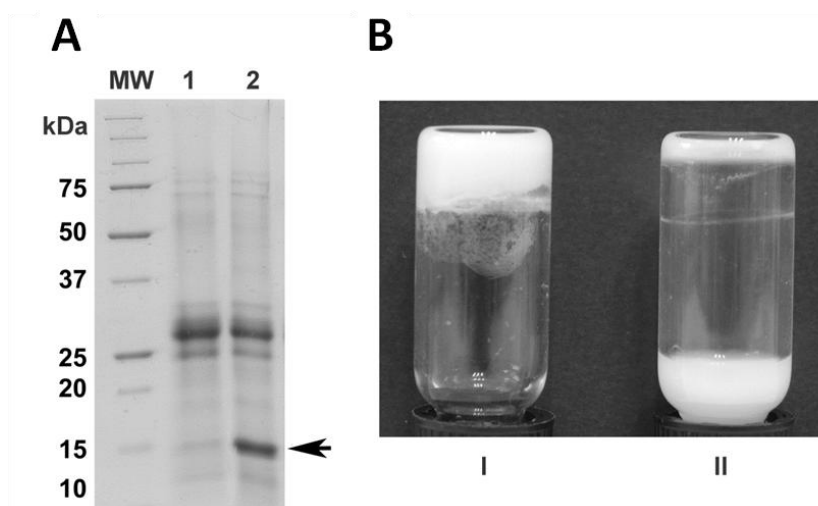


Fig. 24. Análisis electroforético de la hidrólisis de κ -caseína bovina y actividad coagulante de leche de wt-procirsina recombinante purificada. (A) κ -caseína bovina disuelta en buffer ácido acético-acetato de sodio 50 mM, pH 5,5 conteniendo NaCl 100 mM, se incubó a 37 °C durante toda la noche con wt-procirsina recombinante purificada. Los fragmentos de digestión se analizaron por SDS-PAGE con tinción por Coomassie Brilliant Blue: *Calle 1*, control de incubación de κ -caseína sin enzima presente; *Calle 2*, perfil de digestión de κ -caseína en presencia de wt-procirsina recombinante purificada. El principal producto de digestión está indicado con una flecha negra. (B) La actividad coagulante de leche de wt-procirsina recombinante se ensayó por incubación a 37 °C con leche descremada 12% (p/v) reconstituida en CaCl₂ 10 mM de pH 6,5. La formación del coágulo se evaluó mediante observación visual (*tubo I*). La actividad coagulante de leche se inhibió por la presencia de pepstatina A en la mezcla de reacción (*tubo II*), sin formación de coágulo visible.

VI. Caracterización bioinformática de wt-procirsina

Con el objetivo de predecir la estructura tridimensional de procirsina, se construyeron modelos estructurales mediante modelado por homología. La estructura de profitepsina se tomó como referencia, por tratarse de la proteína homóloga más próxima a procirsina según la búsqueda de similitud de secuencias realizada. A partir de ésta, se construyeron 100 modelos para procirsina utilizando el programa Modeller. El mejor modelo seleccionado fue

posteriormente evaluado con el potencial DOPE y el servidor Prosa II (score – 8,5) resultando en un modelo estructural de alta calidad. De acuerdo a este modelo estructural y al análisis evolutivo descrito en la sección *Análisis filogenético*, todos los residuos que participan en el sitio activo (Asp 36 y 223 siguiendo la numeración de profitepsina), así como aquellas Cys que intervienen en los puentes disulfuro, están conservados (Cys 49, 55, 214, 218, 257 y 294 en la proteína madura; 6, 100, 31, 72, 37 y 69 en el dominio PSI, una vez más la numeración se corresponde con la de la proteína profitepsina, Fig. 25).

Modelado de procirsina

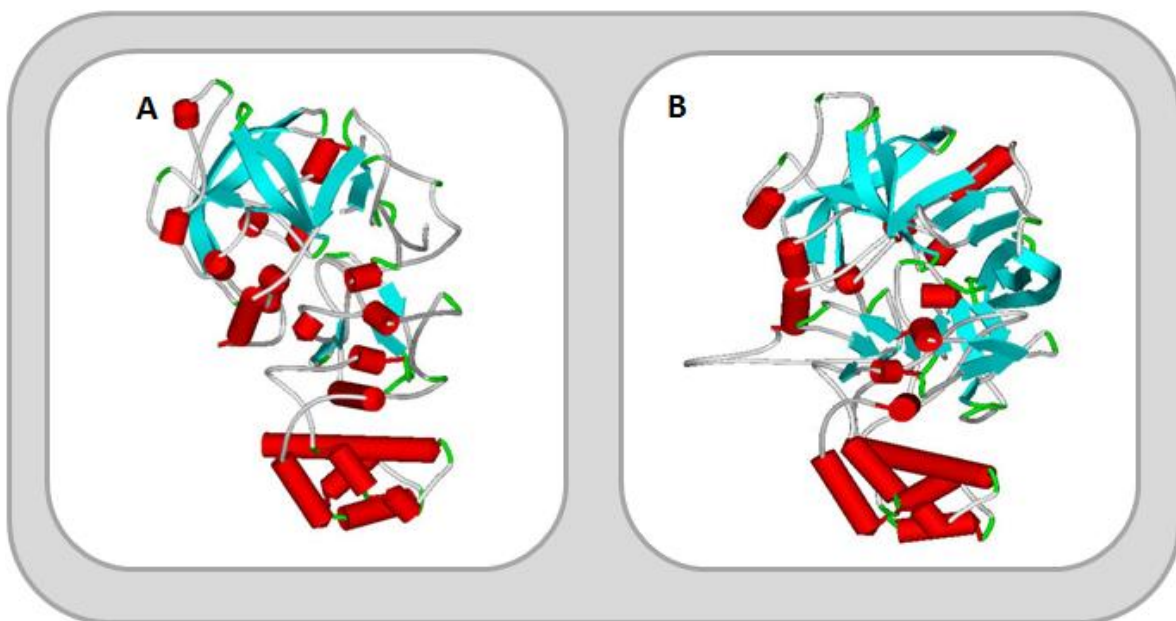


Fig. 25. Modelado de procirsina por homología. La secuencia de procirsina, sin el péptido señal putativo se modeló por técnicas de homología, usando estructuras cristalográficas conocidas relacionadas con procirsina. **(A)** Representación gráfica de la estructura de profitepsina aislada de *Hordeum vulgare* (código PDB: 1qdm). **(B)** Mejor modelo de procirsina obtenido.

Otros residuos de importancia para la función biológica se encuentran también conservados en procirsina al compararse con profitepsina. Estos residuos son descriptos como críticos para prevenir la activación del zimógeno. Como se reportó previamente (Kervinen et al., 1999), el segmento propéptido juega un rol decisivo en la interacción con la proteína madura y el bloqueo de los Asp catalíticos. Entre los residuos que participan en la interacción establecida entre el propéptido y la proteína madura, figuran una Arg (Arg7 en el propéptido de profitepsina) y el par Lys/Tyr (11 y 13 en la proteína madura).

El residuo Arg7 forma una unión iónica entre el propéptido y Glu171 y Asp178 en la proteína madura (Fig. 26). Estas interacciones, junto con otras uniones hidrofóbicas y puentes de hidrógeno, conectan el prosegmento de modo tal que los residuos Lys y Tyr en la proteína madura interactúan con el Asp en el sitio activo, inhibiendo la actividad de profitepsina. Es importante notar que todos estos residuos se encuentran conservados en procirsina haciendo difícil explicar la actividad encontrada para esta proteína a bajos valores de pH. Cabe destacar además, que estos residuos están conservados de modo evolutivo en todas las secuencias en el *cluster* que contiene a procirsina, con excepción de cardosina A y B (ID 75267434 y 7533856, Fig. 14), que no poseen el par Lys11/Tyr13. En la Fig. 27, se muestran las representaciones gráficas del modelo propuesto para procirsina y de la estructura de profitepsina indicando las interacciones recientemente mencionadas.

Alineamiento de las secuencias de los precursores de cirsina y fitepsina

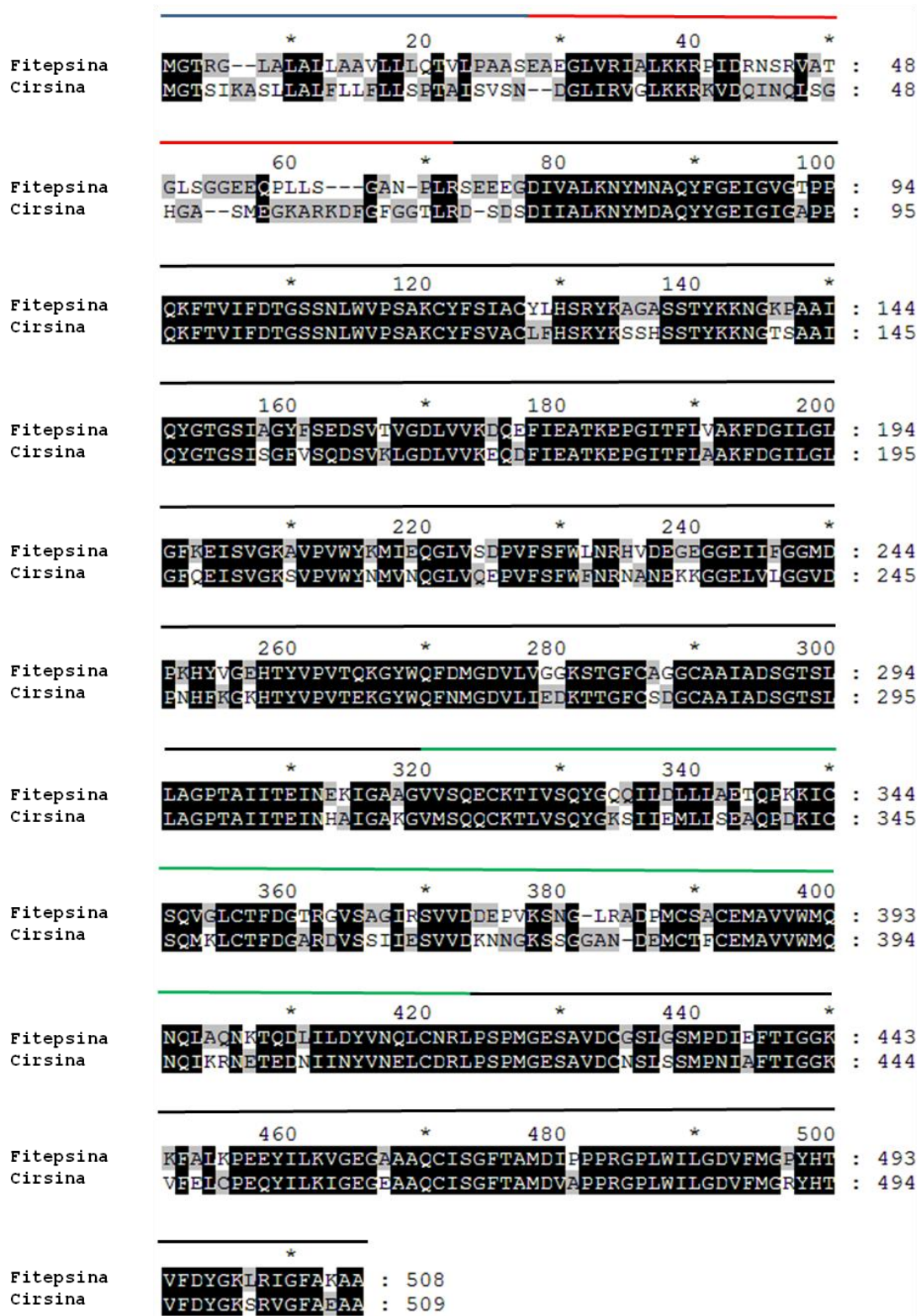


Fig. 26. Alineamiento de las secuencias aminoacídicas de procirsina y profitepsina (número de acceso: P42210). La secuencia pre se encuentra indicada por una línea azul, el prosegmento por una línea roja, la línea verde señala el dominio PSI mientras que la línea negra muestra la secuencia correspondiente a la proteína madura.

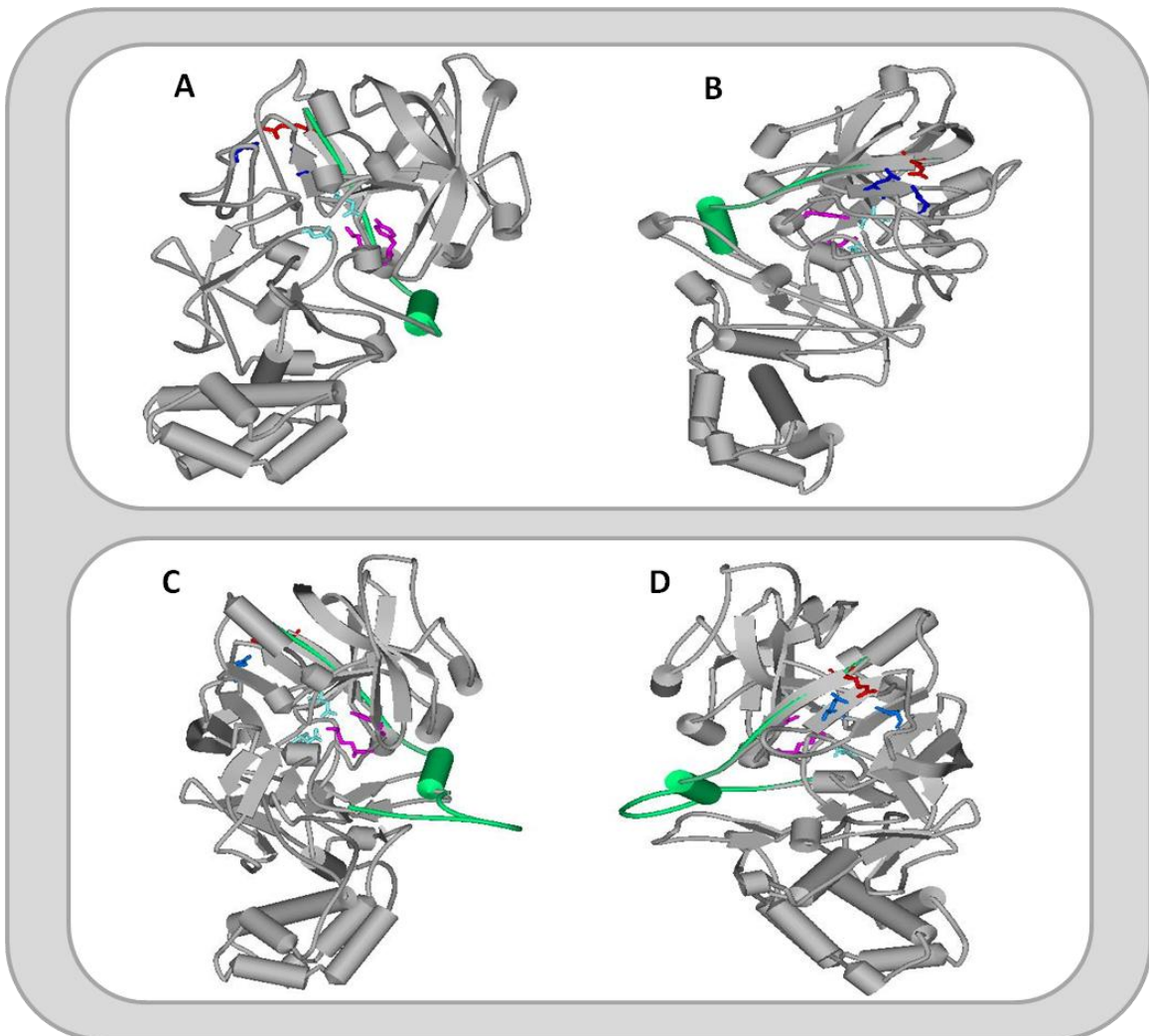


Fig. 27. Representación gráfica de las estructuras de profitepsina y de procirsina. Las figuras (A y B) corresponden a la estructura de profitepsina; las figuras (C y D) pertenecen a procirsina. El plegado principal de las proteínas está mostrado en gris y el prosegmento en verde. (A y C) Interacción de la Lys y Tyr (11/13 respectivamente, en profitepsina y 78/80 en procirsina) representadas en color rosa en las figuras, con los Asp del sitio catalítico, indicados en azul. (B y D) Interacción iónica establecida entre Arg7 del prosegmento (en rojo en ambas figuras) y los residuos Asp, en azul. Esta interacción está descrita en profitepsina, como una de las interacciones responsables de estabilizar el prosegmento en una conformación tal que produce la inhibición de la proteína.

Con el fin de explicar la actividad encontrada para procirsina a valores de pH ácidos, se estimó la variación de la carga en función del pH, empleando los propéptidos de procirsina y de profitepsina. Como se muestra en la Fig. 28, hay una variación significativa de carga alrededor del pH 4,0, donde procirsina mostró mayor actividad. Una gran carga positiva a pH ácido hallada en procirsina en comparación con la carga de profitepsina, podría alterar la correcta localización del propéptido necesaria para impedir la interacción del par Lys/Tyr con los Asp

ubicados en el sitio catalítico, y tornando activa de este modo a procirsina a bajos pH. A través de este análisis, se propone que a medida que el pH aumenta, la carga en el prosegmento resulta menor permitiendo la conformación adecuada del prosegmento para inhibir la proenzima.

Carga neta del prosegmento vs pH del medio

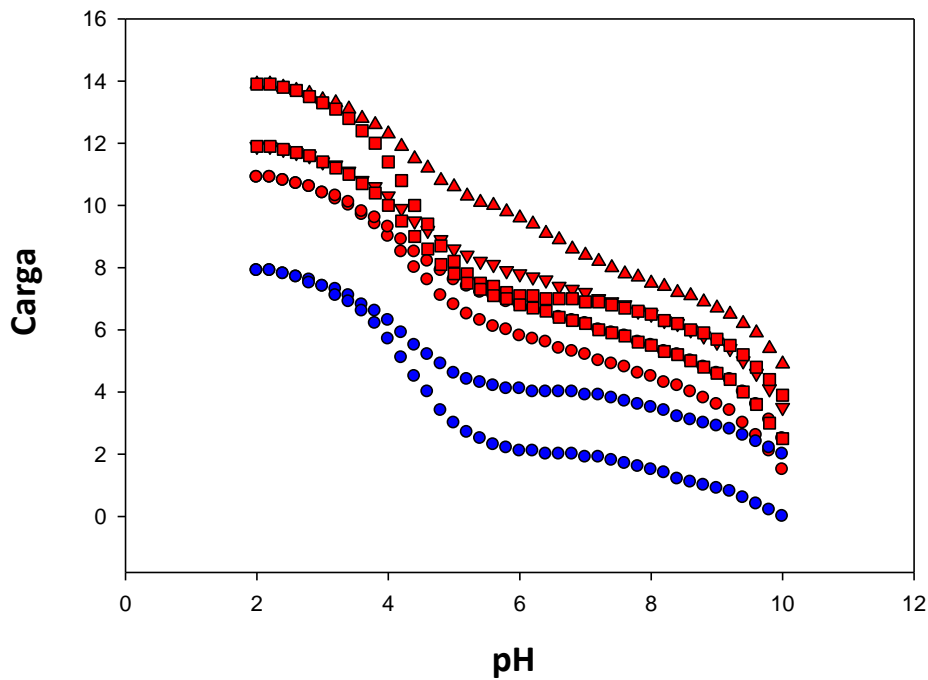


Fig. 28. Estimación de la carga del prosegmento en función del pH. En rojo se muestra el comportamiento del propéptido de procirsina y de sus proteínas más estrechamente relacionadas, compartiendo el mismo *cluster* evolutivo deducido a partir del análisis filogenético (códigos ID de las secuencias: procirsina, 556819, 1665867, 4589716, 75338567, 75267434). En azul se representa el comportamiento de profitepsina (ID 1168536) y de su proteína homóloga más próxima 5822248 (ver Fig. 14).

Discusión—Procirsina

Las PAs típicas de plantas son el grupo de enzimas que incluyen el dominio tipo saposina conocido como *plant specific insert* (PSI), en las secuencias de sus precursores. Aunque el análisis *in silico* ha demostrado que este grupo constituye la minoría de los genes de las PAs de *Arabidopsis thaliana* (Faro and Gal, 2005), sigue siendo el grupo mejor caracterizado entre las PAs de plantas. En este capítulo se reportó la identificación de una nueva PA típica—cirsina—que se expresa naturalmente en flores de *Cirsium vulgare* (Asteraceae). La amplificación del ADNc con cebadores conservados resultó en el aislamiento de al menos otras tres secuencias codificantes de PAs típicas, sugiriendo claramente la presencia en *C. vulgare*, al igual que lo reportado para *C. cardunculus* (Pimentel et al., 2007), de varias enzimas estrechamente relacionadas de este grupo. Con el objetivo de comenzar a estudiar estos nuevos miembros del grupo de PAs típicas de plantas, en esta oportunidad se clonó, expresó, purificó y caracterizó procirsina. La secuencia aminoacídica derivada del ADNc de cirsina completa (preprocirsina), muestra la usual organización de los precursores de PAs típicas de plantas. La similitud global entre la secuencia de aminoácidos de cirsina con las de otras PAs típicas de plantas, junto con la presencia de los motivos característicos de PAs, clasifican a cirsina como un nuevo miembro de esta familia de enzimas.

Resulta llamativa la semejanza de procirsina con procyprosina B de *C. cardunculus* (Cordeiro et al., 1994) que alcanza un 95% de identidad de secuencia. Esto sugiere que cirsina y cyprosina están estrechamente relacionadas previendo una gran similitud entre las estructuras de estas enzimas. Al compararse con cyprosina B más detalladamente, se advierte que la similitud entre las secuencias es más baja en el dominio PSI (93%). Esta observación está de acuerdo con una diversidad más alta entre las secuencias de PSI que podría deberse al resultado de adaptaciones evolutivas (Cordeiro et al., 1994) vinculadas al rol específico del PSI en cada enzima en particular—esto es, en el direccionamiento vacuolar, en el plegamiento correcto o en ambos (Törmäkangas et al., 2001; Brodelius et al., 2005; Terauchi et al., 2006; Duarte et al., 2008).

Las PAs tienden a unir en sus sitios activos 6 a 10 residuos de sus sustratos polipeptídicos, los cuales sufren el ataque nucleofílico de una molécula de agua vinculada a los dos residuos ácido

aspártico catalíticos de la proteasa. Todas las peptidasas de la familia A1 son endopeptidasas y la gran mayoría de ellas son activas principalmente a pH ácido, sin requerimiento de cofactores y son inhibidas por pepstatina A. De manera característica, estas proteasas escinden con preferencia enlaces formados entre residuos con cadenas laterales grandes e hidrofóbicas, aunque varios miembros de esta familia de peptidasas presentan especificidades más restringidas que permiten las funciones de procesamiento de la proteína (Rawlings et al., 2012).

En contraste con otras PAs típicas de plantas que resultaron ser inactivas cuando se expresaron como precursores en sistemas procariotas (Asakura et al., 2000; Castanheira et al., 2005), la expresión de procirsina en *E. coli* condujo a la purificación de una proteína de simple cadena (wt-procirsina) que resultó activa frente al sustrato sintético MCA-Lys-Lys-Pro-Ala-Glu-Phe-Phe-Ala-Leu-Lys(DNP). Este sustrato contiene un enlace entre dos residuos aromáticos y es comúnmente usado en el estudio de PAs típicas.

La proteína recombinante wt-procirsina mostró las mismas características enzimáticas que otras PAs típicas como cyprosina (White et al., 1999) y cardosina A (Castanheira et al., 2005) recombinantes y silpepsina nativa (Vairo Cavalli, 2005; Vairo Cavalli et al., 2009): máxima actividad a pH ácido, inhibición por pepstatina A, dependencia estricta de dos residuos Asp catalíticos para la actividad, preferencia de corte de enlaces peptídicos formados entre aminoácidos con cadenas laterales grandes e hidrofóbicas (Simões y Faro, 2004) y actividad coagulante de leche.

Cuando se empleó la cadena β de la insulina oxidada como sustrato para wt-procirsina recombinante, los fragmentos de digestión mayoritarios resultaron como producto de la ruptura del enlace Leu15-Tyr16. Esta unión es susceptible a la acción de la enzima gástrica pepsina y de fitepsina de cebada (Athauda and Takahashi, 2002; Kervinen et al., 2004), las cuales son consideradas PAs modelo dentro de la familia A1 en mamíferos y plantas, respectivamente. También mostraron capacidad de escindir dicho enlace cardosina A—Leu15-Tyr16, Leu17-Val18 y Phe24-Phe25 (Veríssimo et al., 1995)—y cardosina B, aunque esta última reveló una especificidad ligeramente más amplia que la primera cortando además los enlaces Glu13-Ala14, Ala14-Leu15, Leu15-Tyr16, Leu17-Val18, Phe24-Phe25 y Phe25-Tyr26.

Un patrón de digestión similar al obtenido por acción de procirsina recombinante sobre κ -caseína ha sido observado por acción de la quimosina así como de otras PAs de plantas (cardosina A, cyprosina B, PAs de trigo) sobre este sustrato (Cordeiro, Jakob, Puhán, Pais, &

Brodelius, 1992; Timotijević, Radović, & Maksimović, 2006; Veríssimo et al., 1995) y corresponde al corte del enlace peptídico Phe105-Met106 responsable de desencadenar el proceso de coagulación de la leche (Hyslop, 2003; Jacob et al., 2011).

Como muchas PAs aisladas de flores de varias especies de la familia Asteraceae, procirsina mostró actividad coagulante de leche (Heimgartner et al., 1990; Tamer, 1993; Veríssimo et al., 1995; Llorente et al., 2004; Raposo and Domingos, 2008; Reis and Malcata, 2011). La actividad de procirsina hacia la leche bovina descremada resultó moderada mientras que de acuerdo al patrón de hidrólisis sobre κ -caseína mencionado, esta enzima exhibe también una especificidad de corte significativa. Ambas propiedades son deseables en la manufactura de quesos dado que es sabido que una actividad proteolítica excesiva, usualmente observada en extractos que contienen endopeptidasas cisteínicas (por ejemplo, extractos de *Ficus carica*, *Carica papaya*, *Ananas sativus*), produce características organolépticas indeseables (Walstra et al., 1999).

La actividad catalítica en una proenzima no es completamente inesperada ya que algunos precursores de PAs de mamífero de la familia A1, como BACE y renina, muestran actividad enzimática significativa (Pitarresi et al., 1992; Ermolieff et al., 2000; Shi et al., 2001). Sin embargo, este fenómeno no ha sido aún claramente descrito en el caso de las PAs típicas de plantas.

Hasta el momento, se han descrito tres mecanismos diferentes por los cuales los zimógenos aspárticos de diversos orígenes inhiben sus propias actividades catalíticas. Se ha sugerido que el propéptido de procardosina A, como muchos otros prosegmentos de zimógenos de PAs, tienen un rol en su propio mecanismo de inactivación (Castanheira et al., 2005). La mayoría de las PAs de plantas contienen una secuencia Lys/Tyr en posición equivalente a la secuencia Lys11/Tyr13 de profitepsina (*Hordeum vulgare*) y Lys36p/Tyr37p del pepsinógeno (donde p indica prosegmento). Estudios previos han descrito a estos aminoácidos como responsables de las interacciones electrostáticas con los residuos aspárticos catalíticos que contribuyen a la inhibición de los mismos en la forma inactiva de la enzima al bloquear la hendidura de unión al sustrato.

A partir de la obtención de la estructura cristalográfica de profitepsina en el año 1999 (Kervinen et al., 1999), se sugirió un modelo de inactivación potencialmente común entre los zimógenos de PAs típicas de plantas que contienen en sus prosegmentos los residuos Lys/Tyr en posiciones equivalentes a las secuencias Lys11/Tyr13 de profitepsina. En dicho modelo, el

propéptido se envuelve alrededor de la fitepsina madura y, junto con los primeros residuos de la cadena N-terminal y la horquilla β conteniendo los residuos 75-88 de la enzima madura—porción conocida como “*flap*”—, bloquea estéricamente el sitio activo, conduciendo a su total inactivación. El anclaje de estas porciones de la proteína se produce mediante la interacción de Lys11 con los aspartatos del sitio activo, mientras que Tyr13 e Ile7 ocupan los sitios de unión del sustrato S1 y S1'. Además, los grupos hidroxilos de los residuos Tyr13 y Ser39 forman puentes de hidrógeno con el Asp36 catalítico. El residuo Tyr80, análogo al Tyr75 en el pepsinógeno (Sieleki et al., 1991; Moore et al., 1995), forma un puente de hidrógeno con $N\epsilon_1$ de Trp43 (Trp39 en numeración de la pepsina). En conjunto, el propéptido con los 13 residuos del N-terminal de la enzima madura y el *flap* cubren completamente el sitio activo. En particular, el residuo Arg7 del propéptido está descrito que forma una interacción iónica con los residuos Glu171 y Asp178 de la proteína madura, y de manera conjunta con otros enlaces hidrofóbicos y puentes de hidrógeno posiciona al propéptido de modo tal que el par Lys11/Tyr13 del extremo N-terminal y no del prosegmento, interactúa con el sitio activo produciendo la inhibición de la actividad de profitepsina. Por tanto, la posición y conformación de los primeros 54 residuos de la profitepsina en relación al resto de la proteína, están ancladas por dos interacciones iónicas, por puentes de hidrógeno y por contactos hidrofóbicos, incluyendo interacciones en el bolsillo de unión al sustrato S1 y S1'.

El rol del anclaje interno a los aspartatos catalíticos, asumido en las PAs de animales por el residuo Lys36p del propéptido es cumplido por la Lys11 perteneciente a una porción de la proteína madura de profitepsina. En base a la homología de secuencias, es muy probable que lo descrito hasta ahora sea el caso general para las PAs de plantas, aunque algunas de ellas no presentan una Lys homóloga ni en su prosegmento ni en la proteína madura, como por ejemplo dos isoenzimas de arroz (número de acceso GenBank D12777) y cardosina A y B (Simões and Faro, 2004).

Resulta interesante a su vez, que el prosegmento de procirsina comparte un 40% de identidad con el de profitepsina y, como se describió en los resultados de este capítulo, tanto la Lys inhibitoria como otros residuos críticos a pesar de estar conservados en procirsina, no suprimen la actividad catalítica. Con el objetivo de encontrar posibles diferencias que expliquen la actividad de procirsina recombinante, en este trabajo se realizó el análisis filogenético y el modelado estructural de la molécula. Asimismo, se estimó la variación de la carga neta de los propéptidos de procirsina y profitepsina como una función del pH. El análisis bioinformático mostró que la procirsina comparte un “*cluster*” o grupo con PAs de diversos

organismos en el cual el homólogo más cercano es cyprosina de *C. cardunculus* con un 98% de similitud de secuencia. De acuerdo al modelo estructural obtenido y al análisis evolutivo, no sólo todos los residuos descritos como importantes para la función biológica están conservados en procirsina en comparación con las secuencias restantes del *cluster*, sino que también lo están las posiciones Arg7p, Lys11 y Tyr13. Sin embargo, la gran carga positiva predicha para el prosegmento de procirsina a valores de pH ácidos al compararse con la correspondiente a la región pro de profitepsina, podría alterar la correcta localización del propéptido impidiendo la interacción del par Lys/Tyr con los Asp catalíticos y haciendo activa así a procirsina en estas condiciones. A medida que el pH aumenta, la carga del prosegmento disminuye permitiendo la correcta conformación para llevar a cabo la inhibición del precursor.

Si bien el máximo de la actividad endopeptidásica se observó a pH 4,0, la dependencia de la actividad de procirsina recombinante con el pH no muestra el típico perfil acampanado característico de las PAs maduras, lo que indicaría además que el anclado del pro-segmento en el sitio activo probablemente dependa de una red de interacciones electrostáticas que sólo se perturba en intervalos muy limitados de pH. A fin de designar los residuos que controlan la actividad reportada de procirsina se requiere de estudios posteriores. No obstante, estos resultados proporcionan evidencias adicionales sobre la posibilidad de que en los precursores de PAs típicas de plantas existan diversos arreglos estructurales—desde diferentes mecanismos de inactivación (profitepsina vs procardosina) (Kervinen et al., 1999, Castanheira et al., 2005), hasta formas zimogénicas catalíticamente activas (procirsina)—y ésto podría estar a su vez correlacionado con los patrones de diversidad de PAs típicas de plantas y de sus funciones biológicas.

La activación de los zimógenos de PAs hacia enzimas maduras requiere de la remoción del propéptido. Esta activación ocurre por una reacción intramolecular (autoactivación), una reacción intermolecular (bimolecular) o ambas (Varón et al., 2006). Notablemente, cuando se ensayó la activación autocatalítica *in vitro* de procirsina recombinante a diferentes valores de pH, ninguna conversión definida ni incremento claro en la actividad proteolítica fueron observados, sugiriendo que aunque la forma del precursor de cirsina resultó catalíticamente activa, fue incapaz de disparar la autoactivación o procesamiento. Esta observación se contrapone con lo expuesto para otras PAs típicas de plantas obtenidas en forma recombinante como profitepsina (Glathe et al., 1998), procardosina A (Castanheira et al., 2005) y procenprosina (Domingos et al., 2000) a las cuales se les ha podido asignar los múltiples pasos que llevan a la activación *in vitro* de las enzimas por simple acidificación del medio.

Aunque no puede excluirse que la procirsina requiera condiciones muy específicas para iniciar un evento proteolítico intramolecular, los resultados obtenidos aquí parecen indicar que la procirsina podría sufrir un mecanismo de procesamiento diferente, requiriendo muy probablemente la acción de proteinasas de la célula huésped para alcanzar la conversión completa hacia la forma activa final de la enzima. De hecho, un mecanismo similar fue propuesto para BACE, que también mostró ser activa en la forma precursora e incapaz de autoactivarse (Ermolieff et al., 2000). La expresión del mutante de procirsina sin el dominio PSI condujo a una proteína inactiva, indicando un papel crítico de este segmento interno para el plegamiento de cirsina. Datos similares fueron reportados para la expresión de procyprosina sin el PSI cuando fue expresada en *Pichia pastoris* (White et al., 1999) siendo consistente con la alta similitud en la secuencia entre ambas enzimas. Una vez más, lo visto para procirsina no concuerda con los resultados descritos para la forma de simple cadena de procardosina A sin PSI, que mostró autoactivación *in vitro* por simple caída del pH y exhibió características similares a las de la enzima natural (Castanheira et al., 2005).

En conjunto, estos resultados refuerzan la idea que los precursores de las PAs típicas de planta asumen diversos arreglos estructurales que contribuyen muy probablemente a regular la actividad enzimática, enfatizando la importancia de avanzar en la dilucidación de sus mecanismos de activación/procesamiento.



PSI



Introducción—Plant Specific Insert (PSI)

Las proteasas aspárticas de plantas clasificadas como típicas, contienen una secuencia adicional de aproximadamente 100 aminoácidos, denominada inserto específico de planta (PSI). A pesar de la ausencia de homología con sus equivalentes de origen no vegetal (bacterias, hongos, virus, mamíferos), el PSI posee una elevada similitud estructural con las saposinas, proteínas activadoras de hidrolasas lisomales de esfingolípidos en células de mamíferos. Comparte con las saposinas, la estructura compuesta de cinco α -hélices anfipáticas plegadas en un dominio globular compacto estabilizado por tres puentes disulfuro, la existencia de un sitio de glicosilación conservado y la presencia de prolina y tirosina. Son varias las proteínas que presentan este plegamiento, y debido a esta característica común las mismas constituyen la familia de proteínas similares a saposinas, usualmente denominadas proteínas SAPLIPs (Bruhn, 2005).

Sin embargo, los PSIs han sido llamados “*swaposinas*” (del inglés *swap*: intercambiar) dado que comprometen la porción C-terminal de un tipo de saposina seguida de una porción N-terminal de otro tipo de saposina, unidas por una secuencia de aproximadamente 20 aminoácidos (Brodelius et al., 2005), llamada tentativamente enlazador de saposinas (*S-linker*). Esto es, cada extremo de un determinado PSI está intercambiado con el extremo opuesto de una de dos saposinas diferentes. A pesar de esta diferencia en la localización de los extremos, la estructura global tipo saposina no se encuentra alterada en los PSIs. De hecho, estudios cristalográficos del dominio *swaposina* presente en una PA de papa (StAP-PSI), revelaron para el mismo una configuración global en forma de V similar a la estructura abierta de la saposina C humana (Fig. 29, Bryksa et al., 2011). Además de la estructura cristalina del StAP-PSI, la otra estructura cristalina resuelta en la que se encuentra un dominio PSI, corresponde a la profitepsina (*Hordeum vulgare*). Resulta interesante que dentro de esta proteína, el PSI mostró una estructura terciaria consistente con el plegado tipo saposina “cerrado” (Kervinen et al., 1999).

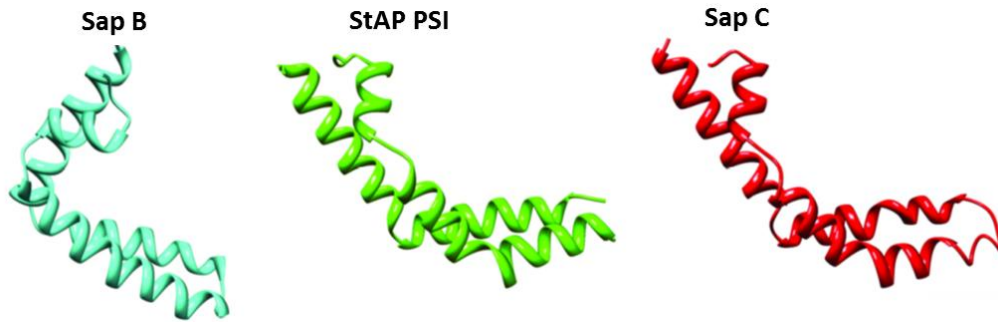


Fig. 29. Comparación del plegado en *boomerang* “abierto” de saposina B (1N69), saposina C (2QYP) y StAP PSI (3RFI). (Figura tomada de Bryksa 2011)

Debido a ello, Bryksa y colaboradores (Bryksa et al., 2011) propusieron que los PSIs de las PAs de plantas, pasan posiblemente de un plegado cerrado a uno abierto al ser liberados de la molécula de la proteasa a la cual pertenecen. A su vez, los autores relacionaron este cambio conformacional propuesto, con un aumento en la interacción con bicapas lipídicas que conduce a la dimerización de los dominios PSI, provocando la fusión de membranas vecinas. Dado que ambas formas, abierta y cerrada, de las SAPLIPs contienen los mismos puentes disulfuro (Bruhn, 2005), el rol primario de las cisteínas del PSI probablemente no sea mantener la estructura global, sino conferir rigidez (Fig. 30).

95

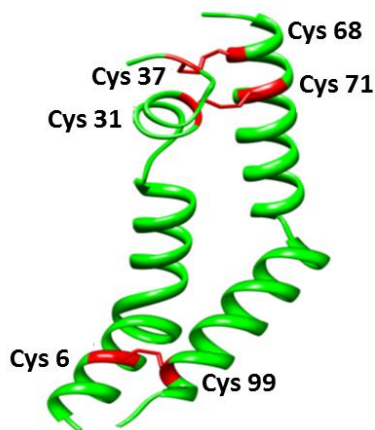


Fig. 30. Localización de los puentes disulfuro (indicados en rojo en la estructura) en StAP PSI. (Figura tomada de Bryksa 2011)

En cuanto a la función del PSI en la planta, no está aún completamente esclarecida. Se han asignado diversos roles a este dominio y algunos autores sugieren que la función del mismo difiere de acuerdo a su origen (Terauchi et al., 2006). Sin embargo, es posible que, al igual que los miembros de la familia SAPLIP, el PSI sea responsable de una interacción directa con la membrana, tal como se desprende de la caracterización estructural del PSI de profitepsina que revela una posible región de unión a membranas incluyendo al PSI y una región adyacente de

la enzima madura (Kervinen et al., 1999). Por otro lado, los efectos sobre membranas caracterizados mediante microscopía de fuerza atómica y dispersión de luz, indicaron que el dominio PSI de la PA de papa, StAP-PSI, posee capacidad de desestabilizar bicapas lipídicas así como también actividad fusogénica. Más aún, se determinó que el mecanismo de dicha interacción con membranas sigue una cinética que se ajusta a la curva de Michaelis-Menten mostrando dependencia con el pH (Bryksa et al., 2011). Años atrás, Egas y colaboradores (2000) estudiaron el PSI de procardosina A y propusieron que el mismo es un dominio desestabilizador de membranas, al encontrar que interactúa con vesículas induciendo la liberación de sus contenidos. Posteriormente, Simões y Faro (2004), basándose en estos estudios, sugirieron que los precursores de PAs son moléculas bifuncionales conteniendo un dominio desestabilizador de membranas (PSI) en adición a su dominio catalítico. Pereira y colaboradores (2008) sumaron más evidencias al debate detectando en las paredes celulares de semillas de cardos, la presencia de precursores de cardosina A que contienen al dominio PSI en su secuencia. Los autores argumentaron que este hecho podría estar implicado en los eventos de reorganización de membranas que ocurren en los tejidos en crecimiento activo (toma de agua, liberación de solutos, aparición de radícula y elongación, entre otros). En otro abordaje, Brodelius y colaboradores (2005) propusieron un rol para el PSI en la biogénesis de membranas en tejidos con altas velocidades de crecimiento, donde se espera un recambio elevado de lípidos.

Además, se ha vinculado el PSI a la función de defensa de las plantas (Egas et al., 2000; Muñoz et al., 2010), al correcto plegamiento del precursor de PAs, y al direccionamiento de los precursores de PAs heterodiméricas durante el transporte intracelular a través del Golgi hacia vacuolas, de un modo similar a la participación de saposinas en el co-transporte de procatepsina D al lisosoma (Runeberg-Roos et al., 1991; Ramalho-Santos et al., 1998; da Costa et al., 2010). En relación a esta última hipótesis, si bien tanto cardosina A como cardosina B, proteasas aspárticas de *C. cardunculus*, entran a la vía de secreción debido a la presencia de un péptido señal responsable de su translocación hacia el RE, mientras que cardosina A se acumula en vacuola (Runeberg-Roos et al., 1994), cardosina B es transportada fuera de la célula, hacia la pared celular y la matriz extracelular en los pistilos (Vieira et al., 2001). Los mecanismos que determinan esta diferente localización continúan siendo materia de discusión. Sin embargo, es de notar que aunque los precursores de ambas cardosinas poseen inserto PSI con elevada similitud entre sus secuencias (75%), la presencia de un motivo de glicosilación en el dominio PSI de procardosina B y ausente en procardosina A, imparte una

diferencia sustancial entre los dos precursores pudiendo estar relacionada con el destino final de la proteasa madura.

Como consecuencia de la actividad membranolítica de las proteínas similares a saposinas, se ha reportado para el PSI de la proteasa aspártica monomérica de papa, StAP, propiedades antifúngicas frente a patógenos de *Solanum tuberosum*, siendo capaz de interactuar con la superficie de los esporos causando la permeabilización de la membrana plasmática y la muerte de éstos, en un modo dosis-dependiente. Este mismo dominio al permeabilizar membranas plasmáticas, también resultó tóxico para patógenos humanos como *S. aureus*, *B. cereus* y *E. coli*, así como frente a linfocitos T tumorales en forma selectiva. Otros efectos evidenciados para el PSI de StAP debido a la acción permeabilizante de membranas, han sido las actividades hemolítica y la espermicida (Mendieta et al., 2006, 2010; Cesari et al., 2007; Muñoz et al., 2010).

En el presente capítulo se reporta la optimización de la expresión del dominio PSI del precursor de cirsina de *C. vulgare* en el sistema de levaduras *Kluyveromyces lactis*, se analiza el efecto de la glicosilación en la secreción de la proteína recombinante en dicho sistema y se determina la actividad antibacteriana y antifúngica del PSI obtenido.

Materiales y métodos—Plant Specific Insert (PSI)

Obtención de PSI en *Kluyveromyces lactis*

Las levaduras se han convertido en un buen sistema de expresión de proteínas heterólogas por su rápido crecimiento y la posibilidad de obtener elevadas densidades ópticas a bajo costo. Por otro lado, establecen una combinación entre la manipulación genética encontrada en bacterias con un ambiente eucariota, que permite modificaciones postraduccionales como procesamiento proteolítico, plegamiento, formación de enlaces disulfuro y glicosilación, ausentes en un sistema procarionta en el cual frecuentemente se generan proteínas eucariotas con plegamientos no nativos, muchas veces insolubles e inactivas. Las levaduras por su parte, no pueden realizar algunas modificaciones postraduccionales más complejas como ser prolilhidroxilación, amidación y algún tipo de glicosilación y fosforilación. Estas modificaciones sí pueden ser llevadas a cabo por sistemas de expresión más complejos como son células de mamíferos y las líneas de insectos infectados por baculovirus. Sin embargo, estos sistemas resultan ser más costosos, presentan un rendimiento de expresión menor y demandan más tiempo y esfuerzo.

La levadura *Kluyveromyces lactis* (New England Biolabs) es un organismo GRAS que permite la expresión de proteínas heterólogas tanto en forma intracelular como extracelular. La secreción de proteínas es el abordaje más común en este sistema y se logra mediante el clonado del gen de interés en el vector pKLAC1 *downstream* al dominio α -mating factor (α -MF), generando la expresión de una proteína de fusión. El dominio α -MF dirige la proteína de fusión para ser eficientemente transportada a través de la vía secretoria de la levadura. La proteína de fusión sufre un procesamiento secuencial que comienza con la escisión de la secuencia señal por una peptidasa (*signal peptidase*) en el retículo endoplasmático (ER) y se completa en el *trans*-Golgi por acción de la proteasa Kex-2, resultando en la secreción de una forma nativa de la proteína de interés hacia el medio de cultivo (Fig. 31).

La expresión en el sistema de *K. lactis* es conducido por una forma mutante del promotor fuerte LAC4 ($P_{LAC4-PBI}$). Para lograr la expresión en *K. lactis*, el vector pKLAC1 conteniendo el gen

de interés clonado es linealizado por *SacII* o *BstXI* para producir un casete de expresión que puede integrarse en el locus LAC4 del genoma de la levadura por recombinación homóloga. El vector pKLAC1 porta el gen de una acetamidasa (*amdS*) que selecciona las levaduras que han integrado el casete de expresión en su genoma al permitirles crecer en un medio mínimo libre de nitrógeno, conteniendo acetamida como única fuente de este elemento. Una ventaja de este método de selección es que enriquece las poblaciones de células transformadas que han integrado copias múltiples del casete de expresión en tándem y que por tanto, producen más cantidad de proteína recombinante.

Finalmente, en cuanto a la cepa *K. lactis* GG799 usada, es una cepa aislada industrialmente que no tiene auxotrofías, crece rápidamente hasta un valor de densidad celular alta, y secreta eficientemente las proteínas heterólogas.

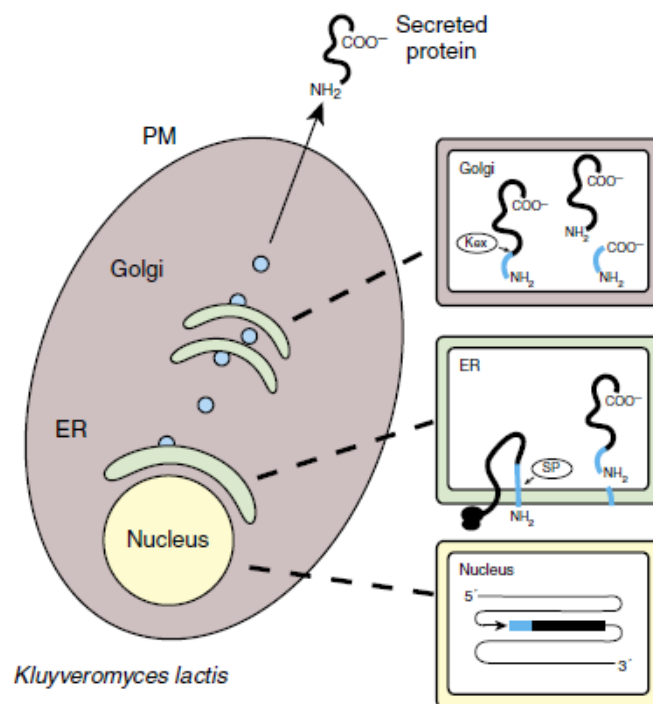


Fig. 31. Proceso de secreción de proteínas heterólogas en el sistema *K. lactis*. El ADN que codifica para una proteína recombinante de interés (negro) fusionada al dominio α -MF (azul) es integrado al genoma de la levadura. La expresión es conducida por el promotor $P_{LAC4-PBI}$ (flecha). Tras la expresión de la proteína de fusión, un péptido señal en la secuencia del α -MF dirige la translocación de la proteína hacia el lumen del retículo endoplasmático (ER) y es removido por la peptidasa señal (SP). Las vesículas secretorias (círculos) transportan la proteína de fusión al Golgi, donde la endoproteasa Kex-2 cliva el prodominio α -MF, liberando una forma madura de la proteína deseada. La proteína de interés es transportada entonces vía vesículas hacia la membrana plasmática (PM) donde es secretada de la célula (Extraído del manual “*K. lactis* Protein Expression Kit—Recombinant Protein Expression in Yeast” de New England Biolabs).

I. Obtención de construcciones para expresión de PSI en *K.lactis*

En este sistema, la secreción de la proteína recombinante fuera de la célula con su N-terminal nativo se consigue de acuerdo a las instrucciones del fabricante, por clonación del gen de interés en el único sitio *XhoI* del vector pKLAC1 *downstream* al dominio α -mating factor (α -MF) y a la secuencia que codifica para el sitio de corte de la proteasa Kex-2 (Fig. 32). No obstante, en este trabajo de tesis se generó además una construcción extra, en la cual la secuencia que codifica para el PSI se clonó inmediatamente *downstream* al péptido señal, sin el pro-dominio α -MF. Por otro lado, debido a que en *K. lactis* son posibles las modificaciones post-traduccionales, para este sistema se armaron las construcciones (con y sin α -MF) que contuvieron la mutación del sitio putativo de glicosilación alterando el aminoácido Asn por Ser en posición 86 (N86S). Se añadió además, una cola de seis histidinas en el extremo C-terminal de las proteínas a fin de ser inmunodetectadas y purificadas mediante cromatografía de afinidad.

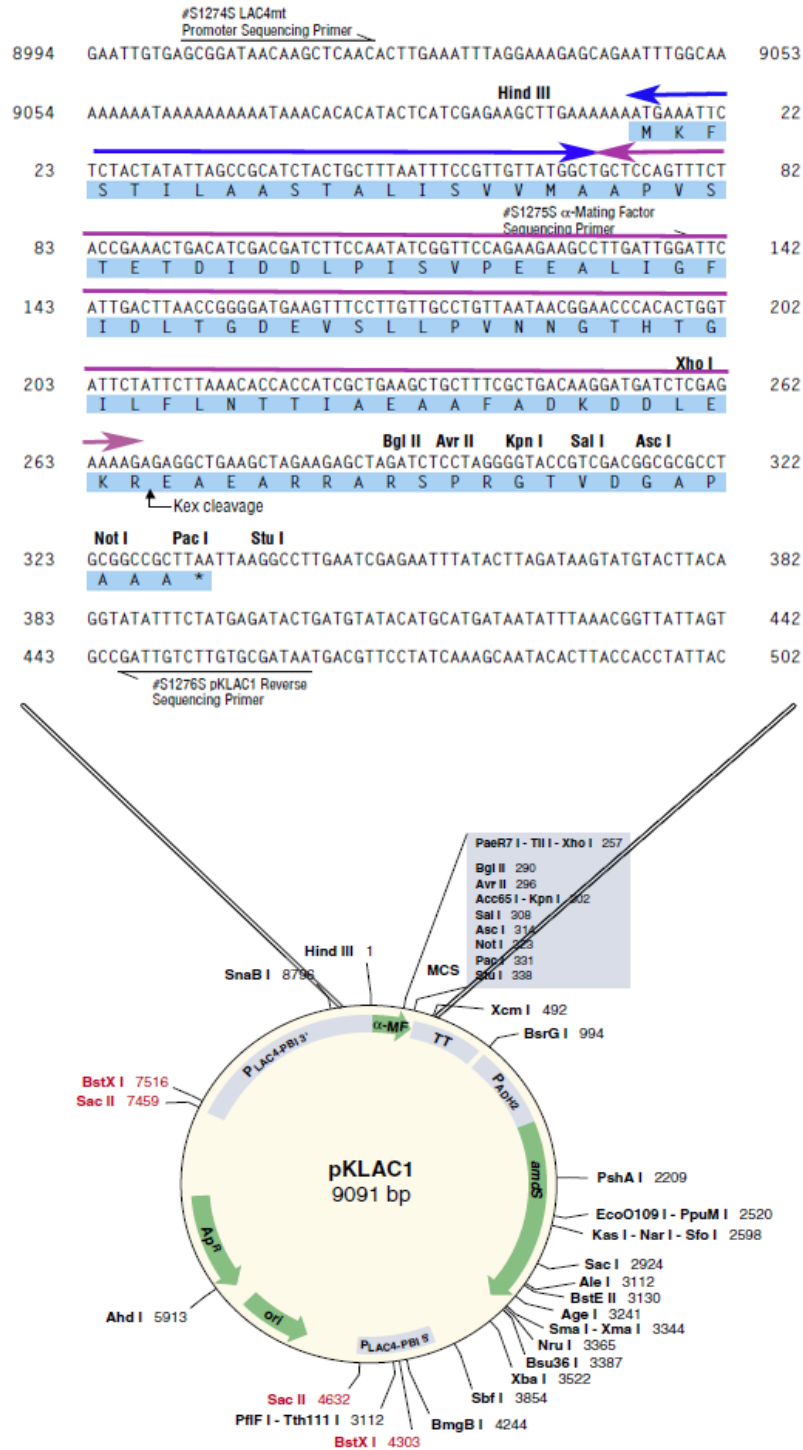


Fig. 32. Representación esquemática del vector pKLAC1. En el esquema se muestra la secuencia de la región correspondiente al sitio de clonado *XhoI* *downstream* al prodominio (círculo negro). La flecha azul indica la secuencia del péptido señal y la púrpura la correspondiente al prodominio α -MF, incluyendo el sitio *XhoI* y el sitio de corte para la enzima Kex-2. (Esquema adaptado del manual *K. lactis* Protein Expression Kit—Recombinant Protein Expression in Yeast de New England Biolabs).



Construcción pKLAC- α -MF-PSI(His)₆ y pKLAC- α -MF-PSIN86S(His)₆

El armado del plásmido pKLAC- α -MF-PSI(His)₆ se llevó a cabo clonando primero, en pGEM-T Easy, el producto de amplificación obtenido con los cebadores Xho_Kex_PSI_F y BglII_PSI(His)₆_stop_R (Tabla 17). Luego a partir de este plásmido, se subclonó el inserto PSI(His)₆ en el vector pKLAC1 entre los sitios *XhoI* y *BglII* del mismo. Los pasos que se siguieron hasta este punto son los ya descriptos en el apartado correspondiente a la obtención de los plásmidos pET-procirisina, pET-procirisinaD103A y pET-procirisina Δ PSI de *Materiales y Métodos—Procirsina*.

La reacción de PCR para amplificar la región del PSI e incorporar la cola de polihistidinas y los sitios de restricción necesarios, se basó en la *Técnica n°2* de *Materiales y Métodos Generales* empleando como ADN molde la construcción pET-procirisina previamente obtenida (*Materiales y Métodos—Procirsina*) y 1,25 U de polimerasa Taq Polimerase (GE Healthcare Life Sciences). El segmento II en el programa del termociclador consistió en 32 ciclos, cada uno de 95 °C, 35 seg; 50 °C, 35 seg; 72 °C, 1 min.

Cebador	Secuencia (5'→3')	T _m (°C) ^(*)
Xho_Kex_PSI_F	<u>CTCGAGAAAAG</u> AGTCATGAGCCAGCAATGC	62
BglII_PSI(His) ₆ _stop_R	AGATCTTTAATGATGATGATGATGATGACTGGGTAAGCGATCA CAC	58
N86S_PSI_F	GGATGCAAACCAAATCAAACGAT <u>TCT</u> GAGACTGAAGATAACAT AATC	69
N86S_PSI_R	GATTATGTTATCTTCAGTCTC <u>AGAT</u> CGTTTGATTTGGTTTTGCAT CC	69

TABLA 17. Secuencias de los cebadores utilizados en la obtención de las construcciones conteniendo el PSI o su mutante de sitio de glicosilación, PSIN86S, downstream al dominio α -MF en el vector pKLAC1. Los sitios de restricción y el codón de fin de la traducción, incorporados en los cebadores, se encuentran subrayados por línea punteada y continua, respectivamente; mientras que las bases que codifican para la cola de histidinas y para el corte de la enzima Kex-2 se indican en los respectivos cebadores mediante línea de trazos. Los nucleótidos utilizados para generar la mutación N86S se muestran en negrita y subrayado.

(*) T_m= 4(C+G)+2(A+T)

Los productos obtenidos se analizaron por electroforesis de agarosa 1% (p/v) desde donde se purificaron para ser clonados en el vector pGEM-T Easy (construcción pGEM-PSI(His)₆) de acuerdo a lo detallado en las *Técnicas n°3* y *n°5*. La selección de clones positivos se realizó

mediante digestión con las enzimas de restricción *Bgl*III y *Xho*I y se eligió un clon positivo entre 12 analizados para su secuenciación automática.

Posteriormente, una alícuota de la construcción pGEM-PSI(His)₆ secuenciada y una alícuota del vector pKLAC1, fueron sometidas a digestión con las mismas enzimas de restricción (*Bgl*III y *Xho*I) a fin de generar extremos cohesivos que permitieran subclonar el inserto PSI(His)₆ en el vector de expresión, siguiendo el procedimiento descrito en la *Técnica n°8*. De esta manera, se generó la construcción pKLAC- α -MF-PSI(His)₆, chequeándose una vez más la presencia del inserto por análisis de restricción con las enzimas *Bgl*III y *Xho*I y confirmándose su secuencia por secuenciación automática.

Con el objetivo de obtener la construcción en la cual se pretendía eliminar el sitio de glicosilación presente en la secuencia del PSI, se emplearon los cebadores N86S_PSI_R y N86S_PSI_F (Tabla 18) y el plásmido pKLAC- α -MF-PSI(His)₆ como molde en la reacción de mutagénesis que condujo a la construcción pKLAC- α -MF-PSIN86S(His)₆. La mezcla de reacción usada se describe en la *Técnica n°2*, utilizándose en este caso 2,5U de polimerasa *Pfu* Turbo Polimerase (Stratagene). Durante la amplificación, se emplearon los parámetros de termociclado indicados en el manual Quick Change Site-Directed Mutagenesis Kit (catálogo n°: 200518).

Finalizada la reacción de PCR, la mezcla resultante se trató con 10 U de enzima *Dpn*I durante 1 hora e inmediatamente se usó para transformar células TOP10F' o XL1-Blue competentes (preparadas según la *Técnica n°4*), siguiendo el protocolo detallado en el manual Quick Change Site-Directed Mutagenesis Kit. Luego, el ADN de una colonia de la placa de transformantes se aisló por medio de un kit comercial (*Técnica n°6*) y se envió al servicio de secuenciación para confirmar el éxito de la mutagénesis y la consecuente obtención de la construcción pKLAC- α -MF-PSIN86S(His)₆.



Análisis por dinámica molecular de la influencia de la cola de polihistidinas terminal en la estructura tridimensional del dominio PSI

Antes de realizar las construcciones sin la región pro- del dominio α -MF a fin de reiterar la inclusión de la cola de polihistidinas en el C-terminal del inserto PSI, o bien agregar un sitio de corte para la enzima TEV que permita eliminarla a posteriori, se analizó mediante herramientas de bioinformática la influencia de las histidinas sobre la estructura tridimensional del PSI.

El PSI se modeló por homología mediante el servidor SWISS MODEL (Arnold et al., 2006; Kiefer et al., 2009) utilizando como referencia la estructura cristalina del dominio StAps-PSI (código PDB: 3RFI) de la PA de *Solanum tuberosum*, StAP-1. A partir del mejor modelo generado por este servidor, se prosiguió con la optimización de la geometría inicial empleando el programa Amber 11 (Case et al., 2010). En esta etapa se utilizó el campo de fuerza ff03 provisto por el programa para llevar a cabo la minimización. Los átomos del PSI se rodearon de una caja periódica de agua del tipo TIP-P32 que se extendió 10 Å a partir del contorno de la proteína. El sistema PSI-entorno se optimizó mediante un paso preliminar aplicando 1000 ciclos de minimización exclusivamente para las moléculas de agua del entorno y luego, 2500 ciclos para el sistema completo. La equilibración del sistema de 0 K a 298 K se realizó en condiciones de volumen constante (equilibración del tipo NTV) durante 20 ps para el sistema levemente restringido (10 kcal/mol Å²), y una segunda equilibración de 200 ps donde se fijó la presión como parámetro constante (equilibración del tipo NTP) para el sistema sin restricciones. Finalmente, se realizó la producción de dinámica molecular por un período de tiempo de al menos 16 ns, salvando imágenes de las coordenadas cada 1000 pasos.

Las estructuras de las moléculas de PSI conteniendo la cola de histidinas o el sitio de reconocimiento para la enzima TEV en el extremo C-terminal se prepararon mediante el modelo por homología descrito anteriormente al que se le agregaron los aminoácidos extras con el programa PyMol (Schrödinger). Para estas moléculas se siguieron los mismos procedimientos de minimización, equilibración y producción de dinámica molecular empleados para la molécula original del PSI.



Construcción pKLAC-PSI(His)₆ y pKLAC-PSIN86S(His)₆

Con el objetivo de obtener las construcciones que dieron lugar a la expresión de proteínas no fusionadas con el dominio α-MF de *K. lactis*, se empleó el método Quick Change Site-Directed Mutagenesis y los cebadores del_MF_PSI-pklac_F y del_MF_PSI-pklac_R (Tabla 18), específicamente diseñados para eliminar la secuencia correspondiente al α-MF y al sitio de corte de la enzima Kex-2 en un único paso de mutagénesis.

La reacción se llevó a cabo empleando la mezcla descrita en la *Técnica nº2* con 2,5 U de polimerasa *Pfu* Turbo Polimerase (Stratagene o Agilent). Como molde de ADN se utilizó la construcción pKLAC-α-MF-PSI(His)₆ o la construcción pKLAC-α-MF-PSIN86S(His)₆ previamente

obtenidas. El programa de amplificación en ambos casos se basó en el indicado en el manual Quick Change Site-Directed Mutagenesis Kit (catálogo nº: 200518).

Cebador	Secuencia (5' -> 3')	Tm (°C) ^(*)
del_MF_PSI-pklac_F	CCGTTGTTATGGCTGCTCCACTCGAGGTCATGAGCCAGCAATG CAAAAC	89,4
del_MF_PSI-pklac_R	GTTTTGCATTGCTGGCTCATGACCTCGAGTGGAGCAGCCATAA CAACGG	89,4
ins_TEV_PSI-pklac_F	GCTGTGTGATCGCTTACCCAGT GAAAATTTGTATTTCCAACATC ATCATCATCAT	79,7
ins_TEV_PSI-pklac_R	GATGATGATGATGATGTT GGAAATACAAATTTTCACTGGGTAA GCGATCACACAGC	79,7

TABLA 18. Secuencias de los cebadores utilizados en la obtención de las construcciones conteniendo el PSI o su mutante de sitio de glicosilación PSIN86S, *downstream* al péptido señal, sin el dominio α -MF en el vector pKLAC1. Las bases que codifican para el sitio de reconocimiento de la enzima TEV se indican mediante línea de trazos.

^(*) Tm = 4(C+G)+2(A+T)

El producto de PCR luego de ser tratado con 10 U de la enzima *DpnI*, se empleó para transformar células TOP10F' o XL1-Blue competentes (preparadas según *Técnica nº4*), como indica el protocolo detallado en el manual Quick Change Site-Directed Mutagenesis Kit.

Una vez obtenidas las construcciones sin las secuencias codificantes para el α -MF y el sitio Kex-2, las mismas se emplearon como molde en la siguiente reacción de mutagénesis en la cual se insertó la secuencia para el reconocimiento de la enzima TEV en el extremo 5' del PSI. La reacción de mutagénesis empleándose los cebadores ins_TEV_PSI-pklac_F e ins_TEV_PSI-pklac_R (Tabla 18), y la transformación de células TOP10F' se llevaron a cabo tal como se describió para el caso de la delección previa.

Finalmente, la presencia de las mutaciones se confirmó por secuenciación automático del ADN aislado a partir de una colonia de la placa de transformantes mediante un kit comercial (*Técnica nº6*).

De esta forma, se obtuvieron las dos construcciones, pKLAC-PSI(His)₆ y pKLAC-PSIN86S(His)₆, para la expresión del dominio PSI y su mutante de sitio de glicosilación, sin la fusión al dominio α -MF del vector pKLAC1 y con la incorporación del sitio TEV para cortar la cola de polihistidinas del C-terminal.

II. Expresión y Purificación de PSI en *K. lactis*

En este sistema se produjeron las siguientes formas del dominio PSI: α -MF-PSI(His)₆, α -MF-PSIN86S(His)₆; PSI(His)₆ y PSIN86S(His)₆.



*Preparación de células *K. lactis* GG799 competentes*

La competencia de la cepa GG799 de *K. lactis* se llevó a cabo en el laboratorio siguiendo un protocolo de dos días. En el primero, se inocularon 5 ml de medio YPGlu con una colonia de GG799 proveniente de una placa YPGlu y se llevaron a 30°C en estufa con agitación continua durante toda la noche. Al día siguiente se realizó una dilución 1:50 del cultivo en medio YPGlu fresco y se creció nuevamente en las mismas condiciones hasta obtener un valor de DO_{600nm} igual a la unidad. Alcanzado ese punto, el sedimento de células obtenido por centrifugación a 1000 x *g* durante 5 min, se lavó dos veces con agua estéril fría y una vez más con solución de sorbitol 1M pre-enfriada. Finalmente, las se resuspendieron en 1 ml de sorbitol 1 M frío.



Linealización de las construcciones y electrotransformación de las células GG799

Cada construcción basada en el vector pKLAC1 y conteniendo la secuencia para el PSI o para su mutante del sitio de glicosilación, se introdujo en células *K. lactis* GG799 competentes por electroporación empleando un equipo Gene Pulser Xcell Total System (Bio-Rad). En cada caso, la construcción circular fue previamente linealizada (obteniéndose el casete de expresión) mediante digestión con la enzima de restricción *SacII* y purificada (*Técnicas n^o7 y n^o3*). Una vez obtenido el ADN lineal (5 μ g), se colocó sobre una alícuota (50-100 μ l) de la suspensión de células competentes en una cubeta de electroporación de 0,2 cm (Bio-Rad) y se aplicó un pulso de 1,5 kV, 25 μ F, 200 Ω . Inmediatamente, las células se se dejaron recuperar en 1 ml de medio YPGlu a 30 °C con agitación durante 1 hora; transcurrido ese tiempo se lavaron se plaquearon en medio YCB-Acetamida. Las levaduras se dejaron crecer a 30 °C hasta aparición de colonias. En forma paralela, un control negativo fue incluido en el cual la transformación de células GG799 se realizó con agua Milli Q sin ADN.

Aproximadamente 25 colonias obtenidas se estriaron en una placa de medio YCB-Acetamida fresca y se dejaron crecer a 30 °C. Esta placa se utilizó para hacer stocks glicerolados e

identificar los clones que integraron en forma correcta el casete de expresión a su genoma (Fig. 33).

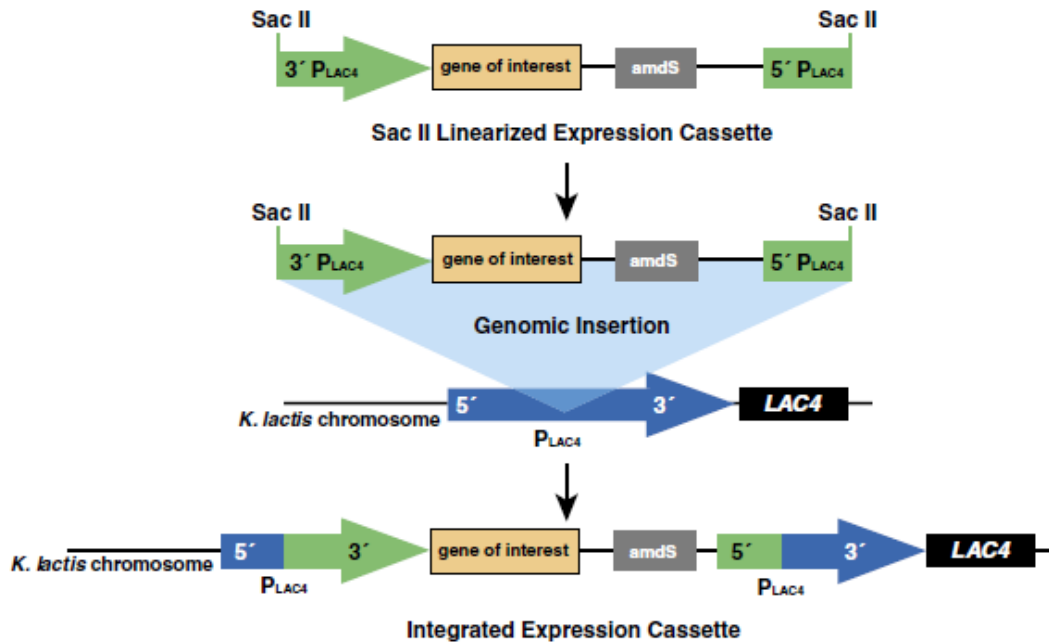


Fig. 33. Esquema de la integración del casete lineal de expresión obtenido por digestión con la enzima *SacII*, en el locus *LAC4* del genoma de *K. lactis*. (Extraído del manual *K. lactis* Protein Expression Kit—Recombinant Protein Expression in Yeast de New England Biolabs).



Identificación de células con el casete de expresión correctamente integrado y selección de transformantes

La presencia del casete de expresión en el genoma de las células GG799 transformadas se puso de manifiesto mediante la técnica *Colony-PCR* utilizando cebadores internos (diseñados a partir de la secuencia del inserto, PSI). Por cada construcción, se analizaron de 10 a 12 células electrotransformadas y se incluyó además, un control que consistió en células sin transformar sometidas al mismo procedimiento (GG799).

En esta técnica se hace uso de la célula entera y no del ADN genómico o plasmídico aislado, como molde para la amplificación por PCR. Para preparar las células, las mismas se resuspendieron desde una placa de medio YCB-Acetamida en sorbitol 1 M y se sometieron a dos ciclos de congelado/descongelado. Luego, la pared celular se digirió parcialmente con 30 U de liticasa (Sigma) a 37 °C por un lapso de 90 min. El digerido resultante y desnaturizado, se utilizó en la mezcla de reacción de PCR que se detalla en la *Técnica nº2*, empleando una

concentración de cebadores de 10 μ M y 2,5 U de polimerasa DFS-Taq Polymerase (Bioron). Los parámetros de termociclado para los 32 ciclos del segmento II del programa indicado en la *Técnica n°2*, fueron: 94 °C, 40 seg; 56 °C, 40 seg; 72 °C, 1 min.

El producto amplificado se analizó luego en un gel de agarosa como se indica en la *Técnica n°1*. En forma simultánea a la confirmación de la presencia del casete de expresión en el genoma de cada clon, se realizó una prueba de expresión en pequeña escala de los mismos clones analizados por PCR y el control. Se emplearon las siguientes condiciones de expresión: 10 ml de medio YPGal (4%), 30 °C, 200 rpm. Se retiró muestra transcurridas 48, 72, 96 y 120 hs de expresión para el análisis del medio extracelular mediante SDS-PAGE y Western blot (*Técnicas n° 14 y n° 18*).



Expresión de α -MF-PSI(His)₆, α -MF-PSIN86S(His)₆, PSI(His)₆ y PSIN86S(His)₆

La expresión de las diferentes formas del PSI derivadas del vector pKLAC1, se llevó a cabo en medio YPGal (4%) a 30 °C, con agitación continua durante 72 hs. Finalizada la expresión, se separó el medio extracelular por centrifugación desde donde se purificó posteriormente la proteína expresada.

III. Análisis de la glicosilación de los productos expresados en *K. lactis*.



*Tratamiento de las formas de PSI obtenidas en *K. lactis* con endoglicosidasas*

La desglicosilación de las proteínas α -MF-PSI(His)₆, α -MF-PSIN86S(His)₆ y PSI(His)₆ se realizó directamente a partir de los medios de cultivos sin purificar. Los ensayos de desglicosilación se hicieron usando dos glicosidasas diferentes:

- PNGasa F (New England Biolabs), una amidasa que corta entre el residuo de GlcNAc más interno y el residuo de asparagina en los complejos de oligosacáridos híbridos de alta manosa de las glicoproteínas N-ligadas. Remueve todo tipo de glicosilación N-ligada (Asn); excepto glicanos N-ligados cuando el residuo de GlcNAc más interno está unido a un residuo α (1-3) Fucosa. Esta modificación se encuentra más frecuentemente en plantas y algunas glicoproteínas de insectos (Fig. 34).

- Endo H (New England Biolabs) es una endoglicosidasa recombinante que corta dentro del centro de quitobiosa de oligosacáridos de alta manosa y algunos híbridos N-ligados de glicoproteínas. Esta enzima, a diferencia de la anterior, permite determinar más cercanamente el tipo de N-glicosilación ya que sólo corta carbohidratos de alta manosa y algunos híbridos (Fig. 34).

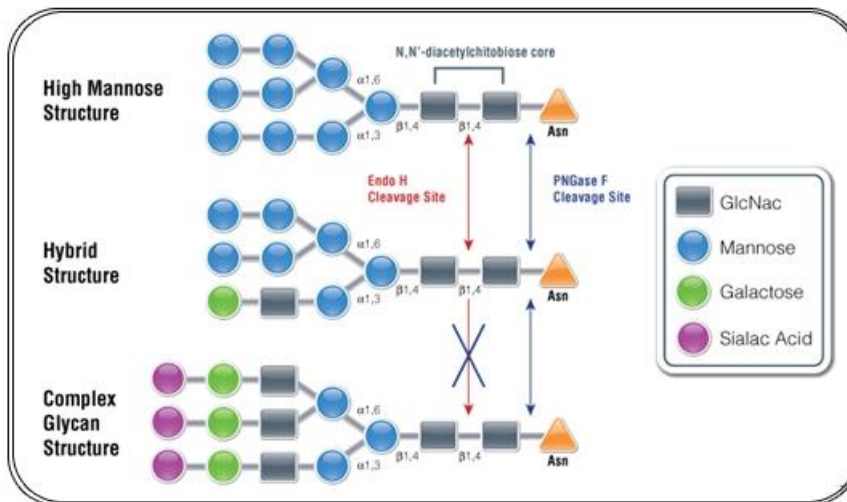


Fig. 34. Esquema de la acción de las endoglicosidasas PNGasa F y Endo H.

En todos los casos, la composición de la mezcla y las condiciones en que transcurrió la reacción fueron las indicadas por el fabricante. Luego de la digestión con cada endoglicosidasa, los productos de reacción se analizaron por inmunodetección empleando anticuerpos anti-His conjugado a fosfatasa alcalina (GenScript) como se describe en la *Técnica nº18*.



Análisis in silico de los sitios putativos de glicosilación

Si bien los sitios aceptores de N-glicosilación sobre secuencias de proteínas no están bien caracterizados, se sabe que la secuencia consenso o motivo Asn-Xaa-Ser/Thr (donde Xaa es cualquier aminoácido con excepción de Pro) es un pre-requisito para esta modificación post-transduccional en residuos Asn. Sin embargo, no todos estos tripletes se encuentran modificados y por tanto la secuencia no es discriminatoria entre las Asn que se encuentran glicosiladas de las que no lo están. El servidor de predicción NetNGlyc 1.0 Server desarrolla una red neuronal artificial sobre el contexto que rodea a estos tripletes, en un intento de discriminar entre un motivo aceptor y uno no aceptor. En una interpretación entrecruzada, las

redes podrían identificar 86% de los motivos glicosilados y 61% de los no glicosilados, con una confianza del 76%.

En cuanto a la O-glicosilación, no se ha identificado ninguna secuencia consenso de reconocimiento para las O-glicosiltransferasas. Julenius y colaboradores (2005) analizaron la base de datos de proteínas (PDB) obteniendo 12 estructuras glicosiladas en las cuales todos los sitios positivos se encontraron en regiones de giros. En base a estos estudios desarrollaron un método para predecir la localización de sitios de glicosilación de tipo mucina usando un enfoque de red neural artificial que resulta en la predicción correcta del 76% de los residuos glicosilados y el 93% de los no glicosilados. NetOGlyc 3.1 puede predecir sitios para proteínas nuevas a partir de propiedades promedio junto con el hecho que los sitios de glicosilación no son necesariamente conservados, es decir teniendo en cuenta que la glicosilación de tipo mucina en la mayoría de los casos es una propiedad del conjunto y no específica de un sitio.

Por tanto, se introdujo la secuencia aminoacídica de la proteína de fusión completa codificada por la construcción pKLAC- α -MF-PSI(His)₆, incluyendo el péptido señal y el prodominio α -MF del vector pKLAC1 en los programas NetNGlyc 1.0 Server y NetOGlyc 3.1 Server (www.expasy.org/tools) para obtener una predicción de los posibles sitios de N- y O-glicosilación presentes en la hipotética proteína de fusión sin procesar en la célula de *K. lactis*.

IV. Producción de las proteínas de secreción.



Purificación de los productos de expresión α -MF-PSI(His)₆ y α -MF-PSIN86S(His)₆, PSI(His)₆, PSIN86S(His)₆

Las proteínas recombinantes obtenidas en *K. lactis* se concentraron por precipitación con (NH₄)₂SO₄ en un 95% de su concentración saturante a 4 °C. Luego, las proteínas precipitadas se resuspendieron en buffer fosfato de sodio 25 mM de pH 7,5 conteniendo 100 mM de NaCl, se ultracentrifugaron a 114.028 × *g* durante 20 min a 4 °C y se pasaron por un filtro de 0,2 μ m. Las muestras filtradas se aplicaron en una columna de exclusión molecular como se describe en la *Técnica n°9*. Cada uno de los eluatos se sembró en dos geles de poliacrilamida en condiciones desnaturalizantes, uno de los cuales se tiñó con Coomassie Brilliant Blue R-250 mientras que el

otro se transfirió a una membrana de PVDF para la detección de la proteína de interés con anticuerpos anti-His (*Técnicas n^o14 y n^o18*).

En un segundo paso de purificación, el eluato obtenido en la primera cromatografía se aplicó a una columna HisTrap HP de 1 ml (Ge Healthcare), pre-empaquetada con Ni⁺² Sepharose High Performance y conectada a un equipo cromatográfico ÄKTA Purifier System (GE Healthcare). Esta cromatografía denominada IMAC, se desarrolló de acuerdo a las etapas y parámetros indicados en la *Técnica n^o11*.

A la salida de esta columna, las fracciones se analizaron nuevamente por SDS-PAGE e inmunoblot. Finalmente, el imidazol se removió de la proteína purificada mediante diálisis frente al buffer fosfato de sodio 25 mM de pH 7,5 conteniendo 100 mM de NaCl y utilizando una membrana de límite nominal de 3,5 kDa (Spectrum).



Corte in vitro del producto α -MF-PSI(His)₆ con la enzima Kex-2

Fracciones purificadas de la proteína recombinante obtenida por expresión del casete α -MF-PSI(His)₆, se sometieron a digestión con la enzima Kex-2 (PeproTech) que es la responsable de procesar al dominio α -MF en la proteína de fusión dentro de la célula de *K. lactis* GG799. Para ello, se mezclaron 37 μ l de la muestra con 1 μ g de Kex-2 y 40 μ l de buffer Tris-HCl 200 mM de pH 7,2 conteniendo CaCl₂ 2 mM. La reacción se dejó transcurrir durante toda la noche a 37 °C y se analizó luego por Western blot empleando anticuerpos anti-His.



Optimización del corte de la cola de histidinas de la proteína PSIN86S(His)₆ con la enzima TEV

Las fracciones de PSIN86S(His)₆ eluidas de la columna HisTrap se dializaron de la manera previamente mencionada, y en presencia de un agente reductor (DTT 1 mM) o el par glutatona reducida (3 mM) / glutatona oxidada (0,3 mM) debido a la existencia de puentes disulfuro en la estructura del PSI.

Posteriormente, sobre volúmenes pequeños de dializado se probaron varias condiciones de corte con la enzima TEV, para la cual se incorporó un sitio de reconocimiento entre la cola de seis histidinas y el PSI durante el clonado de la proteína PSIN86S(His)₆ en el vector pKLAC1. Se analizaron dos relaciones masa en masa de enzima: sustrato (1:10 y 1:25) a dos temperaturas

diferentes, 4 y 20 °C (TA), durante 2,5 hs; 5 hs y ON. La reacción se llevó a cabo con el agregado de EDTA en concentración 1 mM en el buffer de reacción.

Transcurrido el tiempo establecido de reacción, las muestras se centrifugaron a 43000 x *g* durante 20 min a 4 °C y las fracciones solubles se analizaron en un gel de poliacrilamida como se indica en la *Técnica n°14*.



Corte de la proteína PSIN86S(His)₆ con TEV

La digestión con la enzima TEV del volumen restante de dializado conteniendo la proteína PSIN86S(His)₆ se llevó a cabo como se describió en el paso de optimización, fijando las siguientes condiciones: relación enzima a sustrato de 1:10; glutatona reducida (3 mM) / glutatona oxidada (0,3 mM), temperatura y tiempo de reacción 20 °C y ON, respectivamente. Debido a la diferencia de masas entre el PSIN86S y la enzima TEV, la separación de ambas después de la digestión se realizó utilizando un dispositivo de filtro Amicon Ultra-15 (Millipore) con un punto de corte de 30 kDa. En el filtrado se obtuvo el PSI, mientras que el retenato contuvo la enzima.



Análisis de los productos de expresión por técnicas de proteómica

Las bandas proteicas obtenidas en gel de poliacrilamida con las muestras purificadas provenientes de las expresiones de las construcciones pKLAC- α -MF-PSI(His)₆, pKLAC- α -MF-PSIN86S(His)₆ y pKLAC-PSIN86S(His)₆—ésta última analizada pos-tratamiento con la enzima TEV de acuerdo a lo indicado en el punto anterior—fueron cortadas y sometidas a digestión tróptica para la obtención de la huella peptídica de cada una de las bandas, procediendo según lo detallado en la *Técnica n°19*.

Ensayos antimicrobianos realizados con diferentes formas del PSI expresadas en *K. lactis*^(*)



Actividad antibacteriana

Test de dilución en agar

En este método el agente antimicrobiano se incorpora dentro del medio con agar, de manera tal que cada placa contenga una concentración de antibiótico diferente. Una alícuota de 1 a 2 μl (10^4 UFC) del inóculo de cada microorganismo a ensayar se aplica sobre la superficie del agar. Luego, la placa se incuba durante toda la noche a 37 °C sin atmósfera de CO_2 .

Este test se empleó para evaluar la muestra PSIN86S expresada en *K. lactis*, a una concentración fija de 0,018 mg/ml en la placa de LB preparada a pH 5,0. Los microorganismos probados mediante este método han sido: *Escherichia coli*, *Salmonella typhi*, *Staphylococcus aureus*, *Proteus mirabilis*, *Pseudomonas aeruginosa*, *Enterococcus faecalis*.

Test en medio líquido

Este test se realizó con la muestra PSIN86S expresada en *K. lactis*. Un volumen de 40 μl de proteína (0,278 $\mu\text{g}/\mu\text{l}$) en buffer fosfato de sodio 25 mM (pH 7,5), conteniendo NaCl 100 mM se agregó a suspensiones conteniendo 10^5 y 10^6 UFC/ml de *E. coli* y de *S. typhi*. Después de una incubación a 37 °C durante 3 hs, las suspensiones se plaquearon sobre agar LB preparado a pH 5,0. Se incluyeron controles del buffer empleado.

Test de KIRBY-BAUER o de difusión en disco

Se realizó una suspensión de cada cepa bacteriana a ensayar (*E. coli*, *S. typhi* y *S. aureus*) con una turbidez correspondiente al valor 0,5 de la escala de Mc Farland, luego mediante un hisopo se inoculó la superficie completa de una placa agar LB a pH 5,0 y a pH 7,0. Se colocó en cada placa un disco completamente embebido con 80 μl de la muestra PSIN86S (0,278 $\mu\text{g}/\mu\text{l}$) y como controles positivos se incluyeron discos con 75 μl del antibiótico kanamicina (25 mg/ml) y 80 μL de ampicilina (0,6 $\mu\text{g}/\mu\text{l}$).



Actividad antifúngica

Test de co-crecimiento en placa

Sobre la superficie de placas de medio PDA (pH 5,0), PDA (pH 6,8) e YPGal (pH 5,0), se aplicó un disco embebido en células de *K. lactis* que contienen el casete para la expresión de la proteína PSI(His)₆ sin el dominio α -MF, y un segundo disco con células de una de las siguientes especies de hongos: *Lewia infectoria*, *Alternaria alternata*, *Dreschsiera biseptata*, *Botryosphareia obtuse*, *Pleospora herbarum*. Cada una de estas especies se creció a 30 °C también en una placa con la cepa GG799 sin transformar y con un cepa de GG799 expresando una proteína heteróloga cualquiera (GFP) como control del estado metabólico de la levadura GG799 productora de PSI (cepa gentilmente cedida por el bioquímico Pedro Curto de la Unidad de Biotecnología Molecular del Biocant Park). En otra placa los hongos crecieron en condiciones libres, es decir sin presencia de levaduras. En forma adicional, dos bacterias con actividad antifúngica, Fito_B108 y Fito_S127B, se emplearon como control positivo. Luego de ser incubada la placa, la distancia del halo de crecimiento del hongo se midió y se determinó el porcentaje de inhibición del crecimiento micelar (ICM) como sigue:

$$\text{ICM (\%)} = \frac{(\text{Halo de crecimiento sin levadura}) - (\text{Halo de crecimiento en presencia de levadura})}{(\text{Halo de crecimiento sin levadura})} \times 100$$

Test de KIRBY-BAUER o de Difusión en disco

La proteína PSIN86S expresada en *K. lactis* fue sometida a ensayos antifúngicos en los cuales los hongos fueron crecidos en medio PDA (pH 5,0) en un disco a 2,5cm del borde de la placa y a 3 cm del disco conteniendo la muestra de proteína. Las placas se incubaron a 30 °C durante 10 días. Los ensayos se hicieron por triplicado con las especies *Dreschlera biseptata*, *Lewia Infectoria*, *Alternaria alternata*. Se incluyeron controles positivos con Anfotericina B.

(*) Estos ensayos se realizaron mediante la colaboración del Bioquímico Pedro Curto de la Unidad de Biotecnología Molecular del Biocant Park, Cantanhede, Portugal.

Resultados—Plant Specific Insert (PSI)

Obtención de PSI en *Kluyveromyces lactis*

A partir del ADNc que codifica para la cirsinina completa, en su forma de preprocirsina, contenido en el vector pGEM-T Easy, se clonó la región correspondiente al dominio PSI en el vector pKLAC1 que permitió su expresión en el sistema de levaduras. Para ello, se amplificó la secuencia de interés empleando las condiciones de reacción indicadas en *Materiales y Métodos—PSI* y cebadores diseñados específicamente en base a la secuencia obtenida de la preprocirsina (número de acceso: JN703462).

A. Expresión de proteínas secretadas usando pKLAC1

AI. Obtención de los vectores pKLAC- α -MF-PSI(His)₆ y pKLAC- α -MF-PSIN86S(His)₆ para la expresión en *K. lactis*

De acuerdo a las indicaciones del fabricante, en el sistema *K. lactis*, la proteína heteróloga con su N-terminal nativo es secretada al medio extracelular dirigida por la fusión al dominio secretorio *alfa mating factor* (α -MF) de *Sacharomyces cerevisiae* presente en el vector pKLAC1. Este dominio es escindido por la enzima Kex-2 en el *trans*-Golgi previo a la secreción de la proteína esperada.

Para llevar a cabo la expresión de estas proteínas se empleó la estrategia de clonado entre los sitios *Xho*I y *Bgl*II del vector (estrategia de clonado número I presentada en el manual *K. lactis* Protein Expression Kit—Recombinant Protein Expression in Yeast).

La secuencia que codifica el PSI (Fig. 13) se amplificó a partir del plásmido pET-procirsina utilizando los cebadores diseñados específicamente para subclonar el PSI entre los sitios *Xho*I y *Bgl*II del vector pKLAC1. Se obtuvo una banda de aproximadamente 350 pb que se ligó al

vector pGEM-T Easy y el plásmido resultante se amplificó en bacterias TOP10F'. La presencia del inserto en los clones, se comprobó por digestión de los plásmidos pGEM-PSI(His)₆ aislados. La secuencia del PSI proveniente del clon número 12 digerido con *XhoI* y *BglII*, se subclonó en el vector pKLAC1 digerido con las mismas enzimas. El nuevo plásmido generado pKLAC- α -MF-PSI(His)₆, se transformó una vez más en *E. coli* y se repitió el procedimiento de selección de clones positivos. Los resultados de la reacción de PCR y ambos pasos de *screening* se muestran en la Fig. 35.

El plásmido pKLAC- α -MF-PSI(His)₆ se construyó a partir del anterior por mutagénesis dirigida como se detalló en *Materiales y Métodos—PSI*. Luego de expandirlo en células TOP10F', se seleccionó un clon para verificar la introducción de la mutación por secuenciación automática.

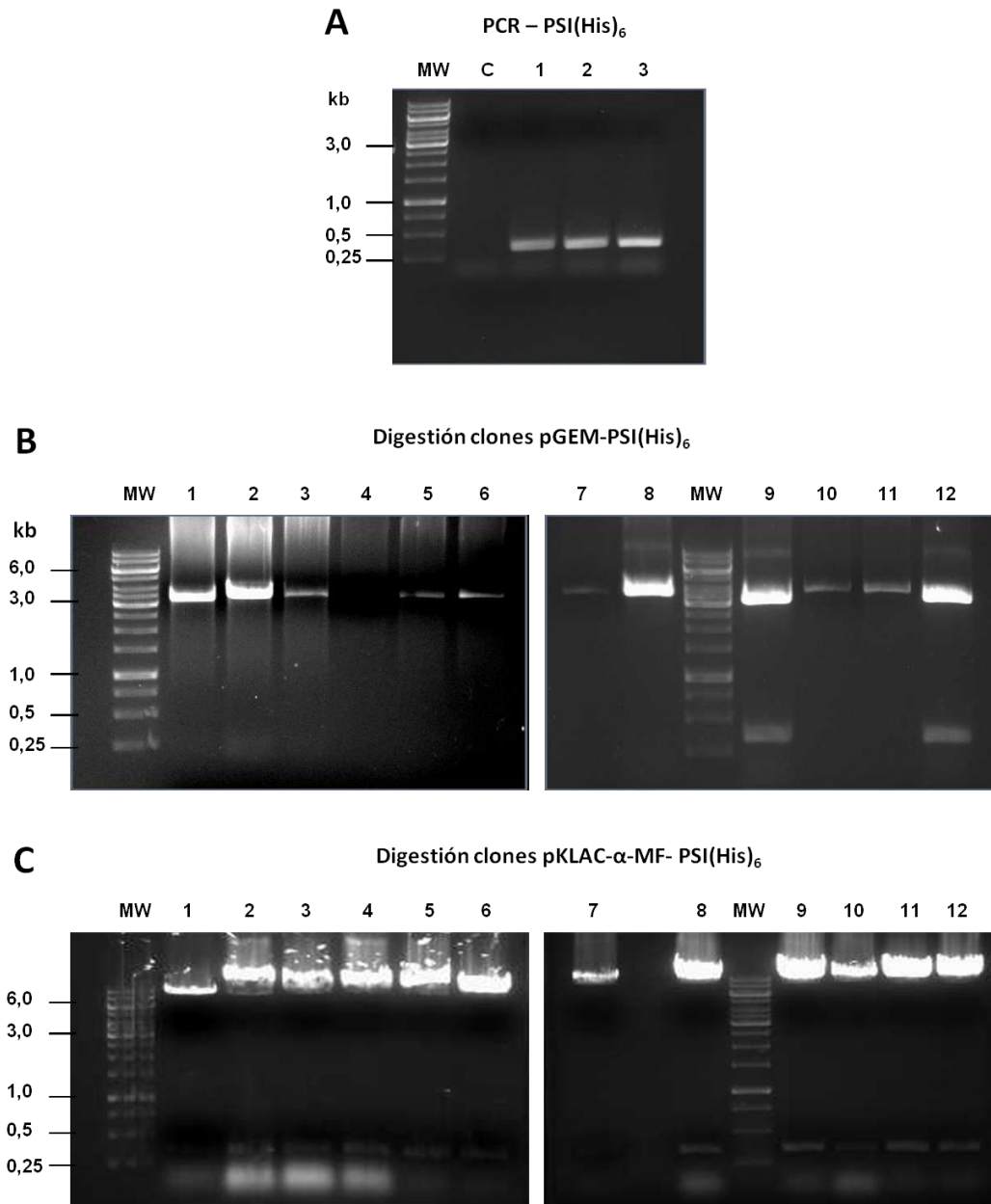
Obtención de la construcción pKLAC- α -MF-PSI(His)₆

Fig. 35. Obtención del plásmido pKLAC1 conteniendo la secuencia del PSI de *C. vulgare* con una cola de seis histidinas en su extremo C-terminal, pKLAC- α -MF-PSI(His)₆. La región de ADNc que codifica para el PSI de *C. vulgare* se amplificó por PCR usando como molde el plásmido pET-procirisina y cebadores específicos que agregaron además, la cola de histidinas. El producto de amplificación obtenido se clonó en pGEM-T Easy y posteriormente, se subclonó en el vector pKLAC1 para su expresión en *K. lactis*. **(A)** Análisis electroforético en gel de agarosa (1%) del producto amplificado mediante PCR: *Calle MW*, marcadores de masas moleculares de ADN 1 kb; *Calle C*, control negativo de PCR; *Calle 1 a 3*, triplicado de la reacción de PCR. **(B)** Análisis por restricción de clones transformados con el plásmido pGEM-PSI(His)₆: *Calles MW*, marcadores de masas moleculares de ADN 1 kb; *Calles 1 a 12*, plásmidos pGEM-PSI(His)₆ aislados de los clones 1 a 12 respectivamente, y digeridos con las enzimas de restricción *XhoI* y *BglII*. **(C)** Análisis por restricción de los clones transformados con el plásmido pKLAC- α -MF-PSI(His)₆: *Calles MW*, marcador de masas moleculares de ADN de 1 kb; *Calles 1 a 12*, plásmidos pKLAC- α -MF-PSI(His)₆ aislados de 12 clones, y digeridos con las enzimas de restricción *XhoI* y *BglII*.

AII. Expresión de PSI en *K. lactis* a partir de las construcciones pKLAC- α -MF-PSI(His)₆ y pKLAC- α -MF-PSIN86S(His)₆



Selección de clones de *K. lactis* conteniendo las secuencias del PSI integradas en el genoma como casete de expresión y optimización del tiempo de expresión

Las construcciones pKLAC- α -MF-PSI(His)₆ y pKLAC- α -MF-PSIN86S(His)₆ generadas en el vector pKLAC1, se transformaron en la cepa GG799 de *K. lactis*. Para ello, los plásmidos se linealizaron por digestión con la enzima *SacII* (Fig. 36) generando los casetes de expresión correspondientes.

Linealización de pKLAC- α -MF-PSI(His)₆ y pKLAC- α -MF-PSIN86S(His)₆ por digestión con *SacII*

—Obtención de casetes de expresión—

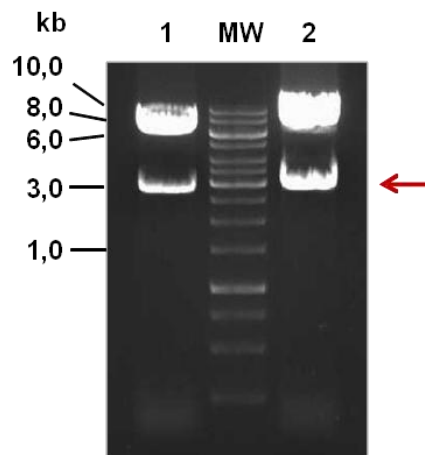


Fig. 36. Generación de los casetes de expresión lineales conteniendo la secuencia del péptido señal y del dominio α -MF correspondientes al vector pKLAC y el inserto PSI(His)₆ o PSIN86S(His)₆. Las construcciones obtenidas subclonando las secuencias que codifican para el PSI(His)₆ o su mutante de sitio de glicosilación, PSIN86S(His)₆, en el vector pKLAC1 se linealizaron mediante digestión con la enzima *SacII* generando los casetes (flecha roja) que permiten la integración en la región del locus del promotor LAC4 en el genoma de *K. lactis*. *Calle 1*, pKLAC- α -MF-PSI(His)₆; *Calle 2*, pKLAC- α -MF-PSIN86S(His)₆; *Calle MW*, marcadores de masas moleculares de ADN 1 kb.

Posteriormente, cada uno de los casetes de expresión purificados se introdujo en una alícuota de células competentes por medio de un pulso de voltaje como se indicó en *Materiales y Métodos—PSI*. Las células transformadas se plaquearon en medio sólido YCB-Acetamida y se incubaron a 30 °C hasta aparición de colonias.

La amplificación de una banda de aproximadamente 350 pb en el análisis por *Colony PCR* permitió identificar 10 clones que integraron eficientemente cada casete de expresión (Fig. 37). El uso de cebadores internos del PSI en esta técnica no permite diferenciar la integración de multicopias—mejores candidatos a producir mayores cantidades de proteína secretada—de la integración de copia única, por lo que esta información se complementó con ensayos de expresión en pequeña escala.

Integración de los casetes de expresión con las secuencias α -MF-PSI-(His)₆ y α -MF-PSIN86S(His)₆ en el genoma de *K. lactis*

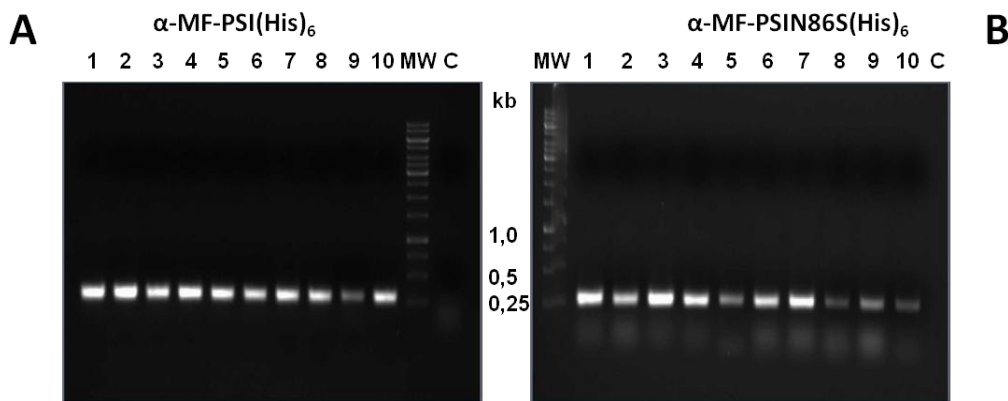


Fig. 37. Identificación de células de *K. lactis* con casetes de expresión correctamente integrados en su genoma. La presencia de la secuencia PSI(His)₆ o PSIN86S(His)₆ en los genomas de las células transformadas de *K. lactis* se detectó por amplificación de una banda de aproximadamente 350 pb utilizando células enteras. Análisis electroforético en gel de agarosa (1%) del producto amplificado mediante *Colony PCR* de células transformadas con el casete conteniendo (A) la secuencia α -MF-PSI(His)₆ y (B) la secuencia α -MF-PSIN86S(His)₆: *Calles MW*, marcadores de masas moleculares de ADN 1 kb; *Calles C*, control negativo de integración con células GG799 sin transformar; *Calles 1 a 10*, productos de amplificación obtenidos con los transformantes 1 a 10, respectivamente.

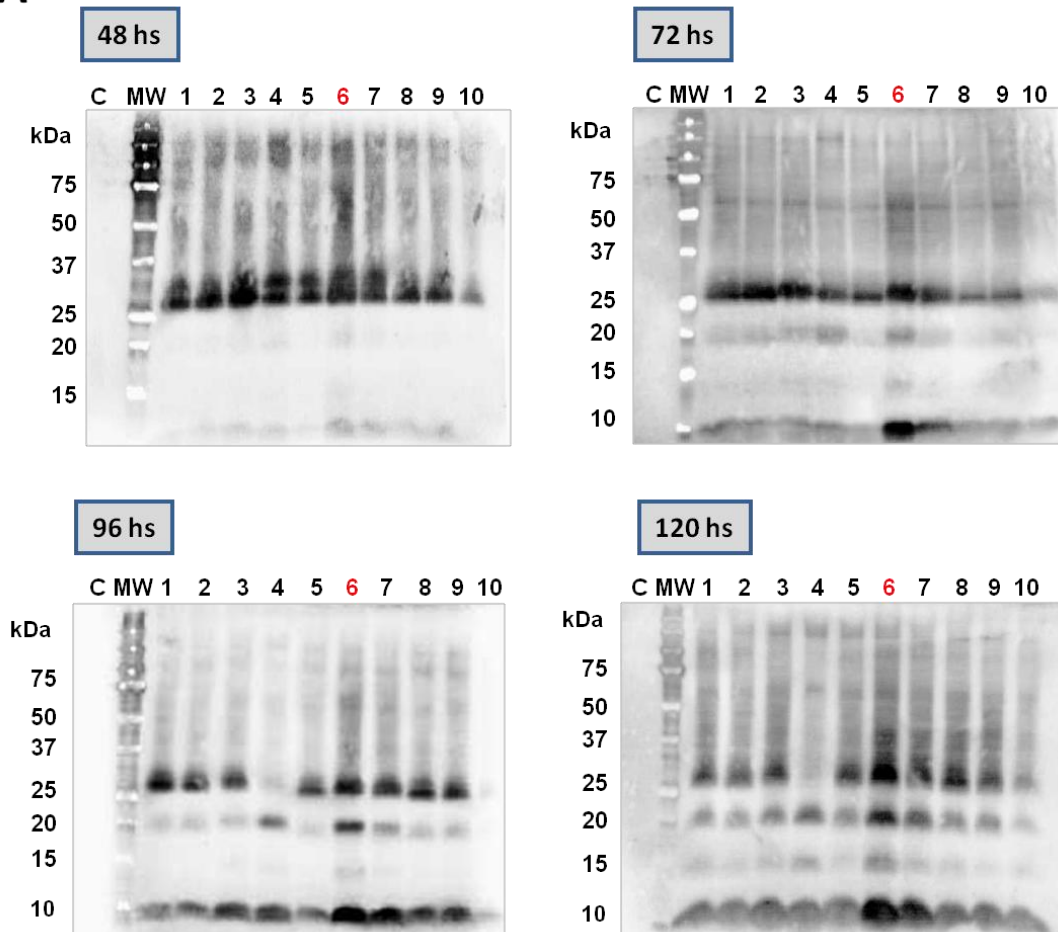
La proteína recombinante secretada se identificó por inmunodetección con anticuerpo anti-His en los ensayos de expresión en pequeña escala de los clones analizados a diferentes tiempos: 48, 72, 96, 120 y 144 hs. En la Fig. 38 A se muestran los blots realizados con las alícuotas tomadas a cada uno de los tiempos de análisis para los clones transformados con el casete conteniendo la secuencia α -MF-PSI(His)₆, mientras que para las expresiones de los transformantes con el casete proveniente de la construcción pKLAC- α -MF-PSIN86S(His)₆, se muestra sólo la inmunodetección realizadas luego de 72hs de expresión (Fig. 38 B). Como se observa en la Fig. 38 se obtuvieron poblaciones de proteínas muy heterogéneas para ambas construcciones. Se observó además, un inusual “chorreado” que se extiende desde la zona de

masas moleculares esperadas (aproximadamente 20 kDa) hasta valores superiores a los 75 kDa. En ambos casos se observa también una banda de aproximadamente 10 kDa que podría deberse a la expresión de formas truncadas de la proteína o bien, a degradación inespecífica por parte de las proteasas del huésped.

A partir de los resultados obtenidos (Fig. 38) se eligió para la expresión en mayor volumen el clon que presentó mayor nivel de expresión en *K. lactis* para cada proteína, indicado en rojo en las respectivas figuras: (A) el clon 6 de α -MF-PSI(His)₆ y (B) el clon 4 para el caso de α -MF-PSIN86S(His)₆.

Inmunodetección de la proteína secretada por los clones de *K. lactis* transformados con los casetes conteniendo las secuencias α -MF-PSI(His)₆ y α -MF-PSIN86S(His)₆

A



B

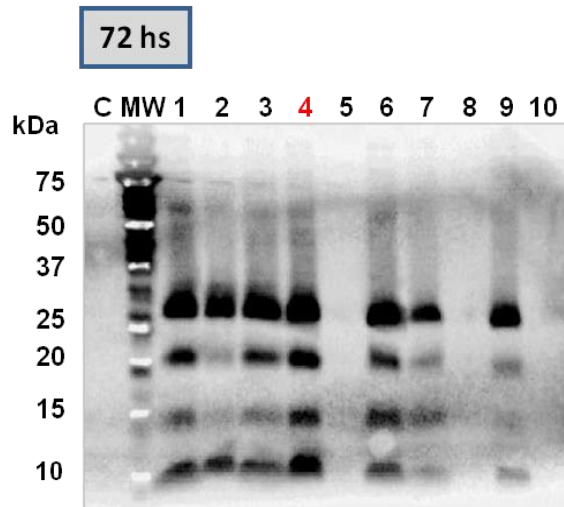


Fig. 38. Ensayos de expresión en pequeña escala de las distintas formas de PSI en el sistema heterólogo *K. lactis*. Se analizó la expresión de proteínas por los mismos clones de transformantes sometidos a *Colony PCR*, en 10 ml de medio YP induciendo con 4% de galactosa y creciendo a 30 °C. Las formas de PSI(His)₆ secretadas al medio de cultivo por los diferentes clones y a distintos tiempos de expresión, se detectaron por Western blot con anticuerpos anti-His: **(A)** Calles MW, marcadores de masas moleculares; Calle C, control negativo de integración con células GG799 sin transformar; Calles 1 a 10, medios de cultivo de las expresiones de los clones 1 a 10 de *K. lactis* transformados con el casete conteniendo la secuencia α -MF-PSI(His)₆ tomados a las 48, 96, 120 y 144 hs de expresión. **(B)** Calle MW, marcadores de masas moleculares; Calle C, control negativo de integración con células GG799 sin transformar; Calles 1 a 10, medios de cultivo de las expresiones de los clones 1 a 10 de *K. lactis* transformados con el casete incluyendo la secuencia α -MF-PSIN86S(His)₆ tomados a las 72 hs de expresión.

En la Fig. 39 se muestran en un único Western blot los diferentes tiempos de expresión para el clon 6 y para el clon 4, con el casete que contiene la secuencia α -MF-PSI(His)₆ y α -MF-PSIN86S(His)₆ integrado en su genoma, respectivamente. De este modo se pudo comparar la intensidad de la señal a cada tiempo y establecer un tiempo de 72 hs para la expresión de ambas formas de PSI obtenidas en *K. lactis*.

Análisis en función del tiempo de los productos de expresión de los casetes con las secuencias α -MF-PSI(His)₆ y α -MF-PSIN86S(His)₆.

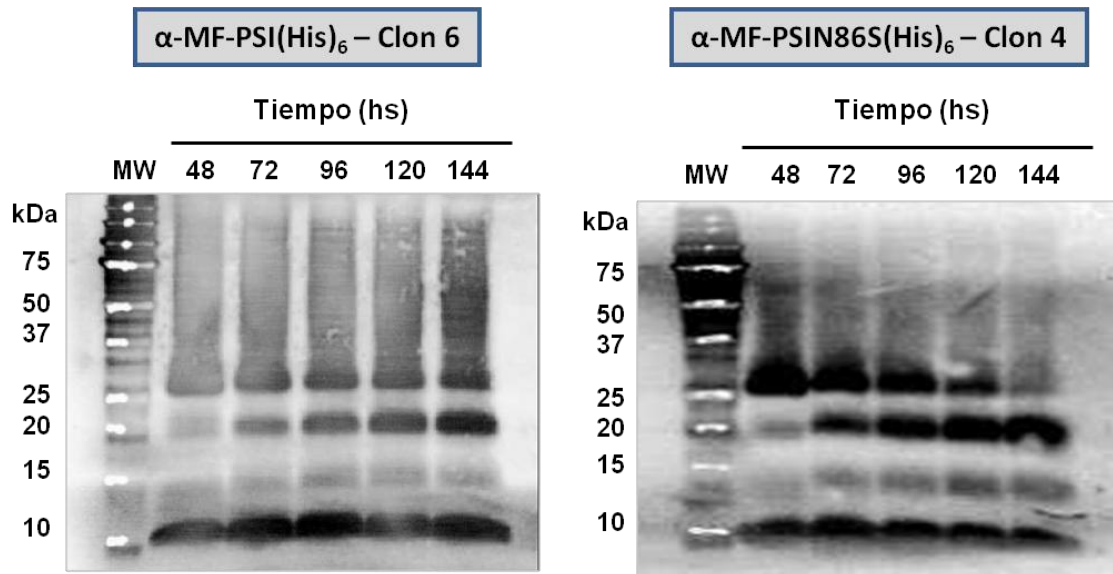


Fig. 39 Análisis por Western blot de la expresión en función del tiempo de los clones 6 y 4 conteniendo los casetes con las secuencias α -MF-PSI(His)₆ y α -MF-PSIN86S(His)₆, respectivamente. La cantidad de proteína secretada al medio de cultivo por cada uno de estos clones en función del tiempo, se comparó en un único Western blot utilizando anticuerpos anti-His: *Calles MW*, marcadores de masas moleculares; *Calles 48, 72, 96, 120, 144*, horas de expresión transcurridas al tomar la muestra de los medios de cultivo.

AIII. Análisis de la glicosilación de los productos expresados en *K. lactis* a partir de los casetes con las secuencias α -MF-PSI(His)₆ y α -MF-PSIN86S(His)₆



*Tratamiento de las formas de PSI obtenidas en *K. lactis* con endoglicosidasas*

Los medios de cultivos de los clones seleccionados durante los ensayos realizados en pequeña escala, se trataron con glicosidasa PNGasa F, la cual corta el enlace entre el residuo de GlcNAc más interno y el residuo de asparagina en los complejos de oligosacáridos híbridos de alta manosa de las glicoproteínas N-ligadas, con el objetivo de contrastar las masas moleculares de las proteínas expresadas con los pesos teóricos calculados con la herramienta Compute pI/Mw (www.expasy.org/tools):

- Peso molecular nominal de PSI(His)₆ = 12571,22 Da
- Peso molecular nominal de PSIN86S(His)₆ = 12544,19 Da

Como se observa en la Fig. 40, cuando las formas wt y mutante obtenidas por expresión de los casetes con las secuencias α -MF-PSI(His)₆ y α -MF-PSIN86S(His)₆ son digeridas con PNGasa F, las masas de las bandas generadas en el tratamiento no coincidieron con los pesos teóricos del PSI(His)₆ y PSIN86S(His)₆ no glicosilados. A su vez, el cambio de movilidad electroforética de las bandas observado para la muestra proveniente de la expresión del casete incluyendo al α -MF-PSIN86S(His)₆ no está de acuerdo con lo esperado ya que, al tratarse del mutante del único sitio de glicosilación presente en el PSI, no debió apreciarse pérdida de masa por acción de glicosidasas. En conjunto, los datos obtenidos por expresión de los casetes con las formas α -MF-PSI(His)₆ y α -MF-PSIN86S(His)₆ sugirieron la presencia de una porción extra secretada junto al PSI.

Tratamiento de los productos de expresión de los casetes que contienen las secuencias α -MF-PSI(His)₆ y α -MF-PSIN86S(His)₆ con PNGasa F

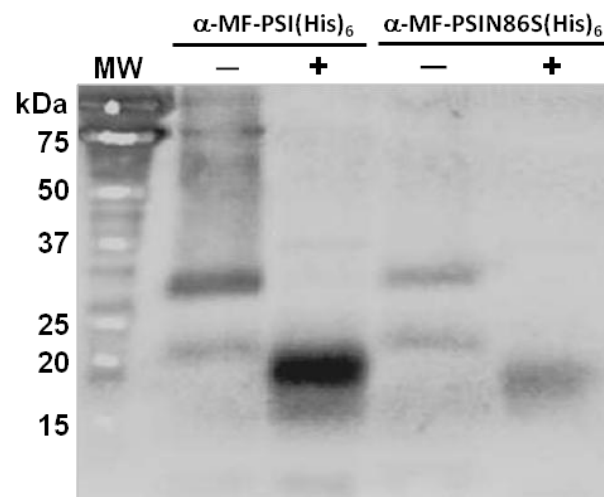


Fig. 40. Deglicosilación de las distintas formas de PSI obtenidas en *K. lactis* por acción de la enzima PNGasa F. Una alícuota del medio de cultivo con la forma α -MF-PSI(His)₆ o la forma α -MF-PSIN86S(His)₆ secretada, se sometió a digestión con la glicosidasa PNGase F durante 3 hs. Los productos digeridos se analizaron por inmunodetección con anticuerpo anti-His: *Calles MW*, marcadores de masas moleculares; *Calles (-)*, alícuota sin digerir; *Calles (+)*, alícuotas tratadas con PNGasa F.

Se calcularon entonces, los pesos moleculares teóricos incluyendo el dominio α -MF, los cuales arrojaron valores mucho más próximos a los estimados mediante los blots de las muestras deglicosiladas:

- Peso molecular nominal de α -MF-PSI(His)₆ = 19593,02 Da
- Peso molecular nominal de α -MF-PSIN86S(His)₆ = 19565,99 Da

La presencia del dominio α -MF en las proteínas secretadas en este sistema indica que el procesamiento de las mismas no se lleva a cabo de la manera esperada en el *trans*-Golgi de la levadura. Para comenzar a abordar esta hipótesis, se evaluó la sensibilidad de las proteínas recombinantes a la endoglicosidasa EndoH que corta el centro de quitobiosa de oligosacáridos N-ligados de alta manosa y algunos híbridos no procesados por la enzima del Golgi alfa manosidasa II (Alberts et al., 2002). Como se observa en la Fig 41, la forma obtenida por expresión del casete conteniendo la secuencia α -MF-PSI(His)₆ resultó sensible a esta enzima, lo que indica que la proteína no ha sido procesada más allá del retículo endoplasmático. El perfil obtenido por tratamiento con EndoH mostró dos bandas de aproximadamente 18 kDa y 12,5 kDa, desapareciendo todas las bandas de alto peso molecular e incluso la mayoritaria de 21 kDa; asimismo, se deduce de la imagen que las bandas de alto peso molecular se corresponden con formas hiperglicosiladas y que la glicosilación es predominantemente N-ligada.

**Tratamiento del producto de expresión del casete con la secuencia
 α -MF-PSIN86S(His)₆ con Endo H**

124

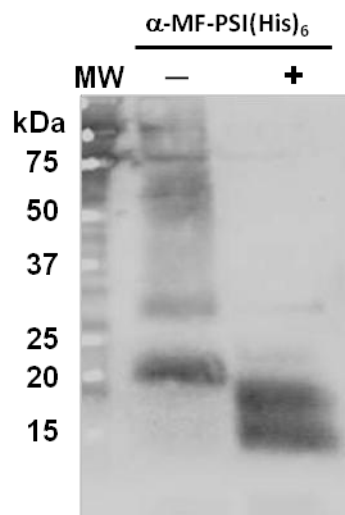


Fig. 41. Deglicosilación del producto de expresión del casete que contiene la secuencia α -MF-PSI(His)₆ obtenido en *K. lactis* por acción de la enzima Endo H. Una alícuota del medio de cultivo de la expresión en pequeña escala del clon transformado con la secuencia α -MF-PSI(His)₆, se trató con glicosidasa Endo H durante 3 hs. Los productos de digestión se analizaron por inmunodetección con anticuerpo anti-His: *Calle MW*, marcadores de masas moleculares; *Calle (-)*, alícuota sin glicosidasa en la mezcla de reacción; *Calle (+)*, alícuota tratada con Endo H.



Análisis in silico de los sitios putativos de glicosilación

Para confirmar lo observado experimentalmente, se analizó la existencia de potenciales sitios de N- y O-glicosilación en el dominio α -MF mediante NetNGlyc 1.0 Server y NetOGlyc 3.1 Server (www.expasy.org/tools), respectivamente. En efecto, estos servidores predicen dos sitios putativos de N-glicosilación en los residuos 58 (**NGT**) y 68 (**NTT**) del α -MF como se muestra en la Fig. 42 B donde ambas posiciones arrojan un potencial (líneas verticales) superior al umbral estándar de 0,5 con alta especificidad (de acuerdo a las 9 redes neurales que componen el método predictivo empleado), y los dos residuos treonina 24 y 26 del motivo **TET** con una calificación del mejor predictor general (G-score) mayor a 0,5 (Fig. 42 C) en que los carbohidratos se unen al oxígeno (O-glicosilación). Por tanto, la presencia del dominio α -MF en la proteína secretada desde las células de *K. lactis* generaría además de un aumento en la masa molecular con respecto a la esperada para el PSI(His)₆, cuatro sitios de glicosilación adicionales que explicarían la notable heterogeneidad de poblaciones glicosiladas como así también la pérdida de masa observada en PSIN86S(His)₆ por tratamiento con PNGasa F y Endo H.

Predicción de sitios putativos de glicosilación en la proteína de fusión α -MF-PSI(His)₆

A

```

MKFSTILAASTALISVVM AAPVSTETDIDDLPI SVPEEALIGFID 45
LTGDEVSLLPVNGTHTGILFLNTTIAEAAFADKDDLEKRVMSQQ 90
CKTLVSQYGKSIIEMLLSEAQPKICSQMKLCTFDGARDVSSIIE 135
SVVDKNNGKSSGGANDEMCTFCEMAVVWMQNQIKRNETEDNIINY 180
VNELCDRLPSHHHHHH* 196
    
```

- Péptido señal
- PSI
- α -MF
- Cola de (His)₆

B

CENTERFO
 RBIOLOGI
 CALSEQU
 ENCEANA
 LYSIS CBS

NetNGlyc 1.0 Server - prediction results

Technical University of Denmark

Asn-Xaa-Ser/Thr sequons in the sequence output below are highlighted in blue.
 Asparagines predicted to be N-glycosylated are highlighted in red.

Output for 'mf-psi-his'

Name: mf-psi-his Length: 196
 MKFSTILAASTALISVVM AAPVSTETDIDDLPI SVPEEALIGFIDLTGDEVSLLPVNGTHTGILFLNTTIAEAAFADKDDLEKRVMSQQCKTLVSQYGKSIIEMLLSEAQPKICSQMKLCTFDGARDVSSIIE SVVDKNNGKSSGGANDEMCTFCEMA VVWMQNQIKRNETEDNIINYVNELCDRLPSHHHHHH

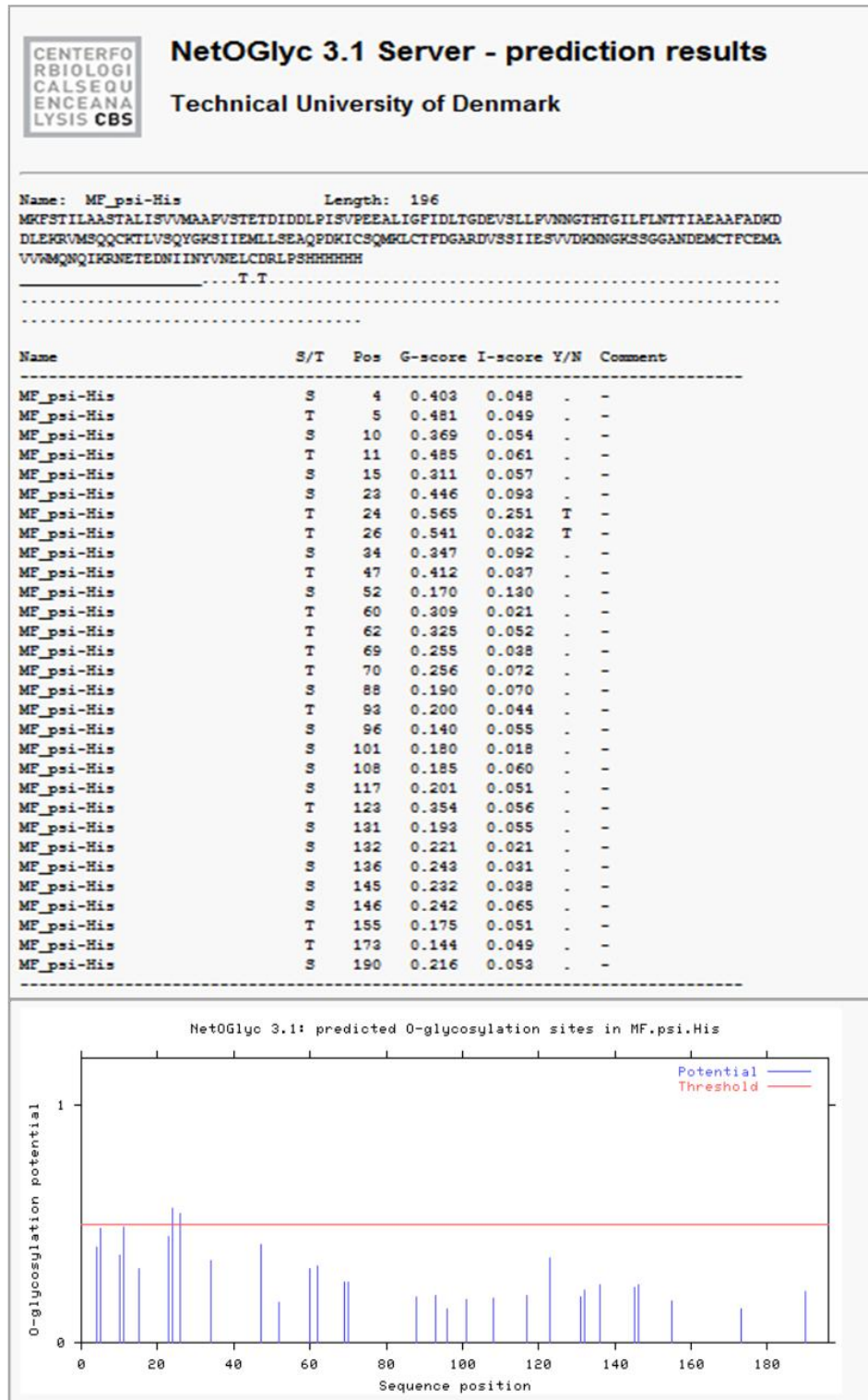
.....N.....N.....	80
.....N.....	160
.....N.....	240

(Threshold=0.5)

SeqName	Position	Potential	Jury agreement	N-Glyc result
mf-psi-his	58 NGTH	0.6640	(9/9)	++
mf-psi-his	68 NTTI	0.6509	(9/9)	++
mf-psi-his	171 NETE	0.5310	(6/9)	+

NetNGlyc 1.0: predicted N-glycosylation sites in mf-psi-his

C



127

Fig. 42. Análisis bioinformático de los potenciales sitios de N- y O-glicosilación en la secuencia correspondiente al α -MF-PSI. (A) Secuencia aminoacídica completa codificada por el vector pKLAC1 al que se insertó la secuencia del PSI(His)₆. En la figura se indican las regiones aportadas por el vector (péptido señal y prodominio α -MF), el PSI y la cola de seis histidinas agregadas. Los residuos KR subrayados representan la secuencia de reconocimiento de la enzima Kex-2, encargada de clivar la porción pro- del α -MF previo a la secreción de la proteína recombinante. **(B)** Sitios de N-glicosilación predichos por el servidor NetNGlyc 1.0 Server (www.expasy.org/tools) sobre la secuencia completa mostrada en A. **(C)** Sitios putativos de O-glicosilación según servidor NetOGlyc 3.1 Server (www.expasy.org/tools).

AIV. Producción de las proteínas de secreción α -MF-PSI(His)₆ y α -MF-PSIN86S(His)₆

Una vez establecido el clon de cada transformante a utilizar y el tiempo de expresión mediante los ensayos en pequeña escala, se llevó a cabo la expresión de cada una de las formas de PSI en 250 ml de medio YPGal (4%) a 30 °C durante 72 hs.



Purificación de los productos de expresión α -MF-PSI(His)₆ y α -MF-PSIN86S(His)₆

Transcurrido el tiempo de expresión, el medio de cultivo se separó por centrifugación, se concentró por precipitación con (NH₄)₂SO₄ y se aplicó a una columna de exclusión molecular para la eliminación de contaminantes y agregados de alto peso molecular. A continuación, debido a la presencia de la cola de seis histidinas en el extremo C-terminal de cada una de las proteínas recombinantes, se hizo uso de la cromatografía de afinidad de iones inmovilizados (IMAC) en la cual los iones Ni⁺² de la resina interactúan con los anillos imidazólicos de los residuos de histidina. A su vez, esta marca permitió seguir cada paso de purificación por inmunodetección con anticuerpos anti-His.

De la Fig. 43 a la Fig. 45 se muestran los cromatogramas típicos obtenidos durante las etapas de purificación de los productos de expresión de las construcciones pKLAC- α -MF-PSI(His)₆ y pKLAC- α -MF-PSIN86S(His)₆ y los Western blots correspondientes.

A partir del análisis de las fracciones recolectadas durante la cromatografía de exclusión molecular de la forma del PSI derivada de la construcción pKLAC- α -MF-PSI(His)₆ (Fig. 43 A), se seleccionaron las fracciones 9 a 13 (Fig. 43 B), las cuales mostraron alto contenido de proteína. Al mismo tiempo, se tuvo en cuenta que el volumen de retención/elución de las fracciones se correspondiera con el rango de masas moleculares esperadas.

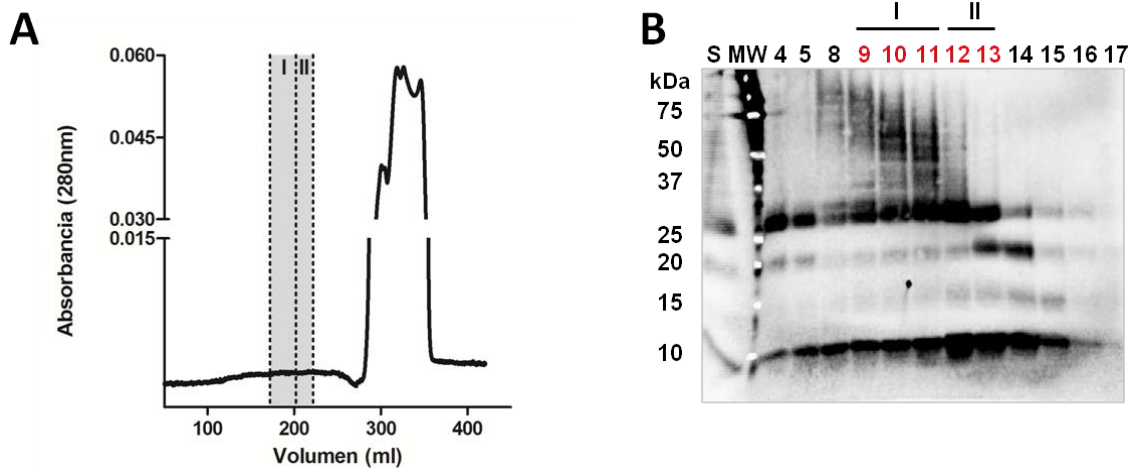
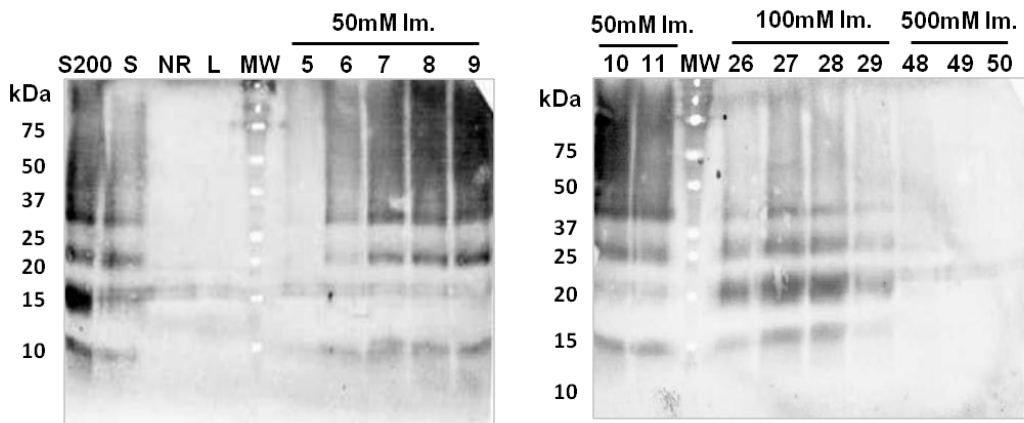
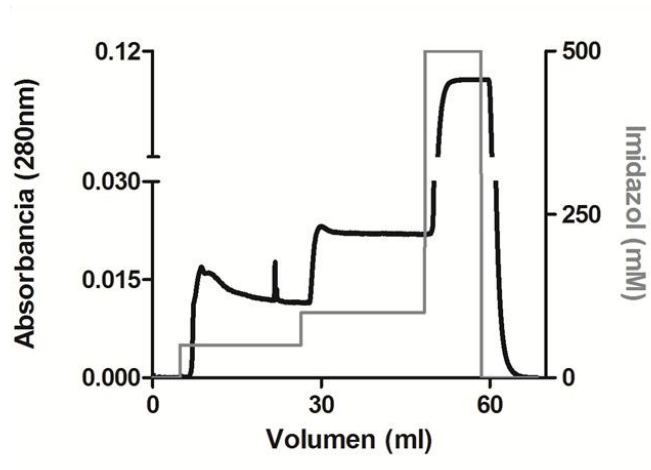
Cromatografía de exclusión molecular de la expresión del casete α -MF-PSI(His)₆

Fig. 43. Purificación parcial por exclusión molecular del producto obtenido en la expresión del casete conteniendo la secuencia α -MF-PSI(His)₆. El dominio PSI expresado y secretado por el clon 6 de *K. lactis* con el casete α -MF-PSI(His)₆ integrado en su genoma, se concentró desde el medio de cultivo por precipitación con $(\text{NH}_4)_2\text{SO}_4$ (90% de saturación). Luego, se sometió a un primer paso de purificación por exclusión molecular, tal como se describió en *Materiales y Métodos de PSI*. **(A)** Cromatografía de exclusión molecular de α -MF-PSI(His)₆ usando una columna Superdex 200 HiLoad 26/60. **(B)** Las fracciones recolectadas se analizaron por inmunodetección empleando anticuerpos anti-His: *Calle MW*, marcadores de masas moleculares; *Calle S*, muestra aplicada a la columna; *Calles 4 a 17*, fracciones eluidas en la respectiva cromatografía.

Se agruparon las fracciones 9 a 11 (*pool I* en la Fig. 43 A) y las 12 a 13 (*pool II* en la Fig. 43 A) de acuerdo al perfil observado para continuar con la purificación en forma separada, Fig. 44 A y 44 B, respectivamente. En ambas cromatografías de afinidad, en el primer paso de elución (50 mM de imidazol) eluyeron de la columna HisTrap las formas con mayor porcentaje de glicosilación, mientras que en el segundo paso de 100 mM se obtuvieron bandas más definidas. A su vez, en la cromatografía mostrada en la Fig. 44 B, se observa el efecto concentrador de la afinidad sobre las bandas intermediarias (23,5 kDa y 15 kDa). Las fracciones eluidas con 100 mM de imidazol se conservaron para posteriores análisis.

Cromatografía IMAC del producto de expresión del casete α -MF-PSI(His)₆.

A



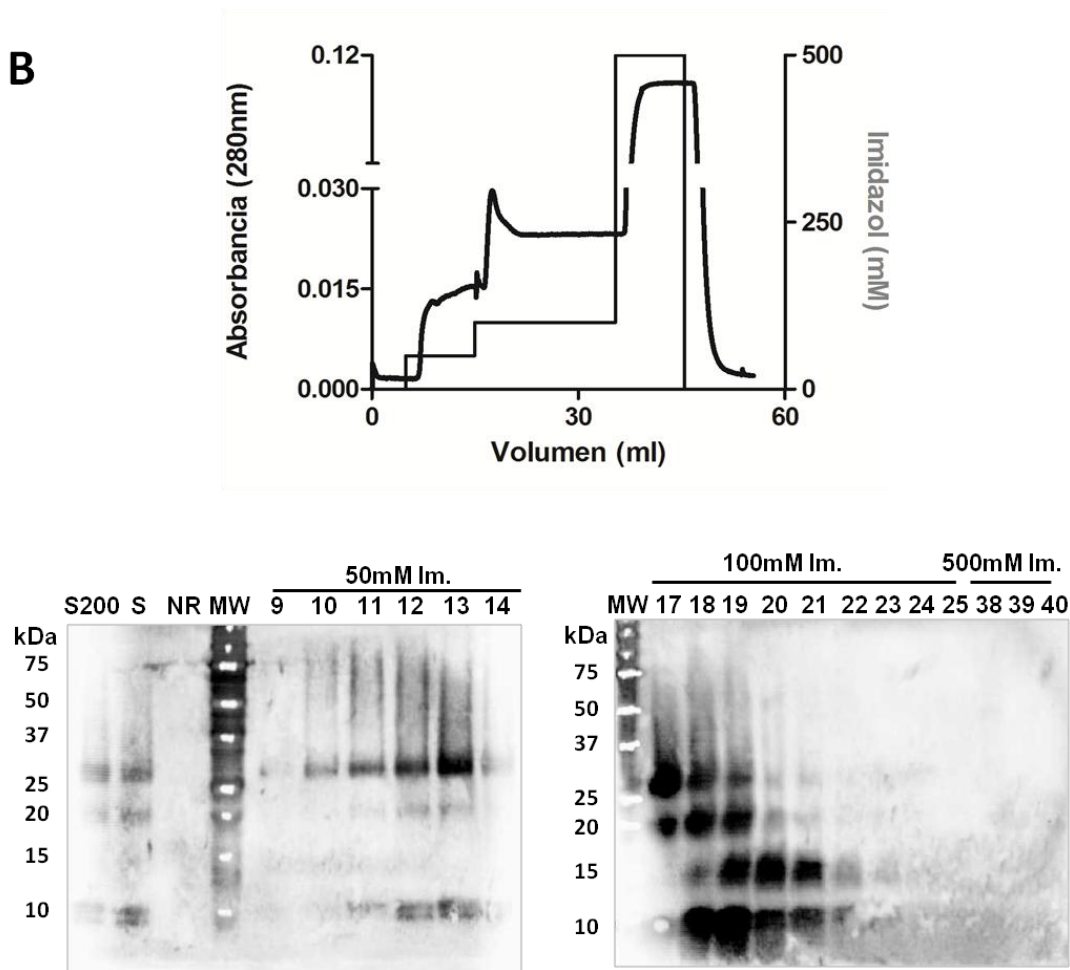


Fig. 44. Purificación por IMAC del producto de expresión del casete conteniendo la secuencia α -MF-PSI(His)₆ parcialmente purificado. A partir de las fracciones 9 a 13 eluidas de la columna Superdex 200 (S-200) se hicieron dos *pools*, cada uno de los cuales fue recromatografiado en una columna HisTrap. **(A)** Cromatografía IMAC del *pool* formado por las fracciones 9 a 11. **(B)** Cromatografía IMAC del *pool* compuesto por las fracciones 12 y 13. Las fracciones recolectadas en cada cromatografía se analizaron por inmunodetección empleando anticuerpos anti-His: *Calles MW*, marcadores de masas moleculares; *Calles S200*, pool proveniente de la columna S200; *Calles S*, muestra aplicada a la columna; *Calles NR*, fracción no retenida, proteínas eluidas durante la aplicación de la muestra; *Calles L*, fracción proteica retenida débilmente por la interacción con Ni⁺² eluida en el lavado de la columna; *Calles 4 a 40*, fracciones eluidas en la respectiva cromatografía. Se indica además, la concentración de imidazol del buffer de elución por grupo de fracciones.

Al purificar el mutante del sitio de glicosilación α -MF-PSIN86S(His)₆ por exclusión molecular (Fig. 45 A), se observó como era de esperar un menor grado de glicosilación y notablemente, que a los diferentes volúmenes de elución el perfil presenta aproximadamente bandas de la misma masa molecular. Esto sugeriría la presencia de formas agregadas en la muestra que se separan como consecuencia del tratamiento desnaturalizante previo al análisis electroforético.

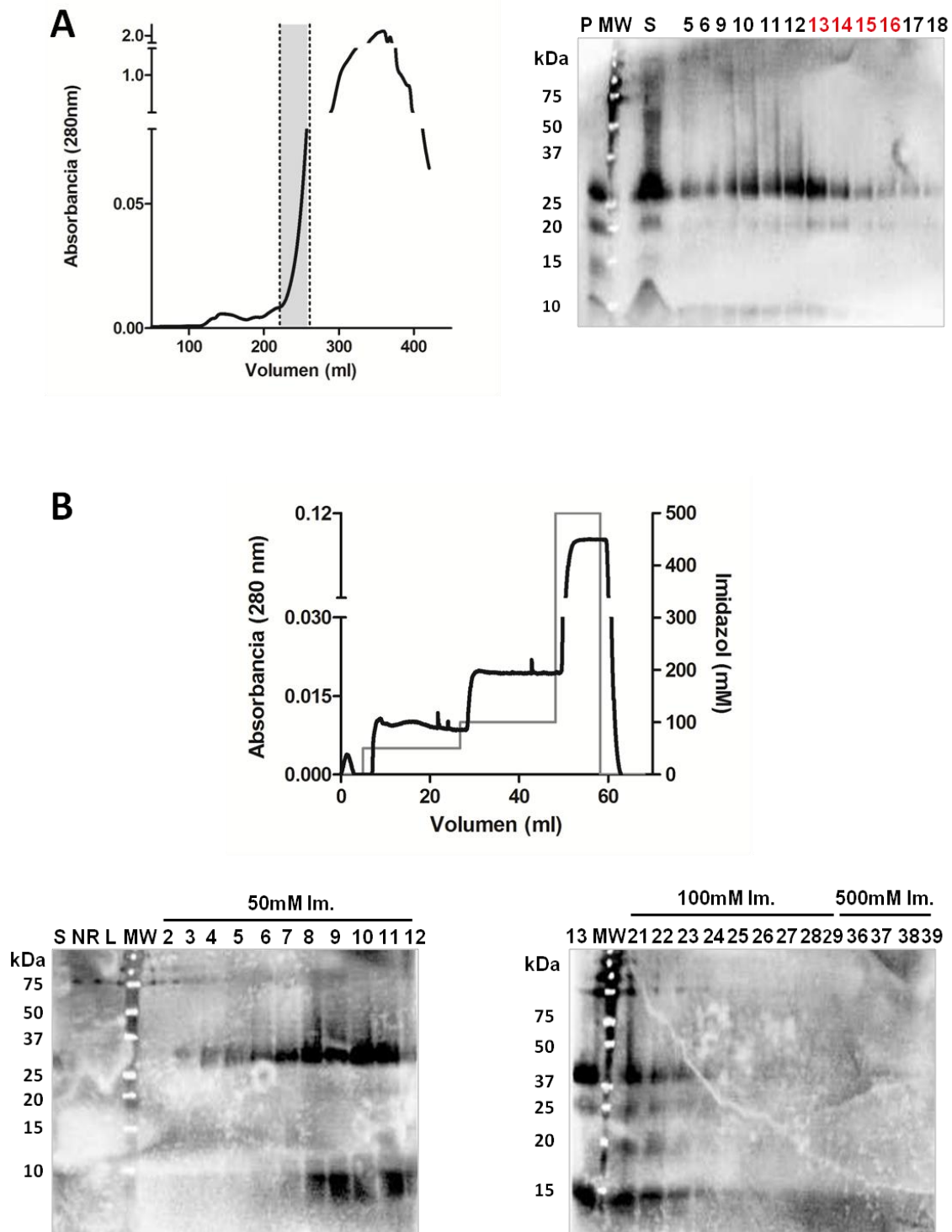
Purificación de la expresión del casete conteniendo la secuencia α -MF-PSIN86S(His)₆

Fig. 45. Purificación del producto obtenido en la expresión del casete conteniendo la secuencia α -MF-PSIN86S(His)₆. El mutante del sitio putativo de N-glicosilación del dominio PSI se expresó por el clon 4 de *K. lactis* con el casete α -MF-PSIN86S(His)₆ integrado en su genoma. La proteína recombinante secretada al medio de cultivo se concentró por precipitación con $(\text{NH}_4)_2\text{SO}_4$ (90% de saturación), purificó y analizó siguiendo el mismo procedimiento que en el caso del dominio *wild-type*. **(A)** Cromatografía de exclusión molecular de α -MF-PSIN86S(His)₆ usando una columna HiLoad 26/60 Superdex 200 y Western blot del eluato: *Calle MW*, marcadores de masas moleculares; *Calle P*, precipitado obtenido por ultracentrifugación

luego de la concentración; *Calles S*, muestra aplicada a la columna; *Calles 5 a 18*, fracciones eluidas en la respectiva cromatografía. **(B)** Las fracciones 13 a 16 provenientes de la exclusión fueron reunidas en un único *pool*, recromatografiadas por IMAC utilizando una columna HisTrap y las fracciones colectadas analizadas por Western blot: *Calles MW*, marcadores de masas moleculares; *Calles S*, muestra aplicada a la columna; *Calles NR*, fracción no retenida, proteínas eluidas durante la aplicación de la muestra; *Calles L*, fracción proteica retenida débilmente por la interacción con Ni^{+2} eluida en el lavado de la columna; *Calles 2 a 39*, fracciones eluidas de la columna HisTrap. Se indica además, la concentración de imidazol del buffer de elución por grupo de fracciones.

Se seleccionaron las fracciones 13 a 16, bajo el mismo criterio utilizado en la purificación de pKLAC- α -MF-PSI(His)₆, obteniéndose un perfil similar para los eluidos de 50 mM y 100 mM de imidazol que el descrito en el párrafo anterior (Fig. 45 B).



Corte in vitro del producto α -MF-PSI(His)₆ con la enzima Kex-2

A fin de obtener información que apoyase la hipótesis de la presencia del dominio α -MF en las proteínas secretadas por expresión de los casetes que contienen las secuencias α -MF-PSI(His)₆ y α -MF-PSIN86S(His)₆, las fracciones purificadas de α -MF-PSI(His)₆ se sometieron a digestión *in vitro* con la enzima Kex-2, responsable del corte *in vivo* del dominio α -MF. El resultado de esta digestión se muestra en la Fig. 46 donde se aprecia un descenso de la masa de las bandas 32 kDa, 24 kDa y 15 kDa de las muestras tratadas con Kex-2, tanto cuando las mismas se desnaturalizaron previamente a 95 °C como cuando se trataron en forma directa. Específicamente, las bandas mencionadas disminuyen su intensidad contribuyendo a la aparición de una única banda de 13 kDa. Las bandas restantes, de 9,5 kDa y 7 kDa, quedaron invariantes.

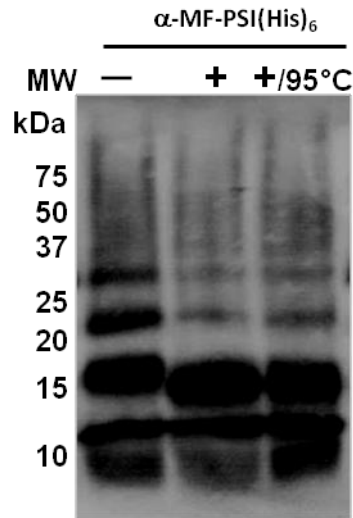
Tratamiento del producto de expresión del casete con la secuencia α -MF-PSI(His)₆ con Kex-2

Fig 46. Tratamiento *in vitro* de la proteína α -MF-PSI(His)₆ con la enzima Kex-2. Alícuotas de la proteína purificada se digirieron con la enzima comercial Kex-2 y se analizaron por Western blot: *Calle (-)*, alícuota sin agregado de Kex-2; *Calle (+)*, alícuota tratada con Kex-2; *Calle (+/95 °C)*, alícuota tratada con la enzima después de un ciclo de desnaturalización a 95°C y. En la figura se indican también los marcadores de masa molecular, (*MW*).



Análisis de los productos de expresión de los casetes conteniendo α -MF por técnicas de proteómica

Para confirmar la secreción del dominio α -MF junto a PSI(His)₆ y a PSIN86S(His)₆, fracciones purificadas de ambas proteínas fueron sometidas a digestión trípica *in situ* para ser analizadas por PMF MALDI-TOF MS. En todos los casos los espectros se adquirieron en dos ventanas de relación masa/carga, una estándar de 400 a 4000 m/z y una segunda ventana de 4000 a 10000 m/z, debido a que el fragmento que se corresponde con el dominio α -MF supera los 4000 Da, al no poseer sitios de corte internos para la enzima tripsina (Fig. 47 B).

En la Fig. 47 A se muestran las bandas electroforéticas analizadas de α -MF-PSI(His)₆, mientras que los espectros obtenidos con cada una de ellas se presentan en las Fig. 48, 49, 50 y 51 junto a los fragmentos de las secuencias a los que mediante el programa Sequence Editor de Bruker Daltonics, se les ha asignado un pico. Previo a la serie de espectros, se indica la secuencia aminoacídica de la proteína que se tuvo en cuenta para realizar los análisis (Fig. 47 B).

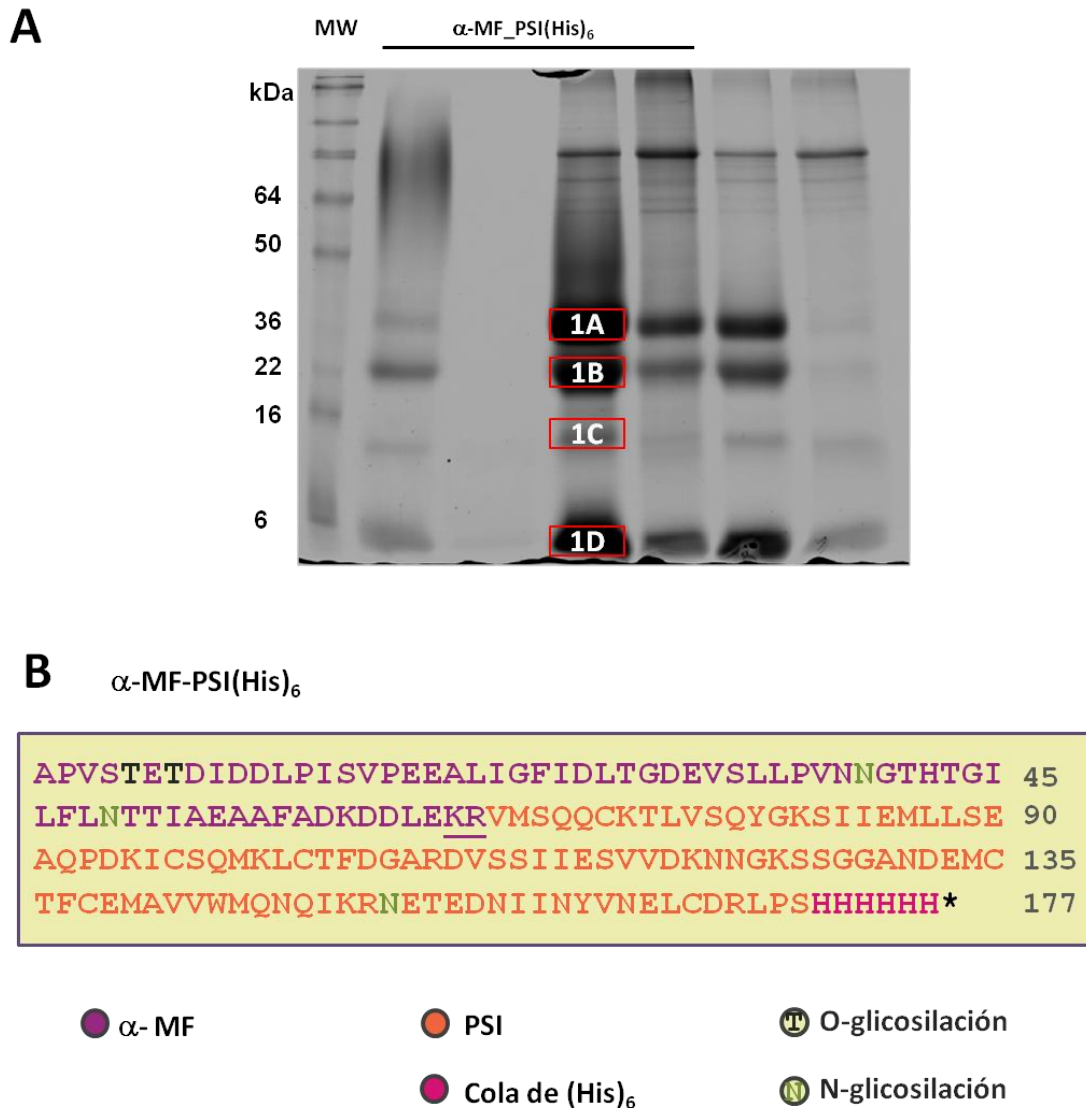
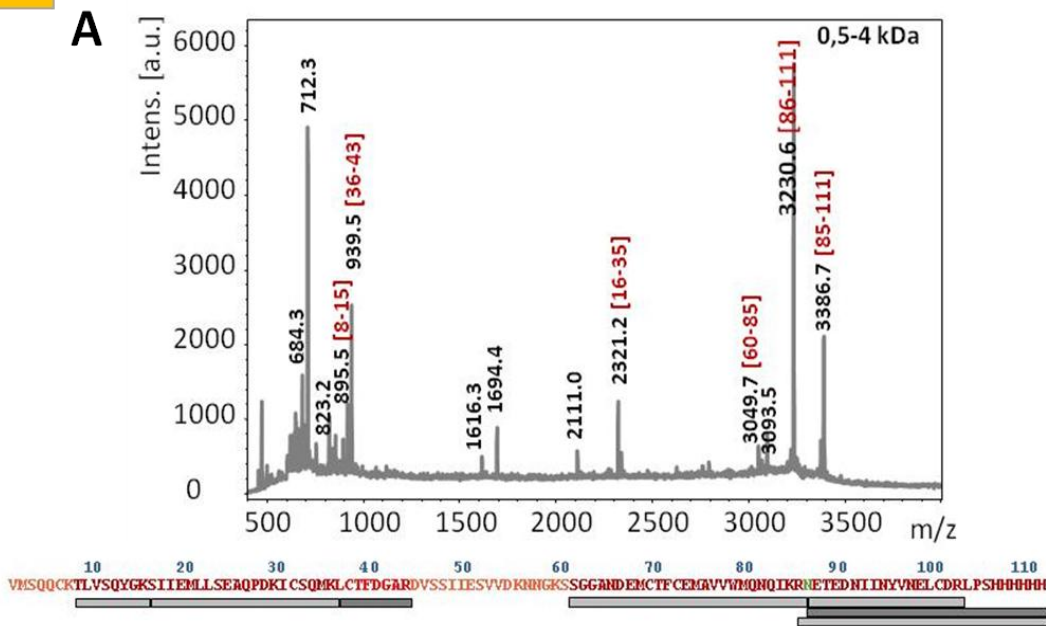
SDS-PAGE del producto purificado de la expresión del casete α -MF-PSI(His)₆ en *K. lactis*

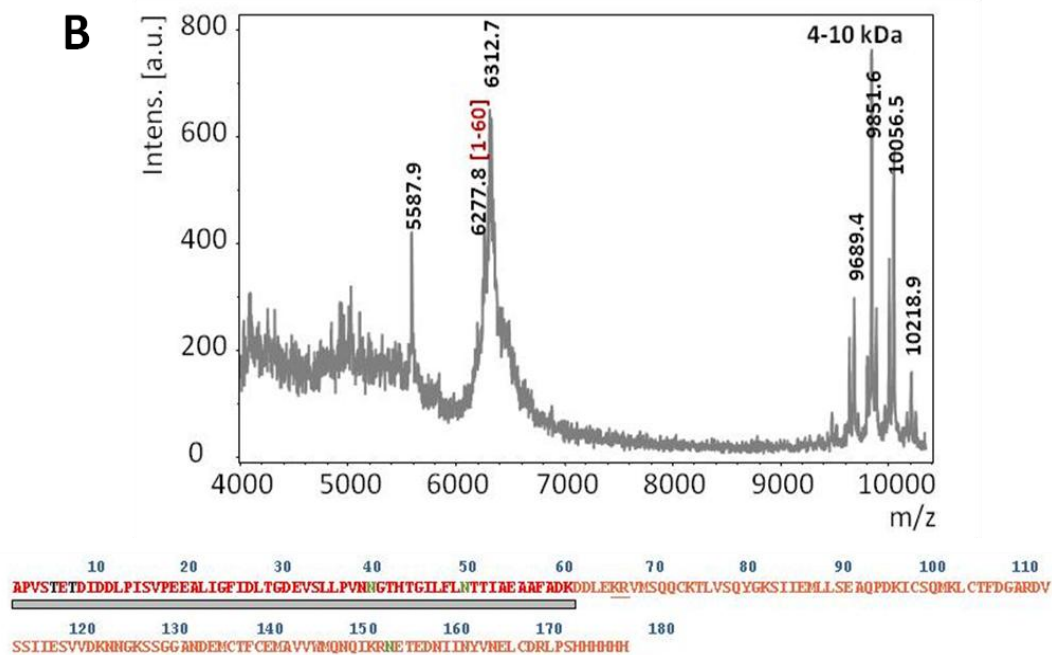
Fig. 47. Muestra obtenida por expresión del casete α -MF-PSI(His)₆ en *K. lactis* y purificada para su análisis por técnicas de proteómica. (A) SDS-PAGE de la muestra, sobre los geles se indican las bandas sometidas a digestión tréptica y analizadas por espectroscopía de masas MALDI-TOF. (B) Secuencia considerada en el análisis de la proteína, en la misma se resaltan el dominio PSI, el α -MF incluyendo el sitio *Xho*I (LE) y el sitio de clivaje de la enzima Kex-2 (KR). Se indican también la cola de polihistidinas y los sitios putativos de O- y N-glicosilación.

PMF de la banda proteica 1A correspondiente a la expresión del casete α -MF-PSI(His)₆

1A



Cobertura de intensidad / secuencia MS: 44,4 % / 79,3 %



Cobertura de intensidad / secuencia MS: 5,4 % / 33,9 %

Fig. 48. Análisis por técnicas de proteómica de la banda proteica 1A de la muestra obtenida mediante expresión del casete α -MF-PSI(His)₆ en *K. lactis*. Espectros adquiridos en dos ventanas de relación masa/carga, (A) de 400 a 4000 m/z y (B) de 4000 a 10000 m/z, sobre cada pico se indica en negro la masa molecular correspondiente y en rojo el intervalo de secuencia reconocido. Debajo del espectro se muestra la secuencia de la proteína considerada y por medio de barras se señalan los péptidos resultantes de la digestión trípica teórica cuya masa molecular se correlaciona con el valor de un pico del espectro. La escala de grises de las barras representa la intensidad de los picos.

Como se muestra en la Fig. 48 A, en función de los valores de cobertura de intensidad y de secuencia (44,4 % y 79,3%, respectivamente) y de la buena distribución que se observan en el primer espectro correspondiente a la banda 1A, se puede afirmar que en la misma se identifica PSI. Dentro de los fragmentos identificados quedó incluido el sitio de glicosilación NET del PSI, así como la cola de poliHis en el extremo C-terminal.

A partir del segundo espectro de esta banda (Fig. 48 B) es posible asignar el pico de masa 6277,87 Da al fragmento de 60 residuos de aminoácidos esperado del dominio α -MF (que representa el 33,9 % de la secuencia total). Así mismo, se observa una serie de picos con diferencias de masas entre ellos de 162 Da (a partir de 9 kDa) que podría deberse a glicosilación heterogénea. Por ende, se concluye que la levadura *K. lactis* expresó la proteína PSI(His)₆ a partir de la construcción pKLAC- α -MF-PSI(His)₆ secretándola hacia el medio extracelular, pero resultó incapaz de procesar el α -MF de la manera indicada en el manual del sistema de expresión.

Resultados similares se observaron en los espectros que surgen de la digestión de la banda 1B (Fig. 49). Aunque el pico de 6277,87 Da está ausente en el espectro de la Fig. 49 B, se observa el patrón de glicosilación que indicaría que en esta banda el α -MF se encuentra completamente glicosilado. Desprendiéndose así las mismas conclusiones que para la banda 1A. La diferencia de movilidad entre ambas bandas electroforéticas podría deberse a diferentes grados de glicosilación.

PMF de la banda proteica 1B correspondiente a la expresión del casete α -MF-PSI(His)₆

1B

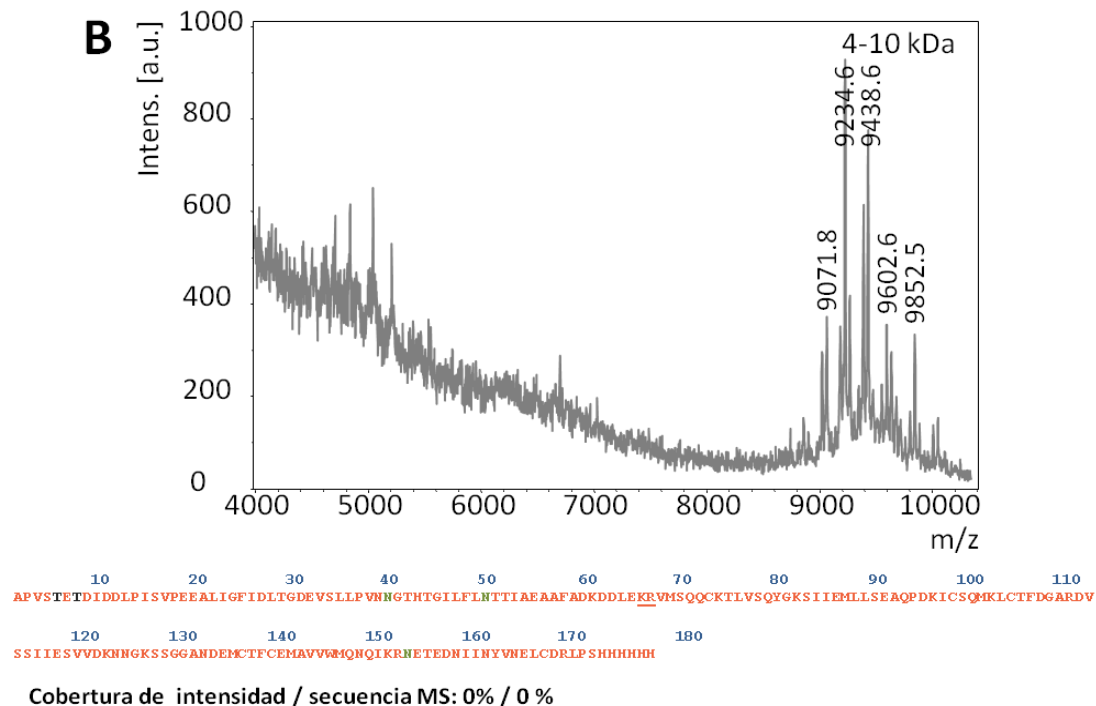
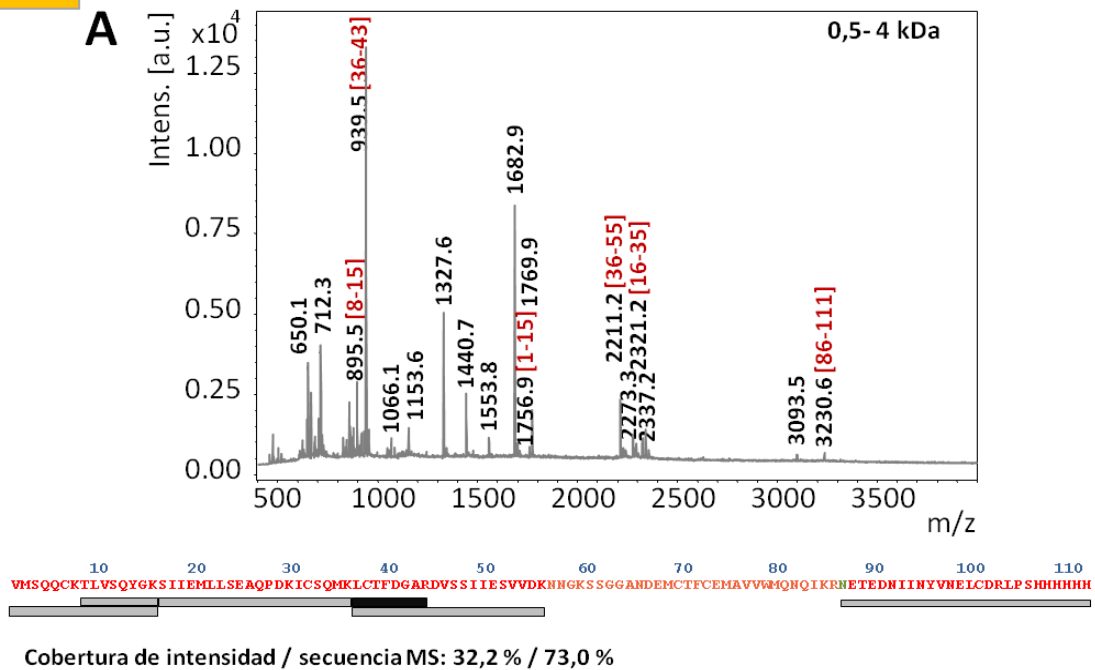
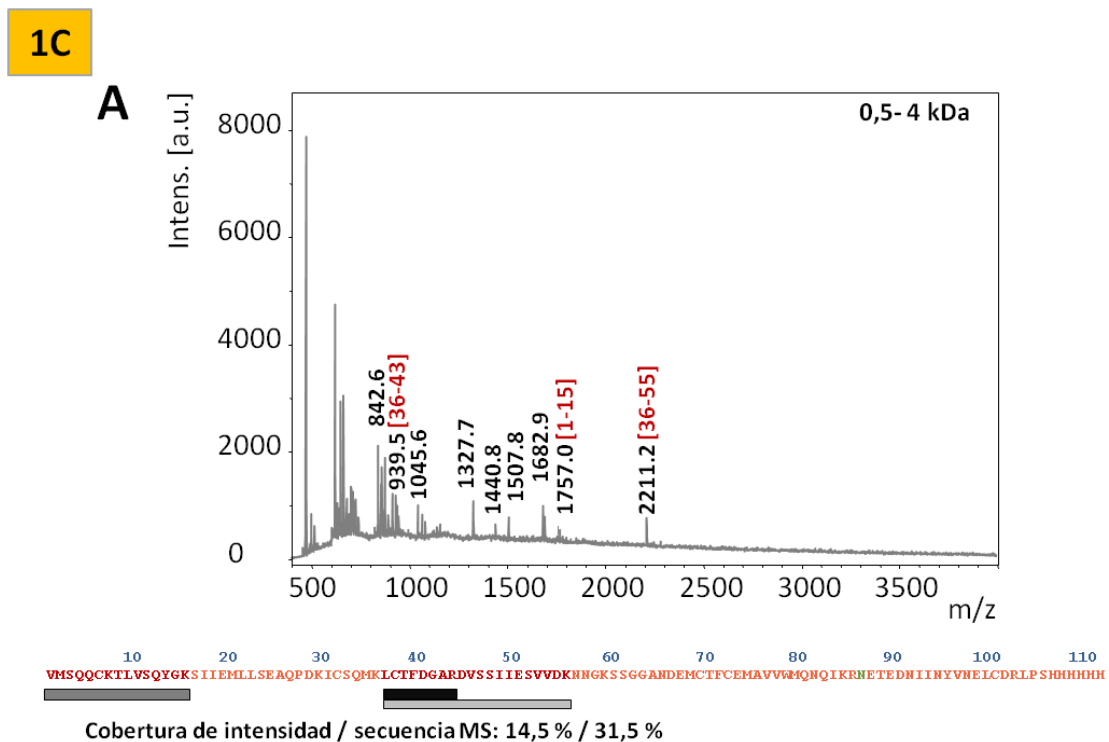


Fig. 49. Análisis por técnicas de proteómica de la banda proteica 1B de la muestra obtenida la expresión del casete α -MF-PSI(His)₆ en *K. lactis*. Espectros adquiridos en dos ventanas de relación masa/carga, (A) de 400 a 4000 m/z y (B) de 4000 a 10000 m/z, sobre cada pico se indica la masa molecular correspondiente. Debajo del espectro se muestra la secuencia de la proteína considerada y por medio de barras se señalan los péptidos resultantes de la digestión trípica teórica cuya masa molecular se correlaciona con el valor de un pico del espectro. La escala de grises de las barras representa la intensidad de los picos.

En los espectros adquiridos para la digestión de la banda 1C (Fig. 50 A) se identificó claramente el PSI(His)₆. Sin embargo, no se vieron indicios de la presencia del α -MF sin glicosilar (pico de 6277,87 Da), ni de este fragmento glicosilado (patrón de picos distanciados en 162 Da en el rango de altos pesos moleculares). Pudiéndose concluir de este modo, que esta banda minoritaria corresponde al PSI(His)₆ correctamente procesado.

PMF de la banda proteica 1C correspondiente a la expresión del casete α -MF-PSI(His)₆

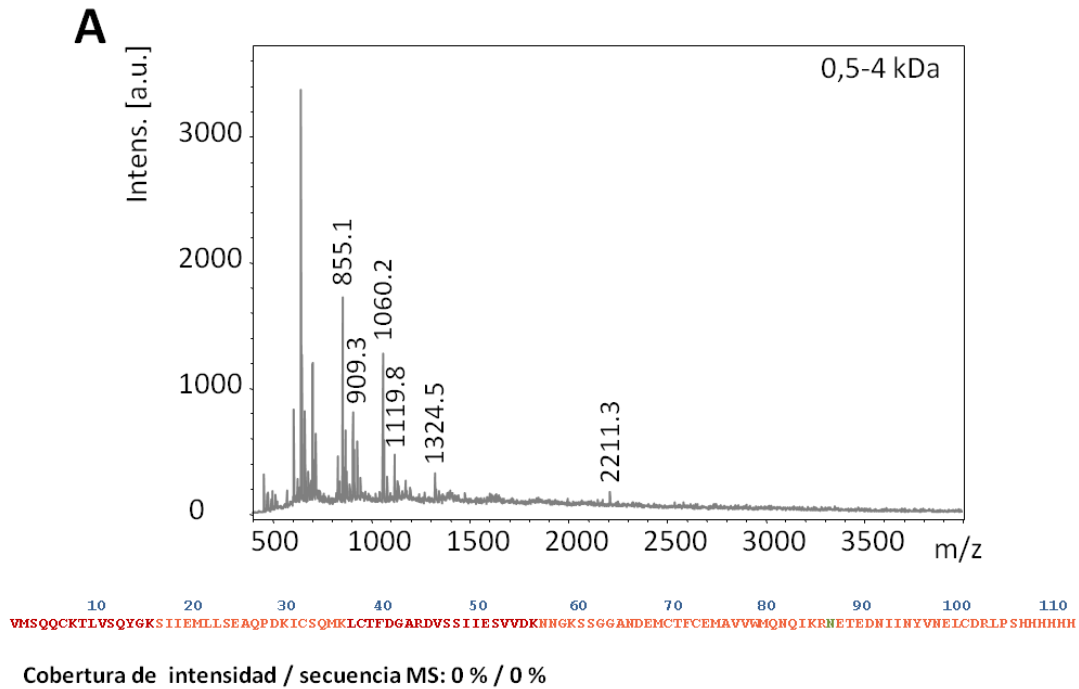


B En la ventana 4-10 kDa no se observaron picos.

Fig. 50. Análisis por técnicas de proteómica de la banda proteica 1C de la muestra obtenida mediante la expresión del casete α -MF-PSI(His)₆ en *K. lactis*. Espectros adquiridos en dos ventanas de relación masa/carga, (A) de 400 a 4000 m/z y (B) de 4000 a 10000 m/z, sobre cada pico se indica la masa molecular correspondiente. Debajo del espectro se muestra la secuencia de la proteína considerada y por medio de barras se señalan los péptidos resultantes de la digestión triptica teórica cuya masa molecular se correlaciona con el valor de un pico del espectro. La escala de grises de las barras representa la intensidad de los picos.

PMF de la banda proteica 1D correspondiente a la expresión del casete α -MF-PSI(His)₆

1D



B En la ventana 4-10 kDa no se observaron picos.

140

Fig. 51. Análisis por técnicas de proteómica de la presencia del α -MF en la banda proteica 1D de la muestra obtenida mediante la expresión del casete α MF-PSI(His)₆ en *K. lactis*. Espectros adquiridos en dos ventanas de relación masa/carga, (A) de 400 a 4000 m/z y (B) de 4000 a 10000 m/z, sobre cada pico se indica la masa molecular correspondiente. Debajo del espectro se muestra la secuencia de la proteína considerada y por medio de barras se señalan los péptidos resultantes de la digestión triptica teórica cuya masa molecular se correlaciona con el valor de un pico del espectro. La escala de grises de las barras representa la intensidad de los picos.

Con respecto a los resultados obtenidos para la banda 1D (Fig. 51), el único pico que podría asignarse al PSI coincide con un pico característico de autodigestión de tripsina. Por ello y siendo que es el único pico asignado, se concluyó que no hay evidencia suficiente para identificar PSI en dicha banda.

La Fig. 52 A presenta las bandas electroforéticas analizadas de α -MF-PSIN86S(His)₆, mientras que en la Fig. 52 B se indica la secuencia aminoacídica de la proteína considerada en el análisis de los espectros.

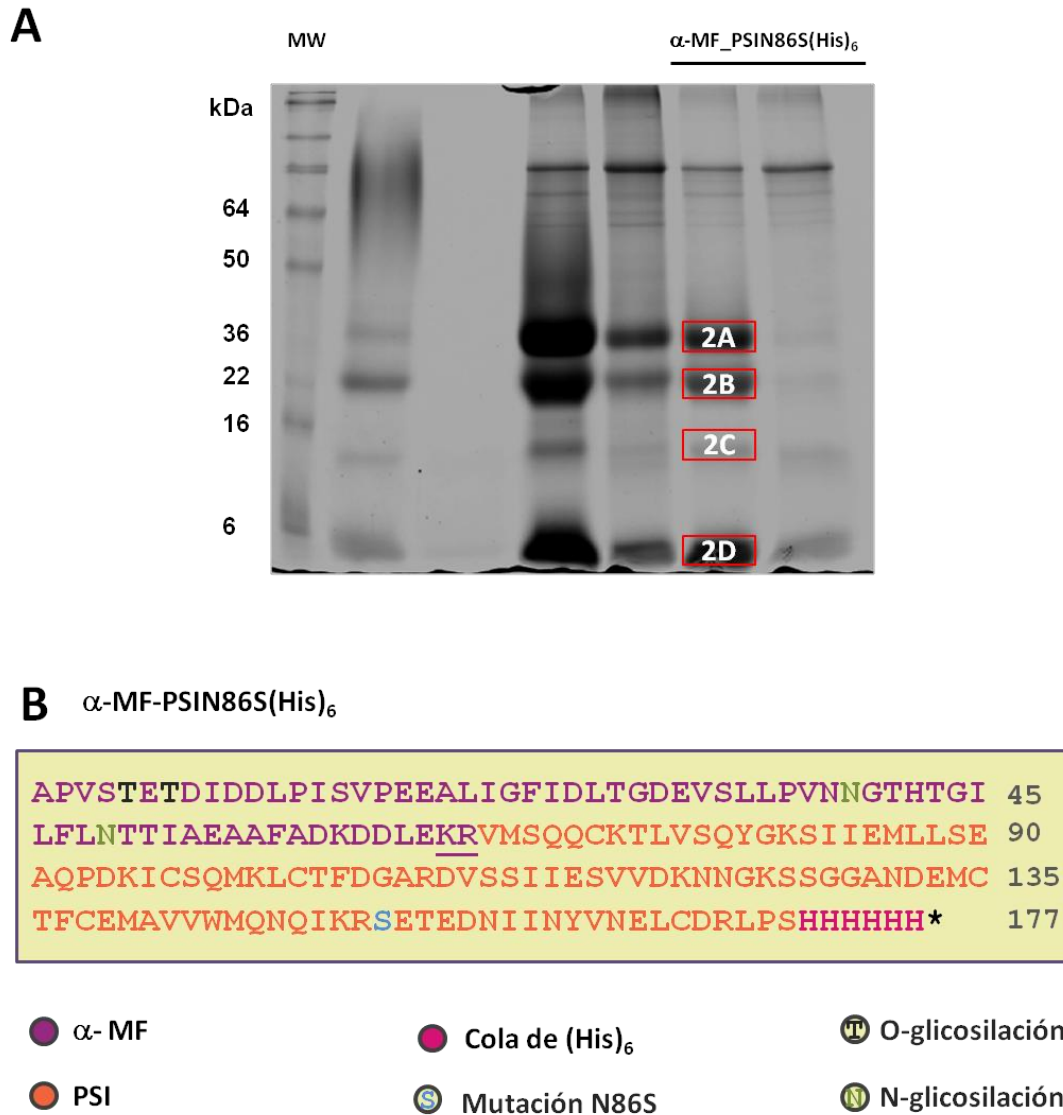
SDS-PAGE del producto purificado de la expresión del casete α -MF-PSIN86S(His)₆ en *K. lactis*

Fig. 52. Muestra obtenida por expresión del casete α -MF-PSIN86S(His)₆ en *K. lactis* y purificada para su análisis por técnicas de proteómica. (A) SDS-PAGE de la muestra, sobre los geles se indican las bandas que han sido sometidas a digestión triptica y posterior espectroscopía de masas MALDI-TOF. (B) Secuencia considerada en el análisis de la proteína, en la misma se resaltan la mutación N86S, el dominio PSI, el α -MF incluyendo el sitio *Xho*I (LE) y el sitio de corte de la enzima Kex-2 (KR), la cola de polihistidinas y los sitios putativos de O- y N-glicosilación.

En la Fig. 53 y 54 se muestran los datos obtenidos para las bandas 2A y 2C de la muestra α -MF-PSIN86S(His)₆. En función de los valores de cobertura de intensidad y secuencia mostrados en la Fig. 53 A para la banda 2A (67,0 % y 55,9 %, respectivamente) y de los dos fragmentos que incluyen la mutación N86S (3203,8 y 3360,0 Da), se pudo identificar la presencia dePSIN86S(His)₆ en esta banda. En este caso, a partir del espectro registrado en la

ventana de altos pesos moleculares (Fig. 53 B) no se pudo asignar el pico de masa 6277,87 Da al fragmento esperado del α -MF; a pesar de ello, aún se observa una serie de picos con diferencias de masas de 162 Da a partir de los 9 kDa que podría deberse, como ya se sugirió, a la glicosilación heterogénea sobre los sitios putativos predichos en la secuencia del dominio α -MF por NetNGlyc 1.0 Server y NetOGlyc 3.1 Server. Una vez más, se concluye que *K. lactis* expresó la proteína mutada PSIN86S(His)₆ aunque sin el procesamiento correcto del α -MF a lo largo de la vía de secreción, encontrándose en el medio extracelular la proteína de fusión α -MF-PSIN86S(His)₆.

Los espectros correspondientes a las bandas 2B, 2C y 2D de la proteína α -MF-PSIN86S(His)₆ no se pudieron analizar, ya sea por no contar con un suficiente número de picos asignables a la secuencia o por verse únicamente señal de la matriz utilizada y picos de polímeros. Por tanto, las mismas no pudieron ser identificadas. A modo de ejemplo, en la Fig. 54 se muestra el espectro obtenido para la banda 2C.

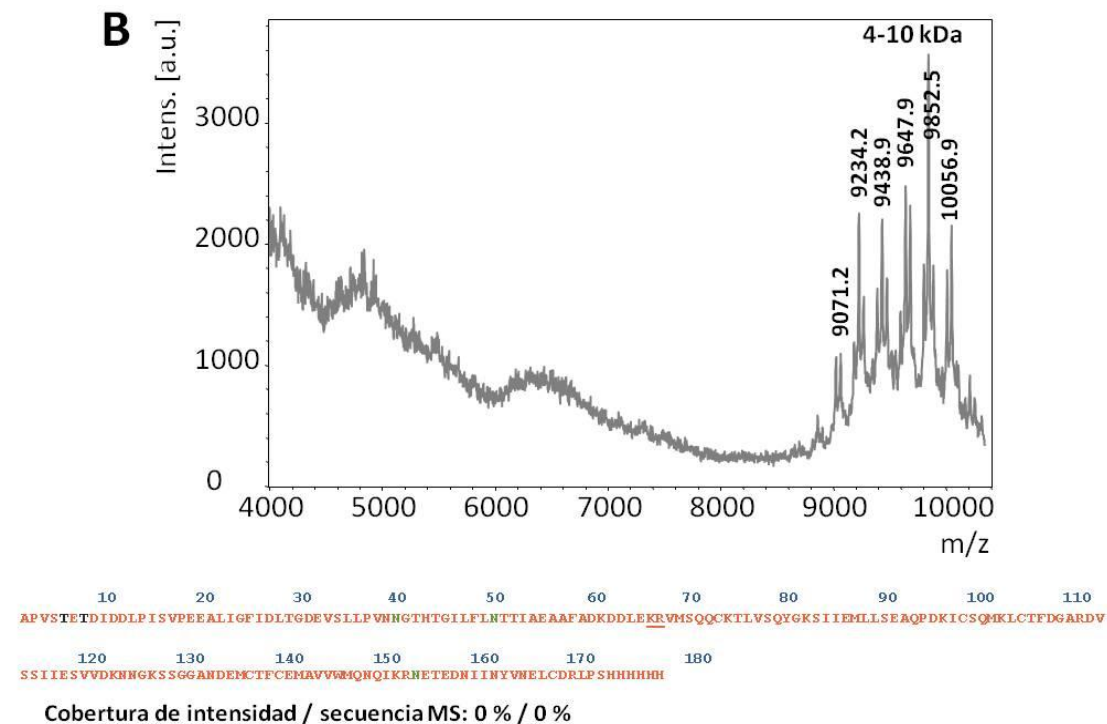
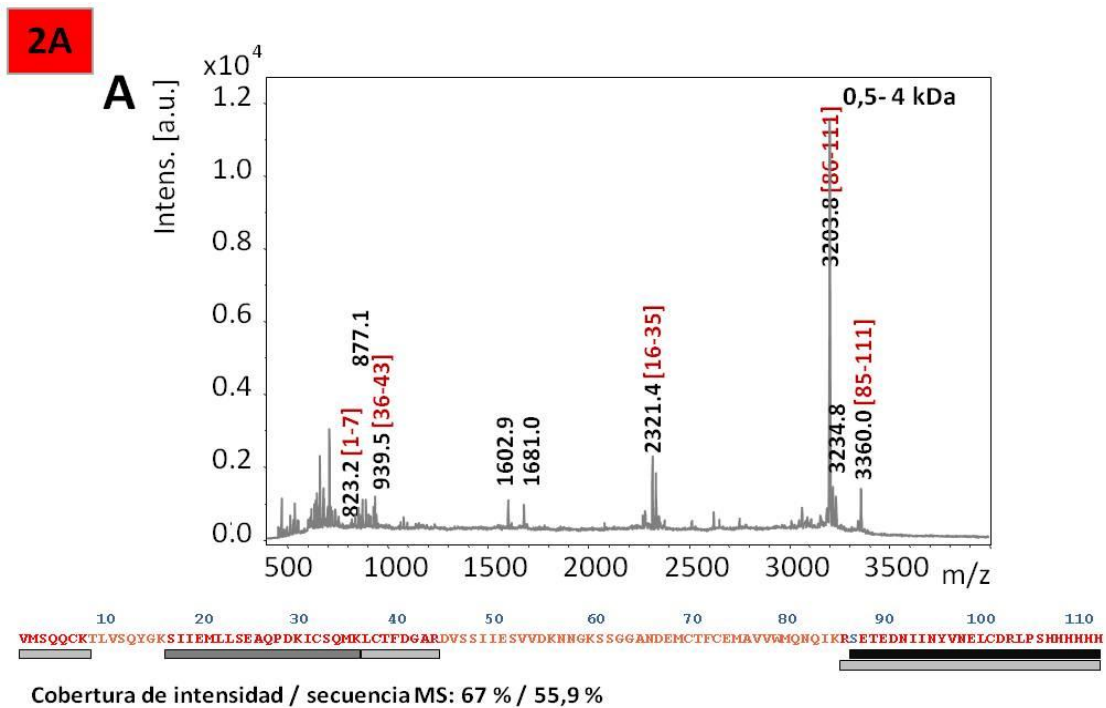
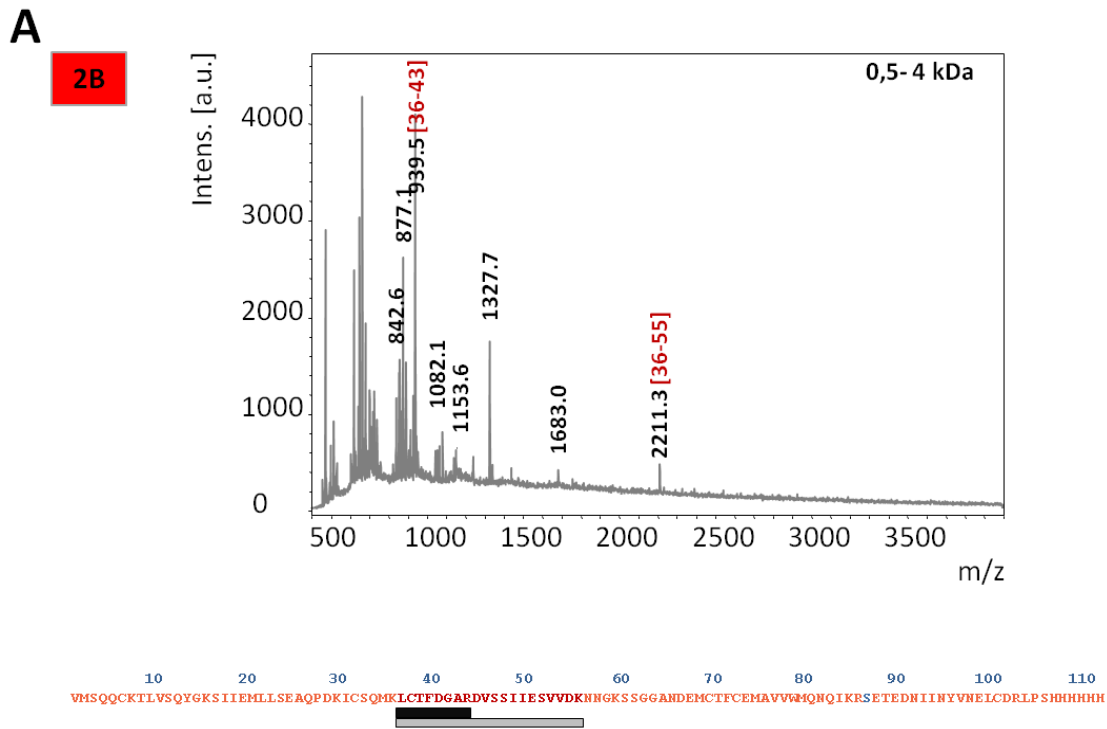
PMF de la banda proteica 2A correspondiente a la expresión del casete α -MF-PSIN86S(His)₆

Fig. 53. Análisis por técnicas de proteómica de la banda proteica 2C de la muestra obtenida mediante la expresión del casete α -MF-PSIN86S(His)₆ en *K. lactis* Espectros adquiridos en dos ventanas de relación masa/carga, **(A)** de 400 a 4000 m/z y **(B)** de 4000 a 10000 m/z, sobre cada pico se indica la masa molecular correspondiente. Debajo del espectro se muestra la secuencia de la proteína considerada y por medio de barras se señalan los péptidos resultantes de la digestión trípica teórica cuya masa molecular se correlaciona con el valor de un pico del espectro. La escala de grises de las barras representa la intensidad de los picos.

PMF de la banda proteica 2C correspondiente a la expresión del casete α -MF-PSIN86S(His)₆

Cobertura de intensidad / secuencia MS: 21,7 % / 18,0 %

B**2B**

En la ventana 4-10 kDa no se observaron picos.

144

Fig. 54. Análisis por técnicas de proteómica de la banda proteica 2A de la muestra obtenida mediante la expresión del casete α -MF-PSIN86S(His)₆ en *K. lactis*. Espectros adquiridos en dos ventanas de relación masa/carga, **(A)** de 400 a 4000 m/z y **(B)** de 4000 a 10000 m/z. Debajo del espectro se muestra la secuencia de la proteína considerada.

B. Expresión de las proteínas PSI(His)₆ y PSIN86S(His)₆ usando una estrategia alternativa en *K. lactis*

Si bien la estrategia presentada en la sección A es la indicada por el fabricante para la obtención de proteínas con su N-terminal nativo en el sistema *K. lactis*, la inesperada presencia del dominio α -MF en las proteínas PSI y PSIN86S, requeriría su corte *in vitro* mediante el uso de la proteasa Kex-2. Este paso adicional haría que el proceso de obtención sea muy poco rentable ya que además de aumentar los costos en la obtención de las proteínas recombinantes, reduciría los rendimientos. Debido a esto, se planteó una estrategia alternativa para producir PSI y PSIN86S en este sistema.

Aunque la falta de un correcto procesamiento podría deberse a una baja eficiencia de corte por parte de la enzima Kex-2, en la elección de la estrategia alternativa se contempló la posibilidad de que el dominio PSI pudiera ser secretado fuera de la célula de *K. lactis* aún en ausencia del factor de secreción propio del sistema. Por consiguiente, se adoptó como estrategia alternativa para producir PSI y PSIN86S, la eliminación de la secuencia codificante para el α -MF de las construcciones en pKLAC1. Se optó a su vez por mantener la secuencia del péptido señal del vector que asegure la entrada de la proteína heteróloga al retículo endoplasmático, simulando así la acción de la región pre- de preprocirsina, entorno natural donde se encuentra insertado el dominio PSI.

Se tuvo en cuenta además, la influencia de la cola de polihistidinas incluídas originalmente en el extremo C-terminal de las proteínas heterólogas durante la primera estrategia empleada. Para llevar a cabo este análisis, se realizó un modelado dinámico del PSI en presencia de las seis histidinas terminales. Alternativamente, se realizó el modelado considerando los residuos remanentes del sitio de corte de la enzima TEV luego de la digestión de la proteína de fusión cuando el mismo se ha localizado entre el C-terminal nativo del PSI y la cola de polihistidinas. Ambos modelos se compararon con el generado para la secuencia del PSI sin ninguna cola C-terminal. En función de los resultados obtenidos en este análisis, los cuales se muestran a continuación, se optó por incorporar el sitio de reconocimiento para la enzima TEV a fin de eliminar las histidinas terminales utilizadas en la detección y purificación de las proteínas recombinantes PSI y su mutante N86S.



Análisis por dinámica molecular de la influencia de la cola de polihistidinas terminal en la estructura tridimensional del dominio PSI

Para predecir la influencia de la cola de polihistidinas terminal en la estructura tridimensional del dominio PSI, se construyeron modelos estructurales de las proteínas PSI, PSI(His)₆ y PSI-TEV (PSI-Glu-Asn-Leu-Tyr-Phe-Gln) mediante modelado por homología. Se utilizó como referencia la estructura cristalina del dominio StAps-PSI (código PDB: 3RF1) de la PA de *Solanum tuberosum*, StAP-1, ya que ésta es la única estructura del dominio PSI aislado—sin formar parte de la proenzima—disponible en base de datos. A partir de ésta, el mejor modelo para el PSI de cirsina generado por el servidor SWISS MODEL arrojó un valor de identidad de secuencia de 38.61% (E-value: 8×10^{-21}) y una calificación QMEAN4 de 0,48 (Z-score -2,99), resultando en un modelo estructural de buena calidad (Benkert et al., 2011).

Esta geometría inicial posteriormente se modeló mediante el programa Amber 11, con los parámetros descritos en la sección correspondiente de materiales y métodos de este capítulo. En el caso de las estructuras de las moléculas de PSI conteniendo la cola de histidinas o el sitio de reconocimiento para la enzima TEV en el extremo C-terminal, se agregaron primero los residuos aminoácidos adicionales con el programa PyMol. Luego sobre cada modelo se realizaron las mismas minimizaciones, equilibraciones y producciones de dinámica molecular de al menos 16 ns, salvando imágenes de las coordenadas cada 1000 pasos.

La comparación de los modelos obtenidos predice un cambio significativo desde una conformación de *boomerang* abierto (forma de V) hacia una conformación más cerrada con forma de C del PSI conteniendo la cola de polihistidinas, la cual en el modelo final se localiza hacia el interior de la cavidad hidrofóbica formada por las hélices α en la estructura del PSI (Fig. 55 A y Anexo 2). La distorsión ocasionada por la cadena de histidinas genera una tensión tal que produce la ruptura de un enlace (indicado en la Fig. 55 mediante una flecha negra). En la Fig. 55 B se observa la posición que adquiere la cola de histidinas transcurridos 16 ns, respecto a la posición inicial del ensayo de dinámica molecular.

De forma contraria, la superposición de las estructuras predichas por el modelado para el PSI original y el PSI con los residuos ENLYFQ (Fig. 56 A), muestra una baja distorsión de la conformación del PSI por el agregado de estos aminoácidos en el C-terminal de la proteína. La Fig. 56 B permite ver que si bien la secuencia de reconocimiento de la enzima TEV tiende a adoptar una estructura secundaria durante el tiempo de la dinámica, a diferencia de la cadena de histidinas, esta secuencia no se desplaza hacia el interior de la estructura del PSI, disminuyendo de este modo los cambios conformacionales de la molécula.

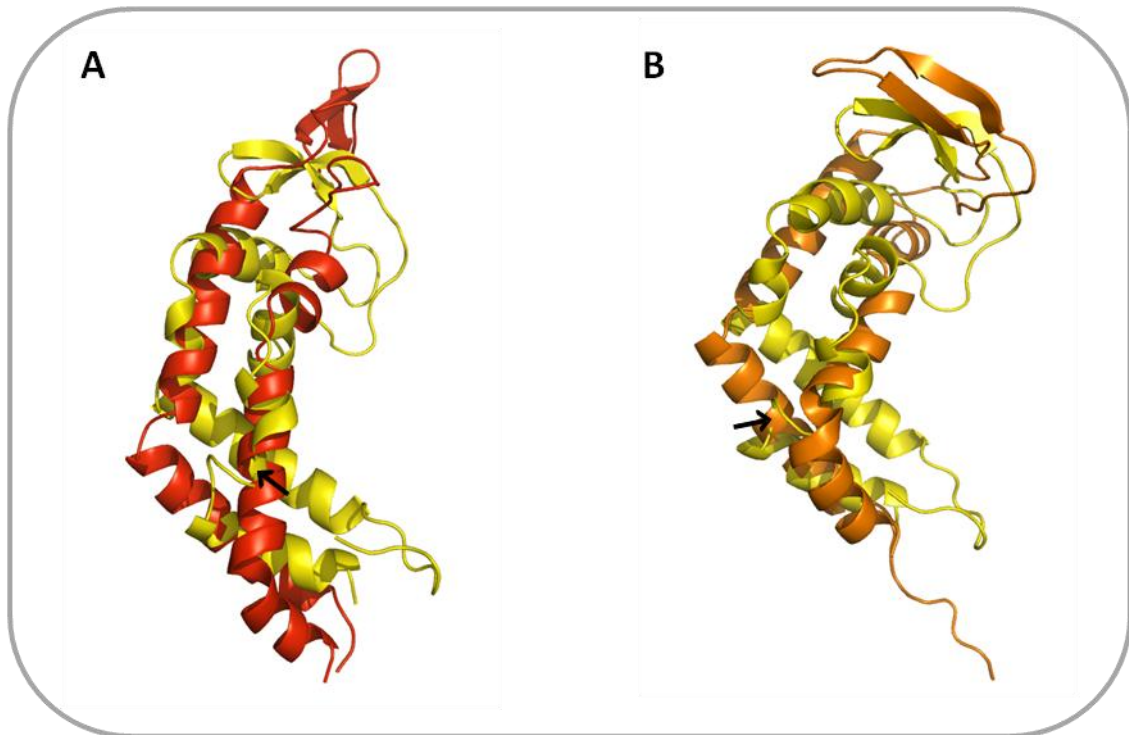
Influencia de la cadena (His)₆ en la conformación del dominio PSI

Fig. 55. Modelos estructurales de las proteínas PSI y PSI(His)₆ obtenidos por dinámica molecular. La secuencia del dominio PSI de procirsina y de PSI(His)₆ fueron modeladas por homología empleando como referencia la estructura cristalográfica de StAps-PSI presente en la PA StAp-1 de *S. tuberosum* (código PDB: 3RF1). **(A)** Superposición de los modelos estructurales obtenidos para el PSI (rojo) y para el PSI(His)₆ (amarillo). **(B)** Superposición del modelo inicial a tiempo cero (naranja) y del modelo final obtenido por producción de dinámica molecular transcurridos 16 ns (amarillo). Mediante una flecha negra se indica el enlace que resulta escindido en la molécula PSI(His)₆ por distorsión de la conformación original.

Influencia de los residuos ENLYFQ en la conformación del dominio PSI

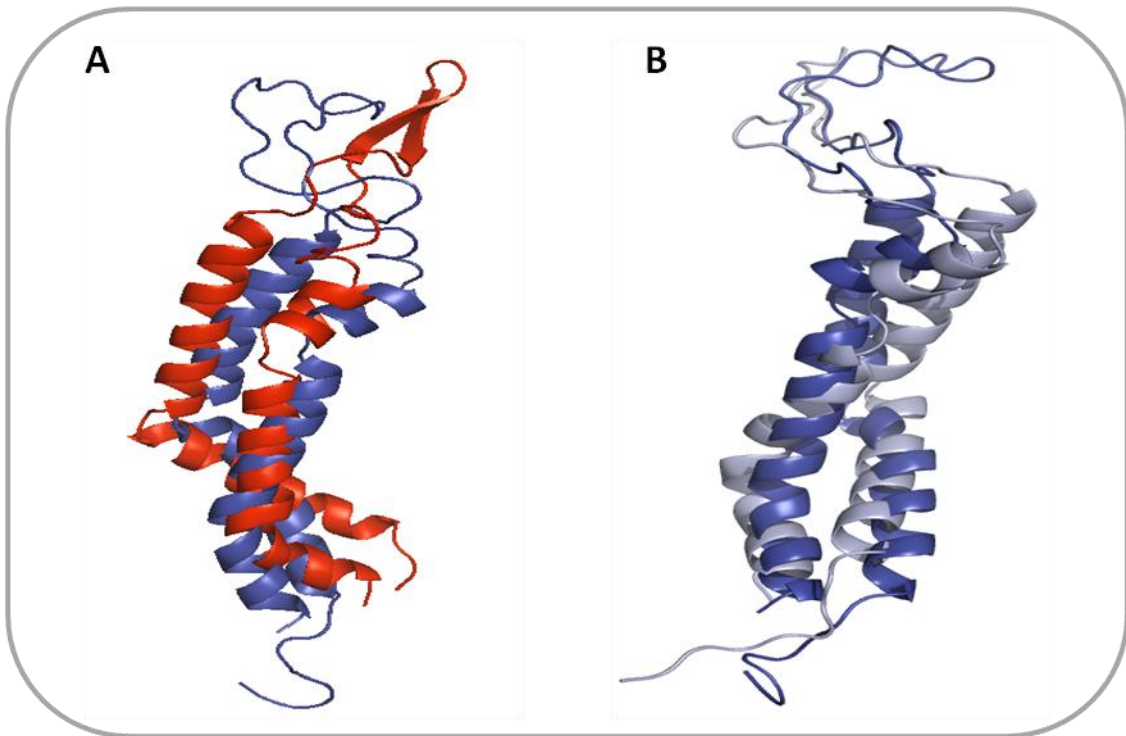


Fig. 56. Modelos estructurales de las proteínas PSI y PSI-TEV obtenidos por dinámica molecular. La secuencia del dominio PSI de procirsina y de PSI-TEV fueron modeladas por homología empleando como referencia la estructura cristalográfica de StAps-PSI presente en la PA StAp-1 de *S. tuberosum* (código PDB: 3RF1). **(A)** Superposición de los modelos estructurales obtenidos para el PSI (rojo) y para el PSI-TEV (azul). **(B)** Superposición del modelo inicial a tiempo cero (celeste) y del modelo final obtenido por producción de dinámica molecular transcurridos 16 ns (azul).

A partir de este análisis, se decidió obtener las construcciones que contuvieran al inserto PSI o PSIN86S, inmediatamente *downstream* a la secuencia del péptido señal del vector pKLAC1 e incorporar las bases que codifican para el sitio de reconocimiento de la enzima TEV entre el extremo 5' de la secuencia del inserto y la secuencia para el cola de seis histidinas: construcciones pKLAC-PSI(His)₆ y pKLAC-PSIN86S(His)₆.

BI. Obtención de los vectores pKLAC-PSI(His)₆ y pKLAC-PSIN86S(His)₆ para la expresión del PSI y su mutante de sitio de glicosilación en *K. lactis*

La obtención de estas construcciones requirió de la eliminación previa de la secuencia del α -MF y del sitio de reconocimiento de la enzima Kex de las construcciones pKLAC- α -MF-PSI(His)₆ y pKLAC- α -MF-PSIN86S(His)₆. Tal como se explicó en materiales y métodos, el proceso se realizó en un único paso de delección, y luego a partir de esa

construcción se insertó el sitio de reconocimiento para la enzima TEV antes de la secuencia codificante para las seis histidinas temprinales. De esta forma, se obtuvieron las construcciones pKLAC-PSI(His)₆ y pKLAC-PSIN86S(His)₆, en las cuales la secuencia del PSI y de su mutante del sitio de glicosilación quedaron inmediatamente después (*downstream*) de la secuencia del péptido señal del vector pKLAC1 sin estar presente la secuencia del dominio α -MF. La ausencia de la secuencia correspondiente al α -MF se corroboró por secuenciamiento automático del plásmido aislado de una colonia de transformantes TOP10F'.

Las construcciones pKLAC-PSI(His)₆ y pKLAC-PSIN86S(His)₆ generadas en el vector pKLAC1, a los que se les eliminó la secuencia correspondiente al dominio α -MF, se transformaron en la cepa GG799 de *K. lactis*. Al igual que en el caso de las construcciones que contenían el α -MF (Sección AII), los plásmidos fueron linealizados con la enzima *SacII* (Fig. 57) generando los casetes de expresión correspondientes y luego de ser purificados, cada uno de ellos se introdujo en una alícuota de células competentes por electroporación, como se indicó en materiales y métodos de este capítulo.

Linealización de pKLCA-PSI(His)₆ y pKLAC-PSIN86S(His)₆ con *SacII*

—Obtención de casetes de expresión—

149

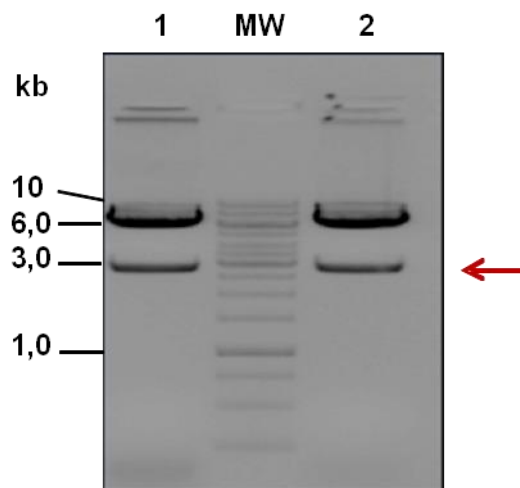


Fig. 57. Generación de los casetes de expresión lineales conteniendo la secuencia del péptido señal de pKLAC y del PSI(His)₆ o PSIN86S(His)₆. Las construcciones obtenidas subclonando las secuencias que codifican para el PSI(His)₆ o su mutante de sitio de glicosilación, PSIN86S(His)₆, en el vector pKLAC1 al que se le removió la secuencia del α -MF original, se linealizaron por digestión con la enzima *SacII* generando los casetes (flecha roja) que permiten la integración en la región del locus del promotor LAC4 en el genoma de *K. lactis*. *Calle 1*, pKLAC-PSI(His)₆; *Calle 2*, pKLAC-PSIN86S(His)₆; *Calle MW*, marcadores de masas moleculares de ADN 1 kb.

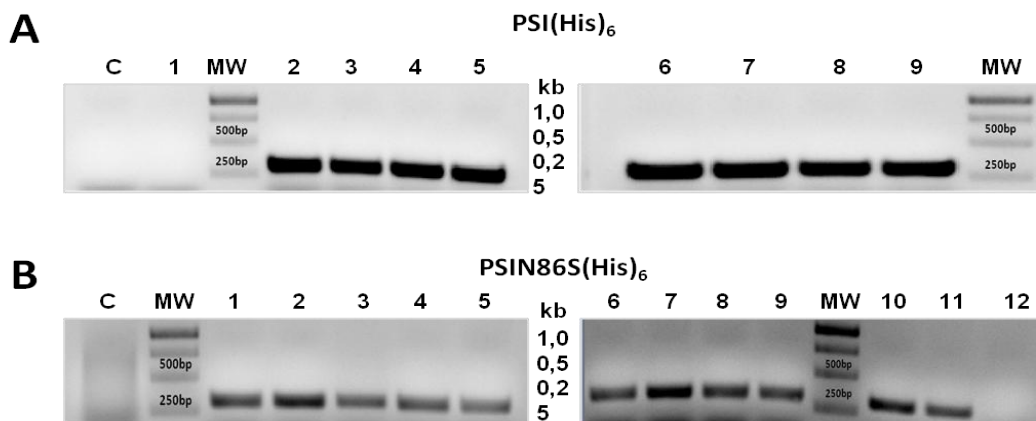
BII. Expresión de PSI en *K. lactis* a partir de las construcciones pKLAC-PSI(His)₆ y pKLAC-PSIN86S(His)₆



Selección de clones de K. lactis conteniendo las secuencias del PSI integradas en el genoma como casete de expresión y optimización del tiempo de expresión

La identificación de los clones crecidos a 30 °C en medio sólido YCB-Acetamida que integraron eficientemente cada casete de expresión (clones positivos), se realizó una vez más mediante *Colony PCR* empleando cebadores internos del PSI. De acuerdo a la amplificación de una banda de aproximadamente 350 pb observada en los geles de agarosa mostrados en la Fig. 58, se identificaron 8 clones positivos de 9 analizados para el casete conteniendo la secuencia PSI(His)₆ y 11 de 12 para el caso de la secuencia PSIN86S(His)₆.

Integración de los casetes de expresión conteniendo las secuencias PSI(His)₆ y PSIN86S(His)₆ en el genoma de *K. lactis*



150

Fig. 58. Identificación de células de *K. lactis* con casetes de expresión con las secuencias PSI(His)₆ y PSIN86S(His)₆ correctamente integrados en su genoma. La presencia de la secuencia PSI(His)₆ o PSIN86S(His)₆ en los genomas de las células transformadas de *K. lactis* se detectó por amplificación de una banda de aproximadamente 350 pb utilizando células enteras. Análisis electroforético en gel de agarosa (1%) del producto amplificado mediante *Colony PCR* de células transformadas con el casete conteniendo **(A)** la secuencia PSI(His)₆ o **(B)** la secuencia PSIN86S(His)₆. *Calle MW*, marcadores de masas moleculares de ADN 1 kb; *Calle C*, control negativo de integración con células GG799 sin transformar; *Calles 1 a 12*, productos de amplificación de los transformantes 1 a 12, respectivamente.

Los mejores candidatos a producir mayores cantidades de proteína secretada se evaluaron de forma complementaria con ensayos de expresión en pequeña escala. La proteína

recombinante secretada se identificó por inmunodetección con anticuerpos anti-His a diferentes tiempos: 48, 72, 96, 120 y 144 hs.

En la Fig. 59 A se muestra el blot realizado con las alícuotas tomadas a las 72 hs de expresión para los clones transformados con los casetes conteniendo el inserto PSI(His)₆ y en la Fig. 59 B, el correspondiente a las alícuotas luego de 24 hs de expresión de los clones transformados con el mutante de sitio de glicosilación PSIN86S(His)₆. El menor tiempo de expresión elegido para el mutante PSIN86S(His)₆ se debió a que a las 24 hs ya pudo detectarse proteína, cuya cantidad pareció ser disminuido transcurridas las 72 hs. Esta observación resulta consistente al rol en la estabilidad de las proteínas al que se vincula la glicosilación (da Costa et al., 2010).

Inmunodetección de la proteína secretada por los clones de *K. lactis* transformados con los casetes conteniendo las secuencias PSI(His)₆.

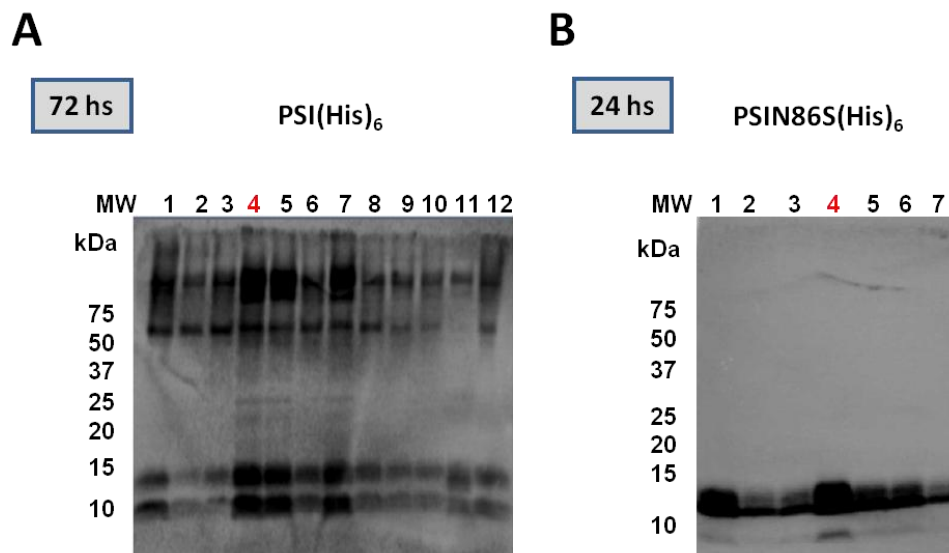


Fig. 59. Ensayos de expresión en pequeña escala del PSI(His)₆ en el sistema heterólogo *Kluyveromyces lactis*. Se analizó la expresión de proteína por los clones de transformantes sometidos a *Colony PCR*, en 10 ml de medio YP induciendo con 4% de galactosa y creciendo a 30°C. Las formas de PSI(His)₆ secretadas al medio del cultivo por los diferentes clones y se detectaron por Western blot con anticuerpos antiHis: *Calles MW*, marcadores de masas moleculares; *Calles C*, control negativo de integración con células GG799 sin transformar; **(A)** *Calles 1 a 12*, medios de cultivo tomados a las 72 hs de expresión de los clones 1 a 12 de *K. lactis* transformados con el casete que contiene al inserto PSI(His)₆; **(B)** *Calles 1 a 7*, medios de cultivo tomados a las 24 hs de expresión de los clones 1 a 7 transformados con el casete que contiene al inserto PSIN86S(His)₆.

En cuanto la región que codifica el α -MF es removida del plásmido pKLAC- α -MF-PSI(His)₆ los perfiles de las proteína expresadas (Fig. 59) se tornan menos heterogéneos con respecto a lo observado cuando la secuencia α -MF está presente en el casete de expresión (Fig. 37). Como se aprecia en los blots de la Fig. 59, una banda de intensidad significativa a la altura de la masa

molecular esperada (12571,22 Da). La banda ancha de elevado peso molecular que se correspondería con formas altamente glicosiladas del PSI(His)₆ observada en la Fig. 59 A, como era esperable está ausente en el mutante del sitio de glicosilación (Fig. 59 B).

En estas expresiones, como en el caso de las provenientes de los casetes con la secuencia del α -MF presente, se observa además una banda de aproximadamente 10 kDa que podría deberse a la expresión de formas truncadas de la proteína o bien, a degradación inespecífica por parte de las proteasas del huésped.

BIII. Análisis de la glicosilación de los productos expresados en *K. lactis* a partir de los casetes con las secuencias PSI(His)₆ y PSIN86S(His)₆.



Tratamiento de las formas de PSI obtenidas en *K.lactis* con endoglicosidasas

La proteína sin purificar expresada por el clon 4 transformado con el casete PSI(His)₆, se trató con la enzima glicosidasa PNGasa F a fin de verificar si el corte del enlace que une los residuos de oligosacáridos al residuo de asparagina del motivo de N-glicosilación (NET), resulta en la aparición de una banda electroforética acorde con la masa molecular teórica calculada con la herramienta Compute pI/Mw: 12571.22 Da.

En el caso de la expresión del casete PSI(His)₆ sin la secuencia del dominio α -MF, las bandas originales se desplazan hacia masas moleculares del orden esperado tanto cuando son tratadas con la enzima PNGasa F (Fig. 60) como cuando se tratan con la enzima Endo H (Fig. 61). El perfil obtenido como resultado de cada tratamiento mostró dos bandas de aproximadamente 12,5 kDa y 9,5 kDa.

Tratamiento del producto de expresión del casete PSI(His)₆ con PNGasa F

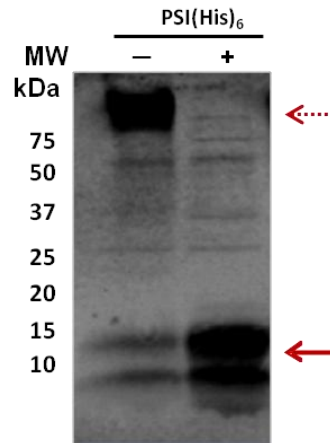


Fig. 60. Deglicosilación del producto de expresión del casete que contiene la secuencia PSI(His)₆ obtenido en *K. lactis* por acción de la enzima PNGasa F. Una alícuota del medio de cultivo con la forma PSI(His)₆ secretada, se sometió a digestión con la glicosidasa PNGase F durante 3 hs. Los productos digeridos se analizaron por inmunodetección con anticuerpos anti-His. *Calles MW*, marcadores de masas moleculares; *Calle (-)*, alícuota sin digerir; *Calle (+)*, alícuotas tratadas con PNGasa F.

La sensibilidad a la enzima Endo H, indica una vez más la ausencia de procesamiento de la glicoproteína PSI(His)₆ más allá del retículo endoplasmático.

Tratamiento del producto de expresión del casete conteniendo la secuencia

PSI(His)₆ con Endo H

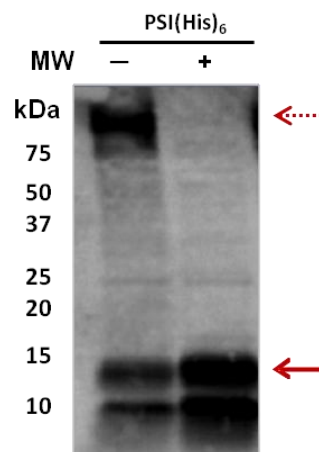


Fig. 61. Deglicosilación del producto de expresión del casete que contiene la secuencia PSI(His)₆ obtenido en *K. lactis* por acción de la enzima Endo H. Una alícuota del medio de cultivo de la expresión en pequeña escala del clon 4 transformado con la secuencia PSI(His)₆, se trató con glicosidasa Endo H durante 3 hs. Los productos de digestión se analizaron por inmunodetección con anticuerpos anti-His. *Calle MW*, marcadores de masas moleculares; *Calle (-)*, alícuota sin glicosidasa en la mezcla de reacción; *Calle (+)*, alícuota tratada con Endo H.

De lo observado en ambas imágenes (Fig. 60 y Fig. 61), se deduce que las bandas de alto peso molecular se corresponden con formas hiperglicosiladas y que la glicosilación es predominantemente N-ligada.



Purificación de los productos de expresión PSI(His)₆ y PSIN86S(His)₆

Debido a que la proteína expresada a partir del plásmido pKLAC-PSI(His)₆, no se pudo obtener en forma pura mediante el procedimiento aplicado, ya que la misma se co-purificó con la proteína sin glicosilar, se eligió trabajar con el mutante del sitio de glicosilación PSIN86S(His)₆ cuya purificación se muestra en las Fig. 62 y Fig. 63.

Cromatografía de exclusión molecular de la expresión del casete conteniendo la secuencia PSIN86S(His)₆

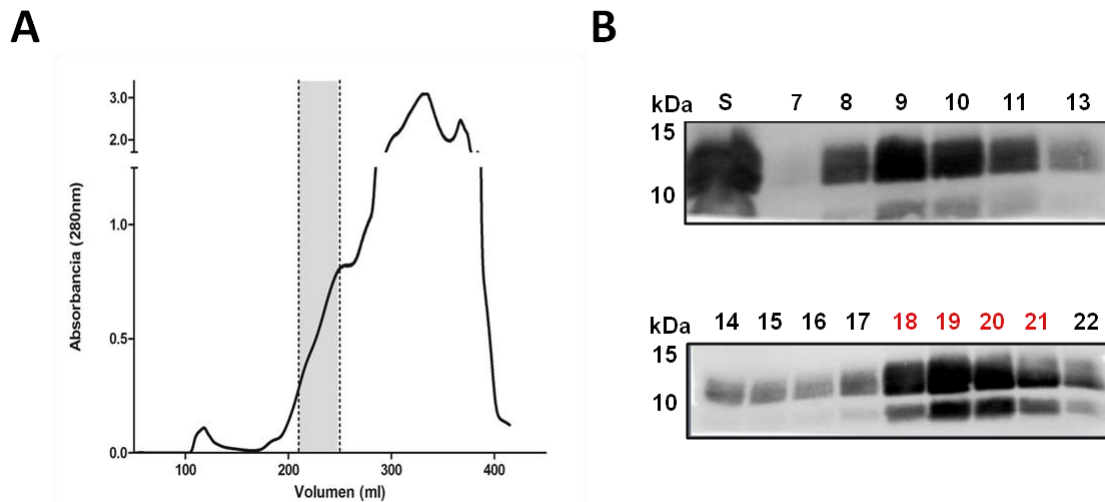


Fig. 62. Purificación parcial de PSIN86S(His)₆. El mutante del sitio de glicosilación (N86S) del dominio PSI se obtuvo a partir del clon 4 de *K. lactis* que contiene el casete PSIN86S(His)₆ sin la secuencia del α -MF integrada en su genoma. La proteína se expresó, concentró y purificó procediendo de la misma manera que en los casos de α -MF-PSI(His)₆ y α -MF-PSIN86S(His)₆. **(A)** Cromatografía de exclusión molecular de PSIN86S(His)₆ usando una columna HiLoad 26/60 Superdex 200. **(B)** Las fracciones eluidas se analizaron por inmunodetección con anticuerpos anti-His: *Calles MW*, marcadores de masas moleculares; *Calle S*, muestra aplicada; *Calles 7 a 22*, fracciones eluidas de la columna.

De la cromatografía de exclusión molecular (Fig. 62 A) se seleccionaron las fracciones 18 a 21, las cuales se juntaron y se aplicaron a una columna HisTrap (Fig. 63). Al comparar las

poblaciones de proteínas observadas en la Fig. 44, se puede concluir que en este caso desapareció el efecto de arrastre y las bandas de masa molecular de 13 kDa y 11,4 kDa separadas, resultaron mucho más nítidas. Las fracciones eluidas con 100 mM y 500 mM de imidazol se juntaron y conservaron para posteriores análisis.

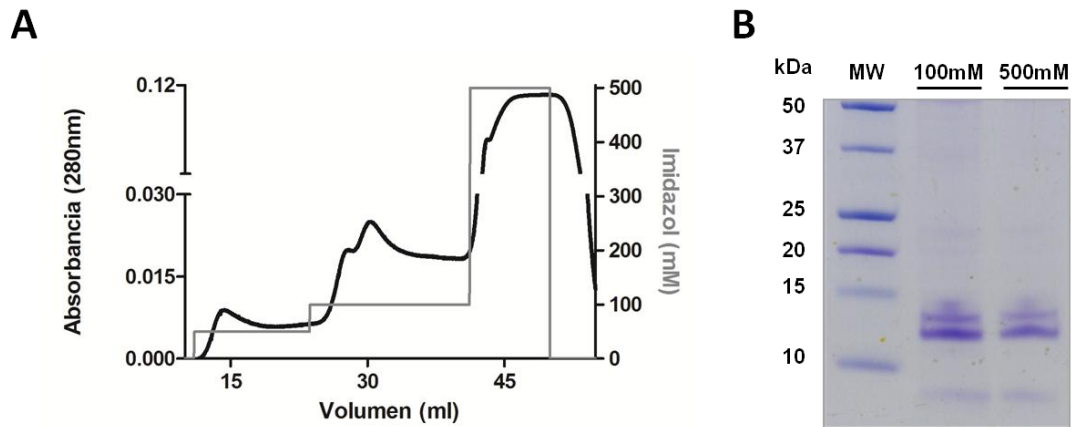


Fig. 63. Purificación por IMAC del producto de expresión del casete conteniendo la secuencia PSIN86S(His)₆ parcialmente purificado. (A) A partir de las fracciones eluidas de la columna Superdex 200 (S-200) mostradas en la Fig. 54, se llevó a cabo la cromatografía IMAC en columna HisTrap. (B) Análisis mediante SDS-PAGE de las fracciones colectadas durante la cromatografía de afinidad. El gel fue teñido con Coomassie Brilliant Blue: *Calle MW*, marcadores de masa molecular; *Calle 100 mM* y *500 mM*, eluatos de columna HisTrap recolectados a concentración 100 mM y 500 mM de imidazol en el buffer de elución, respectivamente.

Del análisis de las expresiones de las formas α -MF-PSI(His)₆, α -MF-PSIN86S(His)₆, PSI(His)₆ y PSIN86S(His)₆, se deduce que la fusión de un dominio secretorio al PSI en el sistema *K. lactis* no fue condición imprescindible para la acumulación de la proteína recombinante en el medio extracelular. Por otra parte, la glicosilación tampoco resultó ser un factor determinante para la secreción del PSI en este sistema de expresión.



Corte de la proteína de fusión PSIN86S(His)₆ con TEV

A fin de separar la cola de seis histidinas incluida en el extremo C-terminal de la proteína PSIN86S(His)₆ para su purificación y detección, se hizo uso del sitio de reconocimiento de la enzima TEV (ENLYFQ) introducido entre la secuencia del PSIN86S y la cola (His)₆. Las condiciones para llevar a cabo el corte de la proteína de fusión con TEV se optimizaron según se describió en *Materiales y Métodos—PSI*, en función de la relación enzima: sustrato, el

ambiente reductor, la temperatura y el tiempo de reacción. En la Fig. 64 se muestran solamente las digestiones con TEV realizadas con relación enzima:sustrato de 1:10; las fracciones resultantes en cada condición probada se analizaron por SDS-PAGE y por Western blot en forma simultánea.

Como se observa en la Fig. 63, las bandas de pesos moleculares intermedios y más bajos se hacen mucho menos intensas en el western blot revelado con anticuerpos anti-His, mediante tratamiento con la enzima TEV a temperatura ambiente en presencia de DTT a partir de las 5 hs y con glutatona luego de toda la noche. Cuando el tratamiento se llevó a cabo a 4 °C, la disminución en las intensidades resultó notoria sólo con incubaciones ON.

Teniendo en cuenta los resultados obtenidos en la optimización, se escogieron las siguientes condiciones de digestión para el tratamiento del volumen total de la muestra:

- Relación PSI:TEV = 1:10
- Condiciones reductoras: Glutatona Reducida (3 mM) / Glutatona Oxidada (0,3 mM)
- Temperatura: 4 °C
- Tiempo: ON

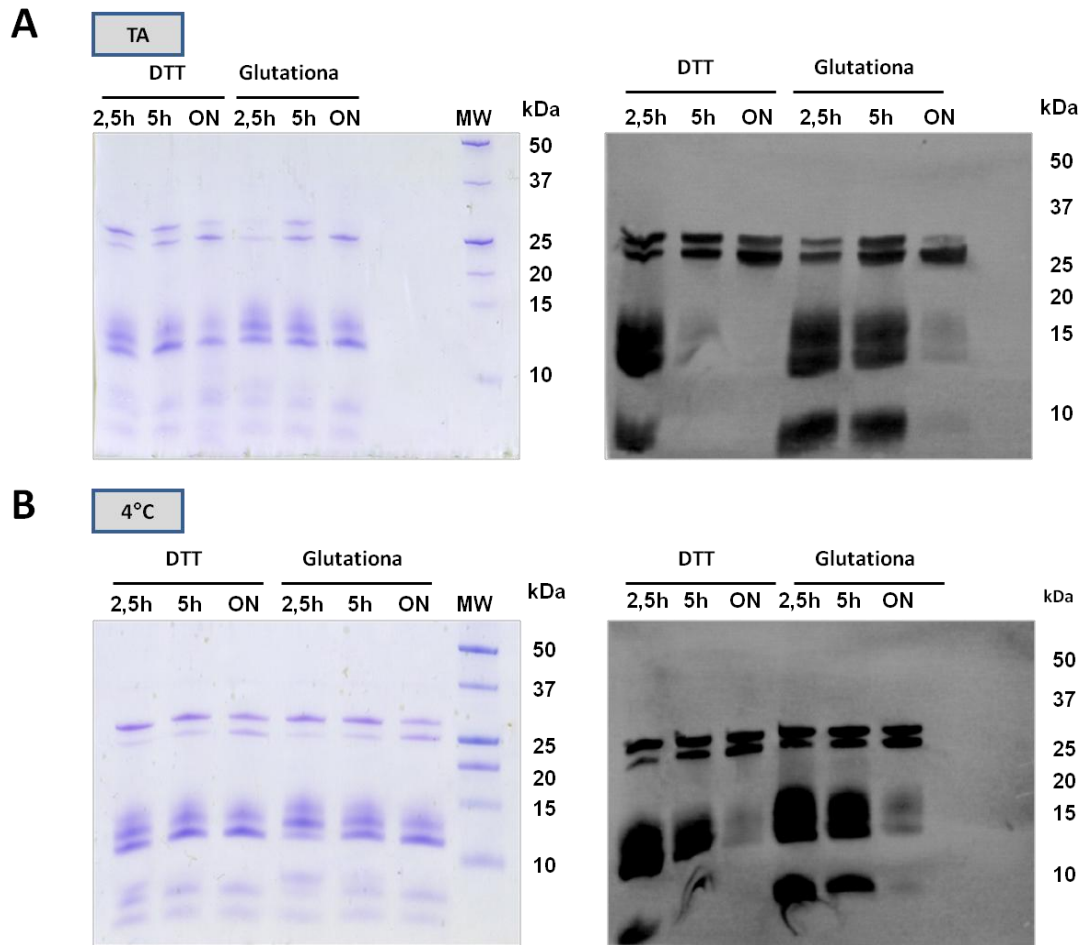
Optimización del corte de la cola de histidinas terminal en PSIN86S(His)₆ con la enzima TEV

Fig. 64. Optimización de las condiciones de reacción con la enzima TEV para separar las seis histidinas de la proteína PSIN86S(His)₆. Fracciones de PSIN86S(His)₆ purificada se sometieron a digestión con la enzima TEV (relación enzima:sustrato de 1:10), en presencia de DTT como agente reductor o del par Glutaciona Reducida/Glutaciona Oxidada, durante 2,5 hs; 5 hs y ON. La eficiencia de corte se analizó mediante SDS-PAGE y Western blot, revelando con anti-His conjugado a fosfatasa alcalina y el reactivo ECF. **(A)** Análisis del producto de la digestión realizada a temperatura ambiente. **(B)** Análisis de la digestión llevada a cabo a 4 °C. *Calles MW*, marcadores de masa molecular; *Calles 2,5 hs, 5 hs, ON*, indican el tiempo de digestión al que fueron sometidas las muestras aplicadas en el gel. La acción de la enzima TEV y por ende, la eliminación de la cola (His)₆ se visualiza en los blots como una disminución de intensidad de la señal. Se indica también, la presencia de DTT o Glutaciona en la mezcla de reacción en cada digestión.

Si bien a temperatura ambiente la reacción prácticamente se completó en tiempos menores con DTT, se prefirió trabajar a 4 °C durante toda la noche a fin de evitar agregación proteica. La eficiencia de corte luego de la digestión de PSIN86S(His)₆ con la enzima TEV y la integridad de la proteína separada, PSIN86S, se analizaron mediante un gel de poliacrilamida en condiciones desnaturalizantes y digestión trípica y huella peptídica (Fig. 65, Fig. 66 y Fig. 67).

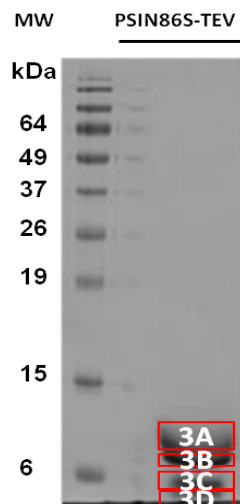


Análisis por técnicas de proteómica del producto purificado de la expresión del casete PSIN86S(His)₆

En la Fig. 65 A se muestran las bandas electroforéticas analizadas de la proteína expresada a partir del casete PSIN86S(His)₆, purificado y digerido con la enzima TEV. Se indica también la secuencia aminoacídica de la proteína que se tuvo en cuenta para realizar los análisis (Fig. 65 B).

SDS-PAGE del producto de expresión del casete PSIN86S(His)₆ en *K. lactis* purificado y digerido con TEV

A



B

PSIN86S-TEV

APLEVMSQOCKTLVSQYGKSI IEMLLSEAQPKICSQMKLCTFDG	45
ARDVSSI IESVVDKNNGKSSGGANDEMCTFCEMAVVWMQNQIKRS	90
ETEDNI INYVNELCDRLPSENLYFQ*	115

● Sitio Xho

● PSI

● Sitio de corte TEV

Ⓢ Mutación N86S

Fig. 65. Muestra obtenida por expresión del casete PSIN86S(His)₆ en *K. lactis*, purificada y tratada con TEV para su análisis por técnicas de proteómica. (A) SDS-PAGE de la muestra, sobre los geles se indican las bandas que han sido sometidas a digestión tróptica y posterior espectroscopía de masas MALDI-TOF. (B) Secuencia considerada en el análisis de la proteína, en la misma se resaltan la mutación del sitio de glicosilación N86S, el dominio PSI, y los residuos codificados por el cebador utilizado en el extremo 5' (Sitio Xho) durante la construcción del vector de expresión y los aminoácidos incorporados en el extremo C-terminal para eliminar la cola de polihistidinas mediante la enzima TEV.

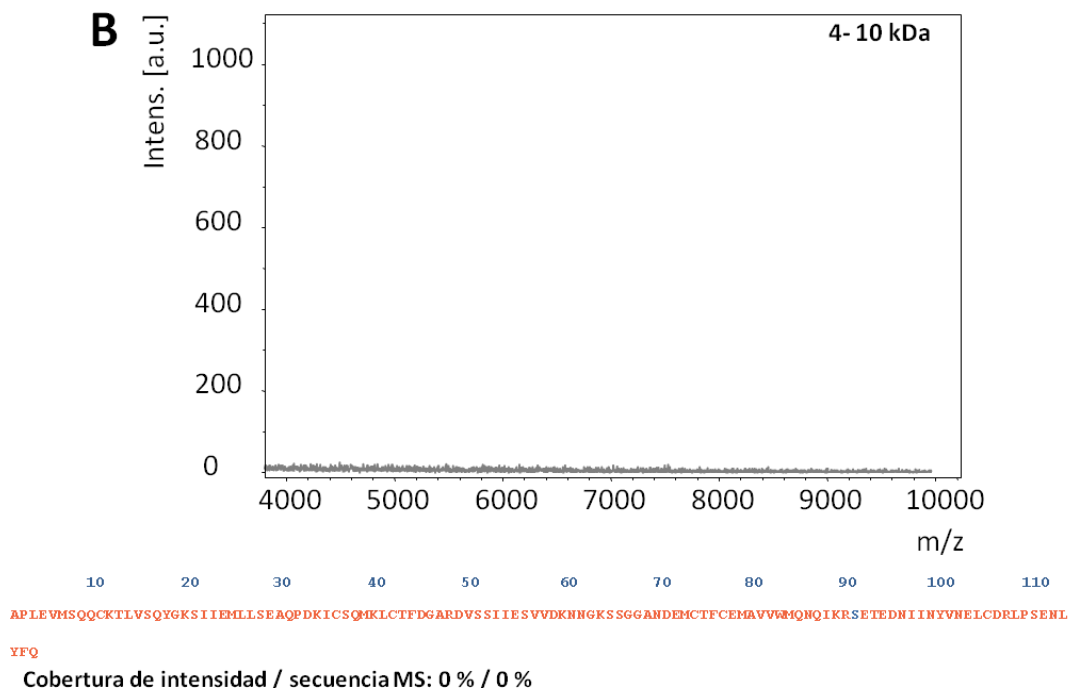
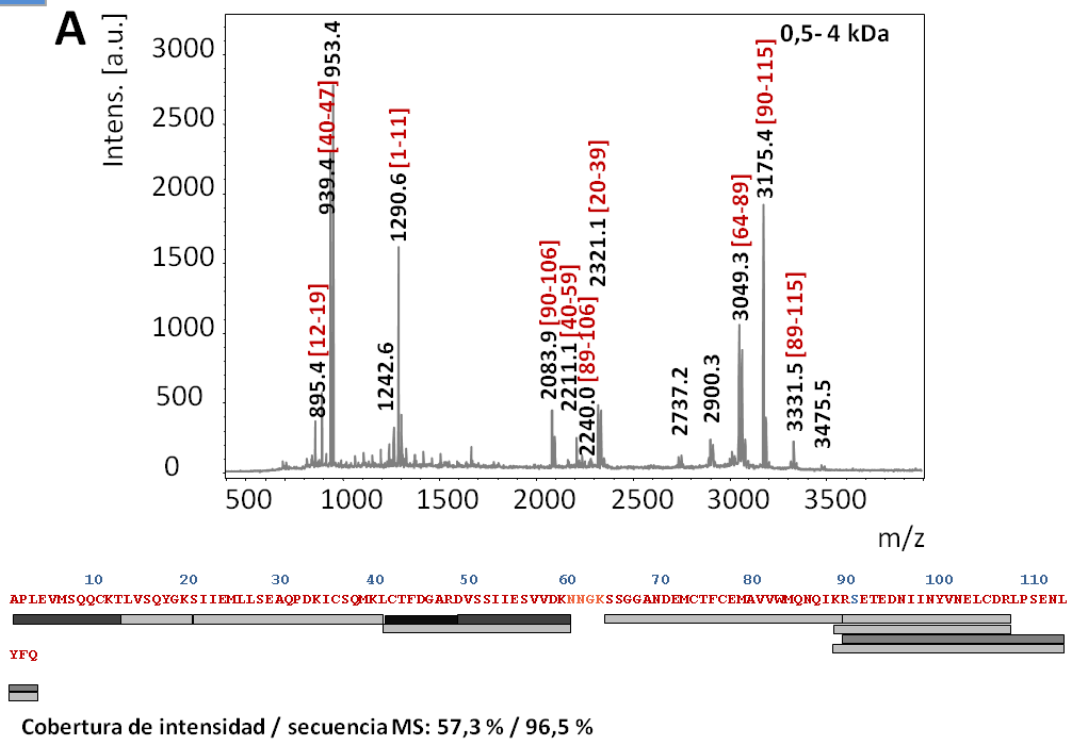
Los espectros obtenidos con las bandas 3A y 3B se presentan en las Fig. 66 y 67 junto a los fragmentos de las secuencias a los que mediante los programas Flex Analysis y Sequence Editor de Bruker Daltonics, se les ha asignado un pico. Al igual que en los casos anteriores, los espectros se adquirieron en dos ventanas de relación masa/carga, una estándar de 400 a 4000 m/z y una segunda ventana de 4000 a 10000 m/z.

Cuando se analizaron las bandas provenientes de la proteína del sitio de glicosilación expresada sin el dominio α -MF, PSIN86S, los picos asignados en los espectros de las bandas 3A y 3B (Fig. 66 y Fig. 67) cubrieron prácticamente la totalidad de la secuencia de la proteína (cobertura de secuencia de 96,5%), incluyendo la mutación del sitio de glicosilación N86S y abarcando tanto el extremo N-terminal como el remanente del sitio de reconocimiento para la enzima TEV luego del corte (ENLYFQ) en el extremo C-terminal. Cabe destacar que en la ventana de mayores valores de m/z tal como se esperaba, no apareció el pico de 6277,87 Da ni el perfil de glicosilación a partir de 9 kDa, ratificando de esta manera la asignación de dichos picos al dominio α -MF sin glicosilar y glicosilado en las proteínas anteriores. Las diferencias en las movilidades electroforéticas de las bandas 3A y 3B (Fig. 65 A) en base a los resultados, no pudieron adjudicarse a variaciones en la secuencia primaria (Fig. 66 y Fig. 67).

En conjunto estos datos permitieron determinar de manera concluyente la presencia de la proteína recombinante en el medio extracelular aun cuando se quitó el dominio α -MF, responsable de garantizar su secreción en el sistema heterólogo *K. lactis*.

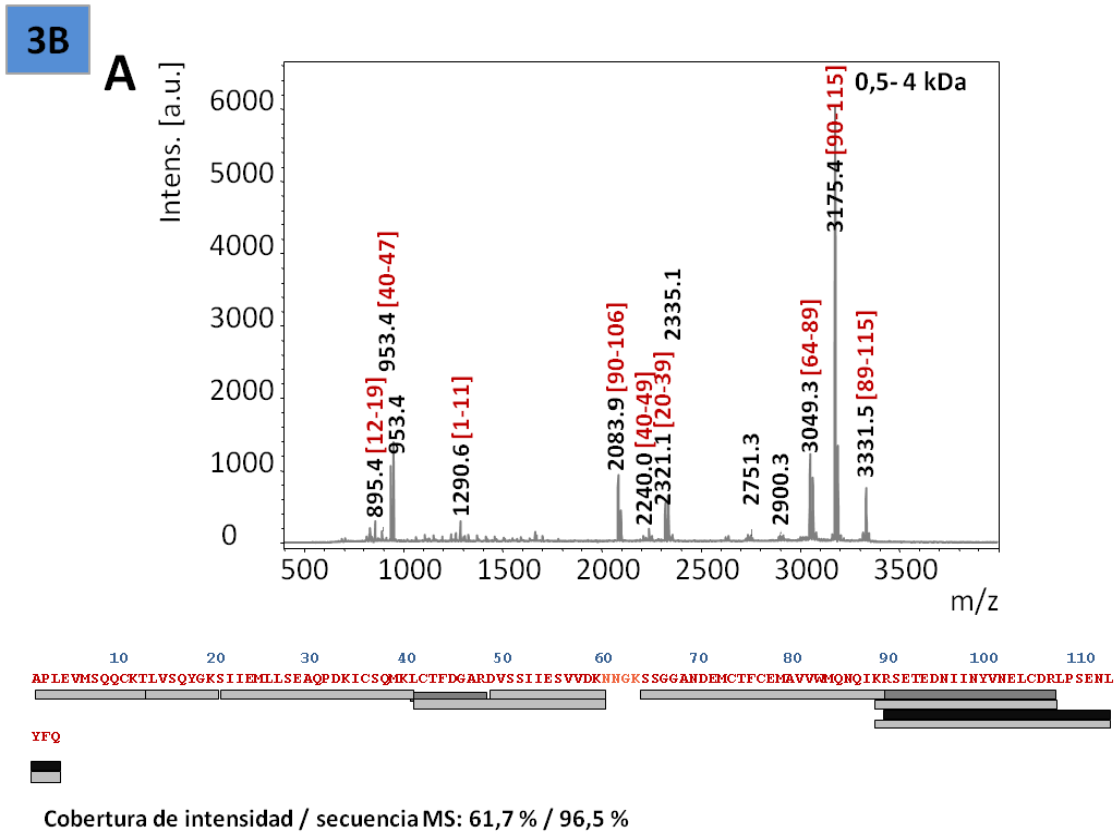
PMF de la banda proteica 3A correspondiente a la expresión del casete PSIN86S(His)₆

3A



160

Fig. 66. Análisis por técnicas de proteómica de la banda proteica 3A de la muestra obtenida mediante la expresión del casete PSIN86S(His)₆ en *K. lactis* Espectros adquiridos en dos ventanas de relación masa/carga, (A) de 400 a 4000 m/z y (B) de 4000 a 10000 m/z, sobre cada pico se indica la masa molecular correspondiente. Debajo del espectro se muestra la secuencia de la proteína considerada y por medio de barras se señalan los péptidos resultantes de la digestión trípica teórica cuya masa molecular se correlaciona con el valor de un pico del espectro. La escala de grises de las barras representa la intensidad de los picos.

PMF de la banda proteica 3B correspondiente a la expresión del casete PSIN86S(His)₆**B**

En la ventana de 4-10 kDa no se observaron picos.

Fig. 67. Análisis por técnicas de proteómica de la banda proteica 3B de la muestra obtenida mediante la expresión del casete PSIN86S(His)₆ en *K. lactis*. Espectros adquiridos en dos ventanas de relación masa/carga, (A) de 400 a 4000 m/z y (B) de 4000 a 10000 m/z, sobre cada pico se indica la masa molecular correspondiente. Debajo del espectro se muestra la secuencia de la proteína considerada y por medio de barras se señalan los péptidos resultantes de la digestión triptica teórica cuya masa molecular se correlaciona con el valor de un pico del espectro. La escala de grises de las barras representa la intensidad de los picos.

Los espectros de las otras dos bandas, 3C y 3D, no brindaron mayor información ya que en el espectro correspondiente a la banda 3C sólo se vieron picos de polímeros y no pudo ser analizado. Mientras que para la banda 3D se asignaron dos picos del espectro, de los cuales uno de ellos abarca el extremo C-terminal de la proteína y el otro pico es coincidente con un valor característico de la autodigestión de la tripsina (2211,147 Da). Por tanto, si bien podría tratarse de una forma de la proteína degradada en su N-terminal, la información no es concluyente.

En resumen, los resultados obtenidos a partir del análisis de los productos de expresión por técnicas de proteómica permitieron arribar a las siguientes conclusiones:

- Las formas de PSI fusionadas al dominio α -MF no son procesadas de manera eficiente en la vía de secreción de *K. lactis*.
- La presencia del dominio α -MF en la proteína recombinante genera un perfil de glicosilación heterogénea de alto peso molecular.
- El PSI consigue avanzar por la vía de secreción del sistema *K. lactis* una vez que es dirigido al retículo endoplasmático por el péptido señal, sin necesidad de estar fusionado a la secuencia secretoria α -MF.
- La glicosilación del único sitio putativo del PSI no influye de modo decisivo en su secreción de la célula de *K. lactis*.

Ensayos antimicrobianos realizados con diferentes formas del PSI expresadas en *K. lactis*

Debido a su estructura, el PSI de cirsina, así como los de otras especies vegetales, pertenece a la familia de las proteínas similares a saposina, la que incluye a NK-lisina, granulosina, proteína surfactante B, esfingomielinasa ácida, amoebapores, entre otras. Dada la capacidad de interactuar con membranas celulares se ha descrito que poseen acción citolítica, antimicrobiana, y también contra células tumorales al causar la muerte celular a través de la formación de poros o desestabilización de la membrana de la célula huésped.



Actividad antibacteriana

Dado los reportes existentes sobre la actividad antimicrobiana de las proteínas de la familia SAPLIP, se propuso evaluar la capacidad del dominio PSI de *C. vulgare* de afectar el crecimiento bacteriano. Estos ensayos, se llevaron a cabo en una primera etapa, mediante las pruebas de dilución en agar y/o de macrodilución a pH 5,0, empleando el mutante del sitio de glicosilación del dominio PSI expresado en *K. lactis* sin el α -MF, PSIN86S con o sin la cola de polihistidinas terminal. Entre los microorganismos Gram (+) evaluados por una de estas dos metodologías se escogieron *Staphylococcus aureus* y *Enterococcus faecalis*, mientras que las cepas *E. coli*,

Salmonella typhi, *Pseudomona aeruginosa* y *Protius mirabilis* se tomaron como representantes Gram (-). Como se informa en la Tabla 20, no se observó actividad antibacteriana de las formas del PSI en ninguno de los ensayos realizados.

Luego, en una segunda etapa a fin de confirmar lo observado con las dos pruebas anteriores, se empleó el método de Kirby-Bauer frente a una cepa Gram (+) y a una Gram (-). Además del pH 5,0 previamente evaluado, el ensayo se repitió a pH 7,0. Una vez más, los efectos del PSI sobre los microorganismos ensayaron resultaron negativos (Tabla 21).

Test de dilución en agar: PSIN86S	
Cepa bacteriana	pH 5,0
<i>Escherichia coli</i>	No se observó actividad
<i>Salmonella typhi</i>	No se observó actividad
<i>Staphylococcus aureus</i>	No se observó actividad
<i>Proteus mirabilis</i>	No se observó actividad
<i>Pseudomonas aeruginosa</i>	No se observó actividad
<i>Enterococcus faecalis</i>	No se observó actividad

Test de macrodilución: PSIN86S	
Cepa bacteriana	pH 5,0
<i>Escherichia coli</i>	No se observó actividad
<i>Salmonella typhi</i>	No se observó actividad

TABLA 20. Actividad antibacteriana del dominio PSIN86S expresado en *K. lactis*

Test de Kirby-Bauer: PSIN86S		
Cepa bacteriana	pH 5,0	pH 7,0
<i>Escherichia coli</i>	No se observó actividad	No se observó actividad
<i>Salmonella typhi</i>	No se observó actividad	No se observó actividad
<i>Staphylococcus aureus</i>	No se observó actividad	No se observó actividad

TABLA 21. Actividad antibacteriana del dominio PSIN86S expresado en *K. lactis*



Actividad antifúngica

Se analizó la actividad antifúngica del dominio PSI de procirsina mediante el test de co-crecimiento en placa frente a las siguientes especies fúngicas: *Lewia infectoria*, *Alternaria alternata*, *Dreschsiera biseptata*, *Botryosphareia obtuse*, *Pleospora herbarum*. En la Fig. 68, se grafican todas las condiciones en que se hizo el ensayo, contrastando los resultados con los obtenidos con controles negativos llevados a cabo con la cepa GG799 de *K. lactis* sin transformar y con un cepa de GG799 expresando una proteína heteróloga cualquiera (GFP) la que a su vez actuó simulando el estado metabólico de la levadura GG799 productora de PSIN86S. Se incluyeron también dos bacterias, Fito_B108 y Fito_S127B, con acción antifúngica como controles positivos.

De acuerdo a lo que muestran las Fig. 68 A y B, el PSIN86S afectó en forma apreciable (entre 30% y 40% de inhibición del crecimiento micelar), a las especies *L. infectoria* y *A. alternata*, respectivamente, en las tres condiciones probadas. No obstante, también se vio una inhibición semejante producida por las cepas utilizadas como controles negativos, sugiriendo la liberación de un metabolito o compuesto con acción antifúngica sobre estas dos especies, por parte de la cepa de *K. lactis* utilizada en la expresión del PSIN86S de procirsina.

En el caso de *D. biseptata*, en la Fig. 68 C se observa que la cepa productora del PSI heterólogo no consiguió inhibir el crecimiento micelar cuando el ensayo se realizó en medio PDA, a los dos valores de pH probados. Si bien en el medio YPGal se observó una inhibición del crecimiento de esta especie fúngica mayor al 60%, al igual que lo acontecido con *L. infectoria* y *A. alternata*, un grado de inhibición similar se dio con los controles negativos.

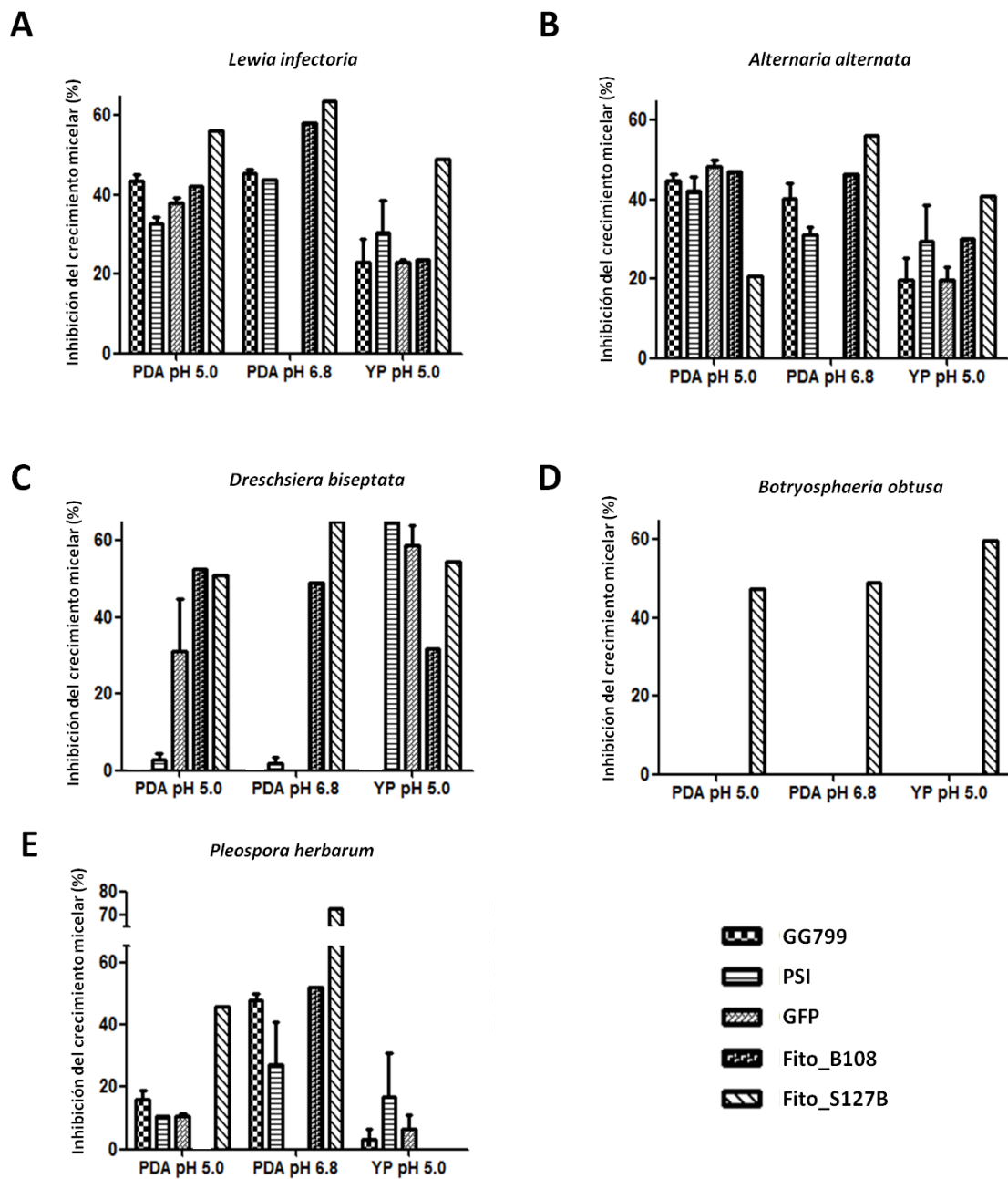
Co-crecimiento *K. lactis* secretora de PSIN86S(His)₆ y distintas especies fúngicas

Fig. 68. Efecto del co-crecimiento de levaduras expresando PSIN86S y especies fúngicas fitopatógenas. Placas de medio PDA a pH 5,5 y 6,8, y de medio YP a pH 5,5 fueron incubadas con discos conteniendo células de *K. lactis* que expresan PSIN86S y un fitopatógeno. Se determinó el porcentaje de Inhibición de Crecimiento Micelar cuando los hongos crecieron en presencia de las levaduras transformadas, en referencia al crecimiento en condiciones libres: (A) *L. infectoria*; (B) *A. alternata*; (C) *D. biseptata*; (D) *B. obtusa*; (E) *P. herbarum*.

Frente a la especie *B. obtuse* el dominio PSIN86S no ejerció efectos inhibitorios (Fig. 68 D), mientras que el crecimiento micelar de *P. herbarum* resultó pobremente inhibido por parte del PSIN86S (valores cercanos a 10%) o incluso, se vio menos afectada por éste que por el co-crecimiento con las células empleadas como control negativo (Fig. 68 E).

Si bien las propias células GG799 mostraron una influencia sobre el crecimiento micelar representando por tanto una interferencia en el ensayo de co-crecimiento, se consideraron las tendencias observadas mediante este ensayo como indicios de presencia o ausencia de actividad antifúngica por parte del dominio PSI secretado por las células transformadas en cada caso.

En base a estas observaciones, se empleó un segundo ensayo más específico en el que se evaluó al dominio PSIN86S purificado—sin las células productoras—, sólo frente a aquellas especies fúngicas para las cuales el ensayo de co-crecimiento mostró indicios de ser afectadas por la proteína. Esta segunda evaluación del PSIN86S recombinante como agente antifúngico se realizó mediante la prueba de difusión en disco o test de Kirby-Bauer, descrito en el capítulo de *Materiales y Métodos—PSI*, frente a las especies *A. alternata*, *D. biseptata* y *L. infectoria*, y utilizando el antifúngico Anfotericina B como control positivo.

Actividad antifúngica de PSIN86S purificado — Prueba de Kirby-Bauer

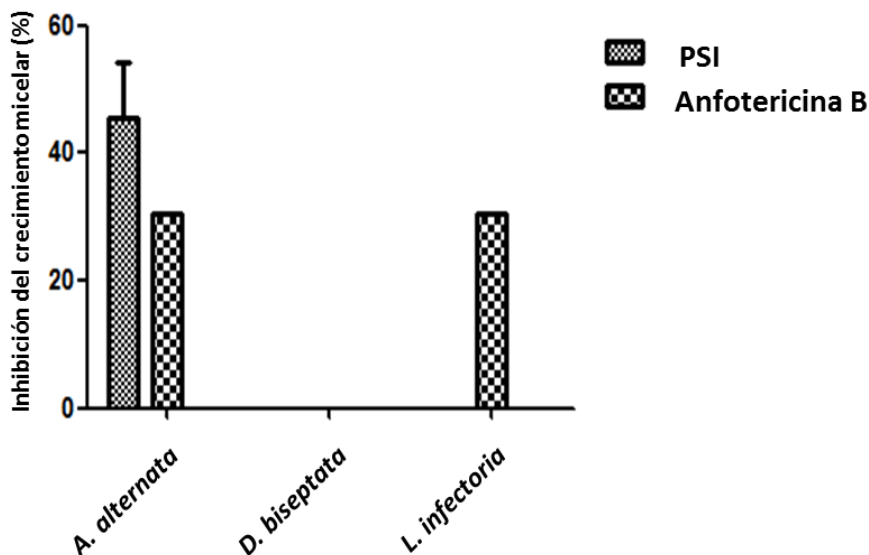


Fig. 69. Efecto antifúngico del mutante de sitio de glicosilación PSIN86S empleando el Test de Kirby-Bauer. En placas de medio PDA (pH 5,0) discos conteniendo las especies *Dreschlera biseptata*, *Lewia Infectoria*, *Alternaria alternata*, se incubaron durante 10 días en presencia de un segundo disco con la proteína PSIN86S(His)₆. Se incluyeron controles positivos con Anfotericina B.

Los resultados obtenidos se muestran en la Fig. 69 donde se ve una vez más una acción pronunciada del PSIN86S como inhibidor del crecimiento fúngico sobre la especie *A. alternata*, aún mayor a la observada con el control positivo. Mientras que el crecimiento de la especie *L. infectoria* resultó ser sólo inhibido por la Anfotericina B. Como en el ensayo anterior, *D. bisepitata* no se vio afectada por el PSIN86S, ni por el control positivo.

Discusión—PSI

La secuencia de aproximadamente 100 aminoácidos conocida como inserto específico de planta (PSI, del inglés *plant specific insert*), constituye un dominio independiente que se presenta exclusivamente dentro de los precursores de las proteasas aspárticas típicas de plantas. Su estructura tridimensional ubica a este dominio entre las proteínas similares a saposinas (SAPLIPs), a las que comúnmente se les asignan propiedades membranolíticas vinculadas a la estructura compartida por los miembros de la familia SAPLIP. La segunda parte del presente trabajo de tesis, se focalizó en el dominio PSI de la nueva PA típica de *Cirsium vulgare* reportada en el capítulo previo—cirsina. Las secuencias correspondientes al PSI y al mutante de su único sitio de glicosilación N86S-PSI, se subclonaron en forma aislada en el vector pKLAC1, se expresaron en el sistema eucariota *Kluyveromyces lactis* y se purificaron. Posteriormente se evaluó la acción del mutante N86S-PSI recombinante frente a diversos microorganismos, bacterianos y fúngicos.

Si bien la expresión en *E.coli* fue planteada como primera tentativa en la obtención del PSI recombinante para su aplicación en estudios biológicos debido a la simplicidad del sistema y en base a datos bibliográficos preexistentes donde esta estrategia se describe como exitosa (Egas et al., 2000; Muñoz et al., 2010), los bajos rendimientos conseguidos al final del proceso completo (expresión, corte de la cola de GST y purificación del PSI aislado) cuando se produjo la proteína de fusión (GST)PSI en la cepa BL21 Star, datos no reportados en esta tesis, hizo que se planteara el uso de un sistema de levaduras.

El primer abordaje para la expresión en *K. lactis* fue la obtención del PSI y su mutante como proteínas de fusión con el dominio secretorio α -MF—región pro de la feromona α -factor de *Saccharomyces cerevisiae*—mediante clonado entre los sitios *XhoI* y *BglII* del vector pKLAC1, y el agregado de la secuencia codificante para seis histidinas en el extremo C-terminal. De acuerdo a los datos informados en este capítulo, esta estrategia dio lugar a la expresión de poblaciones proteicas altamente heterogéneas donde las fracciones mayoritarias se correspondieron con formas secretadas que conservaron el dominio α -MF, sin haber sufrido su procesamiento normal. Resultó notable a su vez, el inusual perfil de glicosilación que se extiende desde la zona de masas moleculares de aproximadamente 20 kDa hasta valores

superiores a los 75 kDa. El tratamiento con endoglicosidasas reveló glicosilación principalmente N-ligada y sensible a Endo H, incluso en el mutante PSIN86S(His)₆, y mostró un perfil de bandas con masas mayores a los pesos teóricos del PSI(His)₆ y PSIN86S(His)₆ no glicosilados. La secreción de proteínas aún fusionadas al dominio α -MF en el extremo N-terminal de estas muestras se confirmó mediante PMF MALDI-TOF MS. Los resultados obtenidos mediante esta técnica de proteómica, permitieron no sólo identificar al segmento α -MF, al PSI y a la mutación N86S en la muestra correspondiente, sino que además permitieron asociar un perfil de glicosilación heterogénea con la presencia del dominio α -MF.

En conjunto, la información aquí presentada sugiere que el procesamiento de las proteínas expresadas a partir de casetes que contienen la secuencia α -MF no se lleva a cabo de la manera esperada en el *trans*-Golgi de la levadura *K. lactis*. A su vez, dan un indicio que el particular perfil se debe a la glicosilación de los tres sitios (Asn-X-Thr) existentes en la secuencia pro del precursor del α -factor (Caplan et al., 1991), y que las cadenas externas de carbohidratos no han sido modificadas más allá del retículo endoplasmático.

Debido a la porción extra de 64 aminoácidos en las proteínas secretadas por las células de *K. lactis* transformadas y a la complejidad que esto ocasionó, se decidió modificar la estrategia de clonado en el vector pKLAC1. Cabe destacar que no se encontraron en la literatura reportes sobre la existencia de un procesamiento ineficaz del dominio α -MF en el sistema de expresión *K. lactis*, y que el manual sólo plantea una segunda estrategia de clonado cuando la secuencia del inserto posee un sitio de reconocimiento para la enzima *XhoI*. En esta estrategia se propone como sitio alternativo de clonado el correspondiente a la enzima *HindIII*, lo cual genera una proteína con un extremo N-terminal no nativo. A pesar de no estar explícitamente documentado en el manual de *K. lactis*, de acuerdo a lo reportado para otros sistemas de expresión como *Pichia pastoris* y *S. cerevisiae*, los residuos que incorpora la secuencia del sitio *HindIII* podrían aumentar la eficiencia de corte del dominio α -MF por acción de la enzima Kex-2 en el *trans*-Golgi (Brake et al., 1984). Sin embargo, la incorporación de estos residuos en el extremo N-terminal mediante la segunda estrategia propuesta en el manual, aportan dos cargas netas negativas, que podrían influenciar en la actividad biológica del PSI a estudiar.

Por otro lado, en la literatura se encuentra reportada para *S. cerevisiae*, la eficiencia en la secreción y procesamiento de la fusión del pre-segmento del precursor de la feromona α -factor a proteínas heterólogas como aminoglucósido fosfotransferasa (APH) y factor estimulante de colonias de granulocitos-macrófagos (hGM-CSF), del mismo modo que las correspondientes fusiones conteniendo además la porción pro de dicho precursor. Más aún, se

reportó también que mientras la fusión de interleuquina-1 β a las porciones pre-pro del precursor de α -factor de *S. cerevisiae* no fue clivada, las fusiones con la región pre únicamente fueron correctamente procesadas durante la secreción. La alta eficiencia de secreción de proteínas fusionadas al pre-segmento han sugerido que la región pro del precursor de α -factor no es requerido para la eficiente secreción de proteínas de fusión (Ernst, 1988).

A partir estas consideraciones, la segunda estrategia de obtención del PSI y su mutante se desarrolló en base a la fusión directa con el péptido señal del sistema de expresión (región pre del precursor de α -factor de *S. cerevisiae*), sin el dominio α -MF (correspondiente sólo al segmento pro del precursor de α -factor). Asimismo, en esta estrategia se analizó la influencia de la cola terminal de polihistidinas en la estructura tridimensional del dominio PSI mediante herramientas bioinformáticas. Los resultados arrojados por este análisis predicen un cambio hacia una conformación de C más cerrada provocada por la ubicación de la cola de His terminal en el interior de la estructura del PSI. Según datos cristalográficos sobre el PSI de una PA de *Solanum tuberosum* (StAP-PSI), este dominio aislado del entorno de la proteasa presenta una conformación de *boomerang* abierto más que cerrado (Bryksa et al., 2011), en contraposición a la conformación de V cerrada observada para el PSI de profitepsina (*Hordeum vulgare*) cuando la estructura resuelta correspondió al precursor entero.

Por esta razón durante la etapa de clonado de la segunda estrategia se incluyó un sitio de reconocimiento para la enzima TEV que permitió remover las seis histidinas de la proteína recombinante.

Contrariamente a lo observado para las formas wt y mutante obtenidas por expresión de los casetes que contienen la secuencia del α -MF, donde las masas de las bandas generadas por la digestión con PNGasa F no coincidieron con los pesos teóricos del PSI(His)₆ y PSIN86S(His)₆ no glicosilados, el producto de la expresión del casete PSI(His)₆ sin la secuencia del dominio α -MF, mostró un desplazamiento de las bandas originales hacia masas moleculares del orden esperado tanto cuando se trataron con la enzima PNGasa F como cuando se trataron con la enzima Endo H—dos bandas de 13 kDa y 11,5 kDa. Más aún, los datos arrojados por PMF MALDI-TOF MS de las bandas correspondientes a la expresión del mutante N86S-PSI sin fusión al α -MF, permitieron determinar de manera concluyente la presencia de la proteína recombinante en el medio extracelular. En conjunto estos resultados sugieren que en efecto, tanto el dominio PSI(His)₆ como su mutante PSIN86S(His)₆ lograron ser eficientemente secretados y procesados en el sistema *K. lactis* aún en ausencia del segmento pro del α -factor de *S. cerevisiae*. Esta observación está en concordancia con lo reportado para la expresión de

la proteína interleuquina-1 β en el sistema *S. cerevisiae*, en el cual mientras la fusión de la proteína heteróloga a las porciones pre-pro del precursor de la feromona α -factor de *S. cerevisiae* no fue clivada, las fusiones con la región pre únicamente fueron correctamente procesadas durante la secreción (Ernst, 1988).

Si bien la glicosilación constituye una modificación pos-traduccional importante que se sabe, influye en aspectos variados de una proteína (plegamiento, solubilidad, antigenicidad, actividad biológica, estabilidad, así como también en las interacciones célula-célula (da Costa et al. 2010), se concluye también que la glicosilación del PSI de procirsina en su único sitio de glicosilación (NET) no es requisito imprescindible para su secreción en el sistema heterólogo *K. lactis*. La susceptibilidad a la endoglicosidasa Endo H mostrada por las glicoproteínas recombinantes obtenidas sin fusión al α -MF, una vez más sugirió que las cadenas de oligosacáridos no sufrieron modificaciones luego del pasaje por el retículo endoplasmático.

Un patrón de glicosilación sensible a Endo H se ha observado también durante la expresión inducible de cardosina B (*C. cardunculus*) en Arabidopsis transgénica y transiente en células epidermales de tabaco (da Costa et al., 2010), sistemas en los cuales además de acumularse en vacuola, cardosina B se detectó en la pared celular. Los autores plantearon que los glicanos resultan inaccesibles a las enzimas procesadoras del Golgi debido a la conformación de cardosina B o bien, que la proteína heteróloga consigue abandonar dicho compartimento muy tempranamente, antes de la acción de las enzimas. En el caso del PSI de cirsina secretado como proteína de fusión con el dominio α -MF la enzima Kex-2, encargada de procesar el α -MF en el *trans*-Golgi, actuó *in vitro* tanto sobre la muestra del PSI pre-tratada a 95 °C como sin tratar. Esto sugirió que el procesamiento incorrecto de la proteína de fusión *in vivo* (*K. lactis*) no se debió a su conformación.

A partir de estudios previamente realizados con profitepsina (*H. vulgare*), Okamoto (2006) propuso que las PAs típicas de plantas son transportadas a las vacuolas vía el complejo de Golgi, usando al PSI como señal de transporte y un posible receptor aún no identificado. Sin embargo, se sabe que proteasas tipo papaína que contienen una cola KDEL en su extremo C-terminal así como dos proteasas de Arabidopsis que no contienen dicha secuencia (la proteasa tipo papaína RD21 y la asparaginil-endopeptidasa γ VPE), son transportadas a la vacuola en una ruta independiente del aparato de Golgi. Las proteasas RD21 y γ VPE están localizadas en vesículas en forma de huso, denominadas “cuerpos del retículo endoplasmático” (en inglés, *ER-bodies*) las cuales derivan de este compartimento y se fusionan con las vacuolas sin pasar por el Golgi (Okamoto, 2006). Teniendo en cuenta que varios

autores sugieren diferentes funciones para el PSI dependiendo de la PA y la especie vegetal a la que ésta pertenece (Törmäkangas et al., 2001; Brodelius et al., 2005; Terauchi et al., 2006; Duarte et al., 2008) y que cirsina posee en su extremo C-terminal la secuencia VGFAEA—equivalente muy probablemente a la reconocida señal de direccionamiento vacuolar, VFAEAI (De Marcos Lousa et al., 2012)—no puede descartarse la hipótesis de que el precursor de cirsina alcance la vacuola en la planta mediante una vía similar a aquellas proteasas con la secuencia KDEL y que por tanto, el PSI expresado en forma aislada en un sistema heterólogo sea secretado sin pasar por el Golgi.

La similitud estructural de los dominios PSI con las saposinas, las cuales participan en el co-transporte de pro-catepsina D (PA animal) al lisosoma, y la capacidad evidenciada del PSI de interactuar con vesículas de fosfolípidos e inducir la liberación de su contenido en forma dependiente del pH y la composición lipídica (Egas et al., 2000; Bryksa et al., 2011) indica que los PSIs de PAs típicas de plantas median la interacción de sus precursores con las membranas fosfolipídicas e inducen su permeabilización. De acuerdo a esto, se sugirió que el PSI, sólo o en conjunto con la acción de la proteasa madre, podría actuar en la defensa contra patógenos y en la autólisis tardía de las células vegetales. Estos resultados sugieren que la función primaria del PSI es facilitar la asociación de los precursores a membrana, más probablemente en el lumen del retículo endoplasmático. Esta asociación podría, en parte, constituir un pre-requisito para el direccionamiento de las PAs a vacuola (da Costa et al., 2010). Las PAs típicas StAPs con actividad antimicrobiana las cuales son inducidas por condiciones de estrés biótico o abiótico, se acumulan en el fluido intercelular de los tubérculos de papa y hojas luego de una herida o infección (Guevara et al., 2002; Pagano et al., 2007). Tanto las StAPs de *S. tuberosum* como su dominio StAP-PSI recombinante, expresado en un sistema procariota con cola de polihistidinas poseen actividad citotóxica frente a las bacterias patógenas de humanos *Staphylococcus aureus*, *Bacillus cereus* y *E. coli* (Muñoz et al., 2010). Sin embargo, debido a que los resultados de dinámica molecular presentados en este trabajo de tesis sugieren que en el caso del PSI de procirsina las seis histidinas fusionadas al extremo C-terminal producirían un cambio conformacional significativo, se optó por realizar los ensayos antibacterianos con la proteína N86S-PSI de procirsina sin la cola de histidinas. Contrariamente a lo esperado, en estos ensayos el PSIN86S no presentó actividad contra los microorganismos ensayados. Mientras que StAP-PSI es similar a saposina C humana (Bryksa et al., 2011) N86S-PSI de procirsina tiene mayor identidad con saposina B estas diferencias en estructura primaria podrían justificar las diferencias en actividad observada entre ambos PSIs recombinantes.

En concordancia con la participación propuesta en la defensa de la planta, y la acción reportada para StAP-PSI sobre los hongos *Phytophthora infestans* y *Fusarium solani* (Muñoz, et al. 2010), los datos experimentales recabados en este trabajo mostraron para el PSI mutante del sitio de glicosilación del precursor de cirsina actividad antifúngica contra la especie fitopatógena *Alternaria alternata* aislada de viñedos. Cabe destacar que si bien, ésta fue la única especie claramente susceptible entre las ensayadas, la escasa difusión de moléculas grandes, como son las proteínas, en agar pudo limitar los resultados observados en el ensayo de Kirby-Bauer y por ende, no debe descartarse la acción del PSI de cirsina sobre las restantes especies.

El hongo *Alternaria alternata* es muy frecuentemente mencionado en literatura fitopatológica como causante de manchas en hojas y frutos o pudrición de estos últimos. Además del efecto directo sobre la calidad de los cultivos, esta especie despierta gran interés debido a su capacidad de producir micotoxinas que afectan a mamíferos y aves a través de la cadena alimentaria en los campos, durante el almacenamiento o manipulación de los cultivos contaminados (Bottalico and Logrieco, 1993; Carrillo, 2003; Tsitsigiannis et al., 2012). Por tanto, las consecuencias económicas de la contaminación por micotoxinas son profundas. Los cultivos con niveles inaceptables de micotoxinas frecuentemente deben ser destruidos o en forma alternativa, son derivados como forraje causando reducción del crecimiento, enfermedades y muertes entre animales susceptibles—además de producir carnes y leche contaminadas. *A. alternata* es descrita como una especie productora de varias toxinas, entre las más conocidas se encuentran altenueno, alternariol, alternariol monometil-éter, altertoxina I y fumonisinas, estas últimas producidas por numerosas especies de *Fusarium*. En particular, alimentos contaminados con fumonisin B1 (FB1) pueden causar alteraciones hepáticas y del sistema inmune en el ganado, bajos rendimientos en aves de corral, leucoencefalomalacia en caballos y edema pulmonar en cerdos. Mientras que maíz contaminado con FB1 ha sido asociado con cáncer esofágico en humanos en África, China y Estados Unidos de América. La similitud de las fumonisinas con esfingosinas sugieren un rol de estas toxinas como agentes de depleción de complejos de esfingolípidos desde las membranas biológicas causando su toxicidad y, quizás, su carcinogenicidad (Tsitsigiannis et al., 2012).

En los últimos tiempos, las restricciones sobre los métodos físicos y químicos empleados para la detoxificación de los productos agrícolas ocasionadas principalmente por los problemas relacionados a la salud (Köhl et al., 2011) así como también la resistencia a los fungicidas desarrollada por varios patógenos, han conducido a un aumento en el interés sobre el control biológico de las plagas, como una alternativa ecológica y no química (Tsitsigiannis et al., 2012).

En este sentido, el PSI del precursor de cirsina siendo sobreexpresado por una especie vegetal o por un microorganismo GRAS, como es el caso de *K. lactis*, representa una alternativa sumamente interesante que merece ser explorada en mayor profundidad.

Conclusiones y Perspectivas futuras



Conclusiones Generales y Perspectivas Futuras

En la presente tesis doctoral se aisló de flores inmaduras de *Cirsium vulgare* el ARNm que codifica para un precursor de proteasa aspártica típica. A partir del mismo, se obtuvo el correspondiente ADNc y el producto proteico codificado—preprocirsina—se expresó en la forma de zimógeno (*wild-type* y dos mutantes) en la cepa BL21 Star del sistema bacteriano *Escherichia coli*. En base a los datos obtenidos mediante los estudios realizados sobre procirsina recombinante, es posible arribar a las siguientes conclusiones:

- Procirsina recombinante mostró ser activa contra el sustrato clásicamente empleado en el estudio de las PAs clasificadas como típicas y comparte con otras proteasas de este tipo, las mismas características enzimáticas: máxima actividad a pH ácido, inhibición por pepstatina A y dependencia estricta de dos residuos catalíticos Asp para su actividad.
- Procirsina recombinante exhibió también una actividad coagulante de leche moderada y una alta especificidad de corte sobre la κ -caseína obteniéndose en la digestión de este sustrato, el perfil electroforético asignado al corte del enlace Phe105-Met106 responsable de desencadenar el proceso de coagulación de leche. Ambas propiedades son deseables en la manufactura de quesos dado que es sabido que una actividad proteolítica excesiva, produce características organolépticas indeseables.
- A pesar de no estar aún claramente descrita la actividad de una proenzima en el caso de las PAs típicas de plantas, la actividad catalítica encontrada para procirsina no es completamente inesperada ya que algunos precursores de PAs de mamíferos de la familia A1, como BACE y renina, muestran actividad enzimática significativa.
- Más aún, la región pro de procirsina parece adoptar conformaciones diferentes como propuso Ermolieff *et al.* (2000) para la proteasa aspártica BACE, permitiendo que el sitio activo se encuentre en equilibrio entre un estado abierto y uno cerrado y habilitando por tanto, el acceso transiente del sustrato a la hendidura catalítica.

- A partir de los resultados aquí presentados, se propone para el precursor de cirsina y PAs más cercanas, un mecanismo de inhibición de zimógenos en el cual a pH ácidos cuando los residuos Asp del sitio activo preformado se protonan, las interacciones con el par Lys11/Tyr13 del extremo N-terminal de la cirsina madura se rompen, dejando disponible el sitio catalítico.
- Aunque es activa en la forma del precursor, procirsina recombinante resultó incapaz de iniciar su propia activación bajo condiciones ácidas, sugiriendo que el procesamiento de procirsina hacia formas maduras (con el prosegmento y el inserto específico de planta PSI, removidos) requiere la acción de otras proteasas o al menos, de condiciones más próximas a las encontradas *in vivo*.

Por otro lado, se logró expresar el dominio tipo saposina (dominio PSI) presente en procirsina en la cepa GG799 del sistema de levaduras *Kluyveromyces lactis*.

Si bien la cepa GG799 consiguió expresar y secretar el dominio PSI hacia el medio extracelular como proteína de fusión soluble, la expresión en *K. lactis* de cuatro construcciones diferentes conteniendo la secuencia del PSI o de su mutante de sitio de glicosilación (N86S), dio lugar a las siguientes observaciones:

- Las formas de PSI fusionadas al dominio α -MF no son procesadas de manera eficiente en la vía de secreción de la célula de *K. lactis*.
- La presencia del dominio α -MF en la proteína recombinante introduce cuatro sitios de glicosilación extra, originando un perfil de glicosilación heterogéneo con formas de alto peso molecular y la coexistencia de varias poblaciones proteicas sumando complejidad al sistema.
- El dominio PSI consigue avanzar por la vía de secreción del sistema *K. lactis* una vez que es dirigido al retículo endoplasmático por el péptido señal α -mating factor, aún en ausencia de la región pro secretoria (dominio α -MF) fusionada a la secuencia del PSI.
- La glicosilación del único sitio putativo del PSI no influye de modo decisivo en su secreción de la célula GG799 de *K. lactis*.
- El dominio PSI del precursor de cirsina expresado en *K. lactis* presentó clara actividad inhibitoria del crecimiento micelar de la especie fitopatógena aislada de viñedos

Alternaria alternata, causante de manchas en hojas y frutos o pudrición de estos, y productora de micotoxinas que afectan la salud de mamíferos y aves a través de la cadena alimentaria en los campos con consecuencias económicas profundas.

En base a lo expuesto, se proponen las siguientes perspectivas futuras resultantes de esta tesis doctoral:

- Debido a la demanda creciente de cuajos alternativos de fuentes no animales en la industria alimentaria, resulta de interés avanzar hacia estudios sobre la actividad coagulante de leche que brinden más información respecto al uso potencial de procirsina en la producción de quesos de leche bovina, caprina y/o ovina de alta calidad (Roseiro et al., 2003; Reis and Malcata, 2011).
- Dada las restricciones sobre los métodos físicos y químicos empleados para la detoxificación de los productos agrícolas ocasionadas principalmente por los problemas relacionados a la salud así como también la resistencia a los fungicidas desarrollada por varios patógenos, y el consecuente aumento en el interés sobre el control biológico de las plagas como una opción ecológica y no química, el dominio PSI de procirsina siendo sobreexpresado por un microorganismos GRAS, como es el caso de *K. lactis*, representa una alternativa sumamente interesante para reducir el uso de pesticidas en el control del crecimiento fúngico y de toxinas de alto riesgo, que merece ser explorada en mayor profundidad.

Referencias



Referencias bibliográficas

- Ahn, V.E., Faull, K.F., Whitelegge, J.P., Fluharty, A.L., Privé, G.G., 2003. Crystal structure of saposin B reveals a dimeric shell for lipid binding. *Proceedings of the National Academy of Sciences of the United States of America* 100, 38–43.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P., 2002. *Molecular Biology of the Cell.*, Fourth Edi. ed. Galard Sciences, New York.
- Almeida, C.M., Pereira, C., Soares da Costa, D., Pereira, S., Pissarra, J., Faro, C., Simões, I., 2012. Chlpsin, a chloroplastidial aspartic proteinase from the green algae *Chlamydomonas reinhardtii*. *Planta* 236, 283–296.
- Anderson, D.H., Sawaya, M.R., Cascio, D., Ernst, W., Modlin, R., Krensky, A., Eisenberg, D., 2003. Granulysin crystal structure and a structure-derived lytic mechanism. *Journal of molecular biology* 325, 355–365.
- Andreeva, N.S., Rumsh, L.E.V.D., 2001. Analysis of crystal structures of aspartic proteinases : On the role of amino acid residues adjacent to the catalytic site of pepsin-like enzymes 2439–2450.
- Arnold, K., Bordoli, L., Kopp, J., Schwede, T., 2006. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* 22, 195–201.
- Asakura, T., Matsumoto, I., Funaki, J., Arai, S., Abe, K., 2000. The plant aspartic proteinase-specific polypeptide insert is not directly related to the activity of oryzasin 1. *European journal of biochemistry / FEBS* 267, 5115–22.
- Asakura, T., Watanabe, H., Keiko, A., Soichi, A., 1997. Oryzasin as an Aspartic Proteinase Occurring in Rice Seeds: Purification, Characterization, and Application to Milk Clotting. *Journal of agricultural and food chemistry* 45, 1070–1075.
- Athauda, S.B.P., Matsumoto, K., Rajapakshe, S., Kuribayashi, M., Kubomura-Yoshida, N., Kojima, M., Iwamatsu, A., Shibata, C., Inoue, H., Takahashi, K., 2004. Enzymic and structural characterization of nepenthesin, a unique member of a novel subfamily of aspartic proteinases. *Biochemical Journal* 306, 295–306.
- Athauda, S.B.P., Takahashi, K., 2002. Cleavage Specificities of Aspartic Proteinases toward Oxidized Insulin B Chain At Different pH Values. *Protein and Peptide Letters*, 9, 289–294.
- Barrett, A.J., Rawlings, N.D., Woessner, J.F. (Eds.), 2004. *Handbook of Proteolytic Enzymes*, Second. ed. Elsevier Academic Press, London.
- Belozersky, Mikhail A. Sarbakanova, S.T., Dunaevsky, Y.E., 1989. Aspartic proteinase from wheat seeds: isolation, properties and action on gliadin. *Planta* 177, 321–326.

- Benes, P., 2008. CathepsinD – many functions of aspartic protease. *Critical Reviews in Oncology/Hematology* 68, 12–28.
- Benkert, P., Biasini, M., Schwede, T., 2011. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* 27, 343–50.
- Bleukx, W., Brijs, K., Torrekens, S., Van Leuven, F., Delcour, J.A., 1998. Specificity of a wheat gluten aspartic proteinase. *Biochimica et biophysica acta* 1387, 317–24.
- Blom, N., Sicheritz-Pontén, T., Gupta, R., Gammeltoft, S., Brunak, S., 2004. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* 4, 1633–1649.
- Blundell, T.L., Guruprasad, K., Albert, A., Williams, M., Sibanda, B.L., Dhanaraj, V., 1998. The aspartic proteinases. A historical overview., in: James, M. (Ed.), *Aspartic Proteinases: Retroviral and Cellular Enzymes*. Plenum Press, New York, pp. 1–13.
- Bottalico, A., Logrieco, A., 1993. Mycotoxins in *Alternaria alternata* - infected olive fruits and their possible transfer into oil. *EPPO Bulletin* 23, 473–479.
- Bradford, M.B., 1976. A rapid and sensitive method for the quantitation of micrograms quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* 72, 248–254.
- Brake, a J., Merryweather, J.P., Coit, D.G., Heberlein, U. a, Masiarz, F.R., Mullenbach, G.T., Urdea, M.S., Valenzuela, P., Barr, P.J., 1984. Alpha-factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America* 81, 4642–6.
- Brodelius, P.E., Hiraiwa, M., Marttila, S., Karadaghi, S. Al, Picaud, S., Brodelius, M., 2005. Immunolocalization of saposin-like insert of plant aspartic. *Physiologia Plantarum* 125, 405–18.
- Bruhn, H., 2005. A short guided tour through functional and structural features of saposin-like proteins. *The Biochemical journal* 389, 249–57.
- Brutti, C.B., Pardo, M.F., Caffini, N.O., Natalucci, C.L., 2012. *Onopordum acanthium* L. (Asteraceae) flowers as coagulating agent for cheesemaking. *LWT - Food Science and Technology* 45, 172–178.
- Bryksa, B.C., Bhaumik, P., Magracheva, E., De Moura, D.C., Kurylowicz, M., Zdanov, A., Dutcher, J.R., Wlodawer, A., Yada, R.Y., 2011. Structure and mechanism of the saposin-like domain of a plant aspartic protease. *The Journal of biological chemistry* 286, 28265–75.
- Caplan, S., Green, R., Rocco, J., Kurjan, J., 1991. Glycosylation and Structure of the Yeast MFotd a-Factor Precursor Is Important for Efficient Transport through the Secretory Pathway. *Journal of Bacteriology* 173, 627–635.
- Carrillo, L. eonor, 2003. *Alternaria*, in: *LOS HONGOS DE LOS ALIMENTOS Y FORRAJES*. Universidad Nacional de Salta, pp. 81–86.

- Case, D.A., Darden, T.A., Cheatham, T.E., III, C.L.S., Wang, J., Duke, R.E., Luo, R., Walker, R.C., Zhang, W., Merz, K.M., Roberts, B., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K.F., Paesani, F., Vanicek, J., Liu, J., Wu, X., Brozell, S.R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Hsieh, M.-J., Cui, G., Roe, D.R., D.H.Mathews, Seetin, M.G., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., Kollman, P.A., 2010. AMBER 11.
- Castanheira, P., Samyn, B., Sergeant, K., Clemente, J.C., Dunn, B.M., Pires, E., Van Beeumen, J., Faro, C., 2005. Activation, proteolytic processing, and peptide specificity of recombinant cardosin A. *The Journal of biological chemistry* 280, 13047–54.
- Cesari, A., Falcinelli, A.L., Mendieta, J.R., Pagano, M.R., Mucci, N., Daleo, G.R., Guevara, M.G., 2007. Potato aspartic proteases (StAPs) exert cytotoxic activity on bovine and human spermatozoa. *Fertility and sterility* 88, 1248–55.
- Ciaffoni, F., Tatti, M., Boe, A., Salvioli, R., Fluharty, A., Sonnino, S., Vaccaro, A.M., Sonnino, S., B, A.M.V.S., 2006. Saposin B binds and transfers phospholipids. *Journal Of Lipid Research* 47, 1045–1053.
- Cimino, C. V, Liggieri, C.S., Priolo, N.S., Bruno, M.A., Vairo-Cavalli, S.E., 2010. *Arctium minus* (Hill) Bernh. (Asteraceae) aspartylendopeptidases with potential application in the formulation of nutraceutical products. *Molecular Medicinal Chemistry* 21, 11–16.
- Coates, L., Erskine, P.T., Mall, S., Gill, R., Wood, S.P., Myles, D. a a, Cooper, J.B., 2006. X-ray, neutron and NMR studies of the catalytic mechanism of aspartic proteinases. *European biophysics journal : EBJ* 35, 559–66.
- Coombs, G.H., Goldberg, D.E., Klemba, M., Berry, C., Kay, J., Mottram, J.C., 2001. Aspartic proteases of *Plasmodium falciparum* and other parasitic protozoa as drug targets. *Trends in parasitology* 17, 532–7.
- Cordeiro, M., Jakob, E., Puhán, Z., Pais, M.S., Brodelius, P.E., 1992. Milk clotting and proteolytic activities of purified cynarases from *Cynara cardunculus* –a comparison to chymosin. *Milchwissenschaft (Milk Science International)* 47, 683–687.
- Cordeiro, M.C., Xue, Z.-T., Pietrzak, M., Pais, M.S., Brodelius, P.E., 1994. Isolation and characterization of a cDNA from flowers of *Cynara cardunculus* encoding cyprosin (an aspartic proteinase) and its use to study the organ-specific expression of cyprosin. *Plant Molecular Biology* 24, 733–741.
- da Costa, D.S., Pereira, S., Moore, I., Pissarra, J., 2010. Dissecting cardosin B trafficking pathways in heterologous systems. *Planta* 232, 1517–30.
- Davies, D.R., 1990. The structure and function of the aspartic proteinases. *Annual Review of Biophysics and Biophysical Chemistry* 19, 189–215.
- de Alba, E., Weiler, S., Tjandra, N., 2003. Solution structure of human saposin C: pH-dependent interaction with phospholip vesicles. *Biochemistry* 42, 14729–14740.

- De Marcos Lousa, C., Gershlick, D.C., Denecke, J., 2012. Mechanisms and concepts paving the way towards a complete transport cycle of plant vacuolar sorting receptors. *The Plant cell* 24, 1714–32.
- de Rzedowski, Graciela C. Rzedowski, J., 2001. *Flora fanerogámica del Valle de México.*, 2nd ed. Instituto de ecología A.C., 2001, Pátzcuaro.
- Doi, E., Shibata, D., Matoba, T., Yonezawa, D., 1980. Characterization of pepstatin-sensitive acid protease in resting rice seeds. *Agric Biol Chem Journal* 44, 741–747.
- Domingos, A., Cardoso, P.C., Xue, Z.T., Clemente, A., Brodelius, P.E., Pais, M.S., 2000. Purification, cloning and autoproteolytic processing of an aspartic proteinase from *Centaurea calcitrapa*. *European journal of biochemistry / FEBS* 267, 6824–31.
- Duarte, P., Pissarra, J., Moore, I., 2008. Processing and trafficking of a single isoform of the aspartic proteinase cardosin A on the vacuolar pathway. *Planta* 227, 1255–68.
- Dunn, B.M., 2001. Determination of Protease Mechanism., in: Beynon, R., Bond, J.S. (Eds.), *Proteolytic Enzymes. A Practical Approach*. Oxford University Press, Oxford, pp. 77–104.
- Dunn, B.M., 2002. Structure and mechanism of the pepsin-like family of aspartic peptidases. *Chemical reviews* 102, 4431–58.
- Dunn, B.M., Hung, S., 2000. The two sides of enzyme ^ substrate speci c city : lessons from the aspartic proteinases. *Engineering* 1477, 231–240.
- Egas, C., Lavoura, N., Resende, R., Brito, R.M., Pires, E., de Lima, M.C., Faro, C., 2000. The saposin-like domain of the plant aspartic proteinase precursor is a potent inducer of vesicle leakage. *The Journal of biological chemistry* 275, 38190–6.
- Ermolieff, J., Loy, J.A., Koelsch, G., Tang, J., 2000. Proteolytic activation of recombinant pro-memapsin 2 (Pro-beta-secretase) studied with new fluorogenic substrates. *Biochemistry* 39, 12450–12456.
- Ernst, J.F., 1988. Efficient Secretion and Processing of Heterologous Proteins in *Saccharomyces cerevisiae* Is Mediated Solely by the Pre-Segment of α -Factor Precursor. *DNA* 7, 355–360.
- Esteves, C.L., 1995. Estudo comparativo das características bioquímicas dos coalhos de *Cynara cardunculus* L., *Cynara scolymus* L. e *Cynara humilis* L.
- Faro, C., Gal, S., 2005. Aspartic proteinase content of the *Arabidopsis* genome. *Current protein & peptide science* 6, 493–500.
- Faro, C., Veríssimo, P., Lin, Y., Tang, J., Pires, E., 1995. Cardosin A and cardosin B, aspartic proteases from the flowers of cardoon, in: Takahashi, K. (Ed.), *Aspartic Proteinases: Structure, Function, Biology, and Biomedical Implications*. Plenum Press, New York, pp. 373–377.
- Faust, P.L., Kornfeld, S., Chirgwin, J.M., 1985. Cloning and sequence analysis of cDNA for human cathepsin D *Biochemistry* : 82, 4910–4914.

- Feijoo-Siota, L., Villa, T., 2011. Native and Biotechnologically Engineered Plant Proteases with Industrial Applications. *Food and Bioprocess Technology* 4, 1066–1088.
- Foltman, B., 1988. Structure and function of propeptides in zymogens for aspartic proteinases. *Biological Chemistry Hoppe Seyler* 369, 311–314.
- Francky, A., Francky, B.M., Strukelj, B., Gruden, K., A., R., Krizaj, I., Kregar, I., Pain, R.H., Pungercar, J., 2001. A basic residue at position 36p of the propeptide is not essential for the correct folding and subsequent autocatalytic activation of prochymosin. *European Journal of Biochemistry* 268, 2362–2368.
- Frazão, C., Bento, I., Costa, J., Soares, C.M., Veríssimo, P., Faro, C., Pires, E., Cooper, J., Carrondo, M. a, 1999. Crystal structure of cardosin A, a glycosylated and Arg-Gly-Asp-containing aspartic proteinase from the flowers of *Cynara cardunculus* L. *The Journal of biological chemistry* 274, 27694–701.
- Friedman, R., Caflisch, A., 2010. On the orientation of the catalytic dyad in aspartic proteases. *Proteins* 78, 1575–82.
- Fruton, J., 2002. A history of pepsin and related enzymes. *Quarterly Review of Biology* 77, 127–147.
- Gevaert, K., Goethals, M., Martens, L., Van Damme, J., Staes, A., Thomas, G.R., Vandekerckhove, J., 2003. Exploring proteomes and analyzing protein processing by mass spectrometric identification of sorted N-terminal peptides. *Nature biotechnology* 21, 566–9.
- Glathe, S., Kervinen, J., Nimtz, M., Li, G.H., Tobin, G.J., Copeland, T.D., Ashford, D. a, Wlodawer, a, Costa, J., 1998. Transport and activation of the vacuolar aspartic proteinase phytepsin in barley (*Hordeum vulgare* L.). *The Journal of biological chemistry* 273, 31230–6.
- Glickman, J., Kornfeld, S., 1993. Mannose 6-phosphate-independent targeting of lysosomal enzymes in I-cell disease B lymphoblasts. *The journal of cell biology* 123, 99–108.
- González-Rábade, N., Badillo-Corona, J.A., Aranda-Barradas, J.S., Oliver-Salvador, M.D.C., 2011. Production of plant proteases in vivo and in vitro - A review. *Biotechnology advances* 1–14.
- Guevara, M.G., Daleo, G.R., Oliva, C.R., 2001. Purification and characterization of an aspartic protease from potato leaves. *Physiologia plantarum* 112, 321–326.
- Guevara, M.G., Oliva, C.R., Huarte, M., Daleo, G.R., 2002. An aspartic protease with antimicrobial activity is induced after infection and wounding in intercellular fluids of potato tubers. *European Journal of Plant Pathology* 131–137.
- Guevara, M.G., Oliva, C.R., Machinandiarena, M., Daleo, G.R., 1999. Purification and properties of an aspartic protease from potato tuber that is inhibited by a basic chitinase. *Physiologia Plantarum* 106, 164–169.

- Guevara, M.G., Veríssimo, P., Pires, E., Faro, C., Daleo, G.R., 2004. Potato aspartic proteases: induction, antimicrobial activity and substrate specificity. *Journal of plant pathology* 86, 233–238.
- Guindon, S., Delsuc, F., Dufayard, J.F., Gascuel, O., 2009. Estimating maximum likelihood phylogenies with PhyML. *methods in molecular biology* 537, 113–137.
- Guruprasad, K., Törmäkangas, K., Kervinen, J., Blundell, T.L., 1994. Comparative modelling of barley-grain aspartic proteinase: a structural rationale for observed hydrolytic specificity. *FEBS letters* 352, 131–6.
- Hartsuck, J., Koelsch, G., Remington, S., 1992. The high-resolution crystal structure of porcine pepsinogen. *Proteins* 13, 1–25.
- Hecht, O., van Nuland, N., Schleinkofer, K., Dingley, A., Bruhn, H., Leippe, M., Grötzinger, J., 2004. Solution structure of the pore-forming protein of *Entamoeba histolytica*. *The Journal of biological chemistry* 279, 17834–17841.
- Heimgartner, U., Pietrzak, M., Geertsen, R., Brodelius, P., da Silva Figueiredo, A.C., Pais, M.S., 1990. Purification and partial characterization of milk clotting proteases from flowers of *Cynara cardunculus*. *Phytochemistry* 29, 1405–1410.
- Horimoto, Y., Dee, D.R., Yada, R.Y., 2009. Multifunctional aspartic peptidase prosegments. *New biotechnology* 25, 318–24.
- Huesgen, P.F., Overall, C.M., 2012. N- and C-terminal degradomics: new approaches to reveal biological roles for plant proteases from substrate identification. *Physiologia plantarum* 145, 5–17.
- Hussain, I., Powell, D., Howlett, D.R., Tew, D.G., Meek, T.D., Chapmand, C., Gloger, I.S., Murphy, K.E., Southan, C.D., Ryan, D.M., Smith, T.S., Simmons, D.L., Walsh, F.S., Dingwall, C., Christie, G., 1999. Identification of a Novel Aspartic Protease (Asp 2) as β -Secretase. *Molecular and Cellular Neurosciences* 14, 419–427.
- Hyslop, D., 2003. Enzymatic coagulation of milk, in: Fox, P.F., Mcsweeney, P.L.H. (Eds.), *Advanced Dairy Chemistry. Volumen 1: Proteins*. Kluwer Academic/Plenum Publishers, Berlin, pp. 839–878.
- Jacob, M., Jaros, D., Rohm, H., 2011. Recent advances in milk clotting enzymes. *International Journal of Dairy Technology* 64, 14–33.
- James, M.N., 2004. Catalytic pathway of aspartic peptidases, in: Barret, A.J., Rawlings, N.D., Woessner, J.F. (Eds.), *Handbook of Proteolytic Enzymes*. Elsevier Academic Press, London, pp. 12–19.
- Jaroszewski, L., Rychlewski, L., Li, Z., Li, W., Godzik, A., 2005. FFAS03: a server for profile--profile sequence alignments. *Nucleic Acids Res* 33(Web Server issue).
- Julenius, K., Mølgaard, A., Gupta, R., Brunak, S., 2005. Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology* 15, 153–64.

- Kervinen, J., Sarkkinen, P., Kalkkinen, N., Mikola, L., Saarma, M., 1993. Hydrolytic specificity of the barley grain aspartic proteinase. *Phytochemistry* 32, 799–803.
- Kervinen, J., Tobin, G.J., Costa, J., Waugh, D.S., Wlodawer, a, Zdanov, a, 1999. Crystal structure of plant aspartic proteinase prophytepsin: inactivation and vacuolar targeting. *The EMBO journal* 18, 3947–55.
- Kervinen, J., Wlodawer, A., Zdanov, A., 2004. Phytepsin, in: Barret, A., Rawlings, N., Woessner, J. (Eds.), *Handbook of Proteolytic Enzymes*. Elsevier Academic Press, London, pp. 77–78.
- Khazanovich Bernstein, N., James, M.N., 1999. Novel ways to prevent proteolysis — prophytepsin and proplasmepsin II. *Current Opinion in Structural Biology* 9, 684–689.
- Kiefer, F., Arnold, K., Künzli, M., Bordoli, L., Schwede, T., 2009. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Research*. 37, D387–D392.
- Kishimoto, Y., Hiraiwa, M., O'Brien, J., 1992. Saposins: structure, function, distribution, and molecular genetics. *The Journal of Lipid Research* 33, 1255–1267.
- Kleifeld, O., Doucet, A., auf dem Keller, U., Prudova, A., Schilling, O., Kainthan, R.K., Starr, A.E., Foster, L.J., Kizhakkedathu, J.N., Overall, C.M., 2010. Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products. *Nature biotechnology* 28, 281–8.
- Klemba, M., Goldberg, D., 2002. Biological roles of proteases in parasitic protozoa. *ANNUAL REVIEW OF BIOCHEMISTRY* 71, 275–305.
- Köhl, J., Postma, J., Nicot, P., Ruocco, M., Blum, B., 2011. Stepwise screening of microorganisms for commercial use in biological control of plant-pathogenic fungi and bacteria. *Biological Control* 57, 1–12.
- Liepinsh, E., Andersson, M., Ruysschaert, J.M., Otting, C., 1997. Saposin fold revealed by the NMR structure of NL-lysin. *Natural Structural & Molecular Biology* 4, 793–795.
- Lin, X., Koelsch, G., Wu, S., Downs, D., Dashti, A., Tang, J., 2000. Human aspartic protease memapsin 2 cleaves the beta-secretase site of beta-amyloid precursor protein. *Proceedings of the National Academy of Sciences of the United States of America* 97, 1456–1460.
- Llorente, B.E., Brutti, C.B., Caffini, N.O., 2004. Purification and characterization of a milk-clotting aspartic proteinase from globe artichoke (*Cynara scolymus* L.). *Journal of agricultural and food chemistry* 52, 8182–9.
- López-Otín, C., Overall, C.M., 2002. Protease degradomics: a new challenge for proteomics. *Nature reviews. Molecular cell biology* 3, 509–19.
- Lufrano, D., Cavalli, S.V., Priolo, N., 2007. Acción de peptidasas aspárticas de flores de *Carduus acanthoides* y *Carduus thoermeri* sobre caseinatos bovino y caprino 52–55.

- Lufrano, D., Faro, R., Castanheira, P., Parisi, G., Veríssimo, P., Vairo-Cavalli, S., Simões, I., Faro, C., 2012. Molecular cloning and characterization of procirsin, an active aspartic protease precursor from *Cirsium vulgare* (Asteraceae). *Phytochemistry* 81, 7–18.
- Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Lu, F., Marchler, G.H., Mullokandov, M., Omelchenko, M. V, Robertson, C.L., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Zheng, C., Bryant, S.H., 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic acids research* 39, D225–9.
- Marzocca, A., Marsico, O.J. V., del Puerto, O., 1976. Marzocca, A. 1976. Manual de malezas. Hemisferio Sur. Buenos Aires, Argentina. Hemisferio Sur, Buenos Aires.
- Mazorra-Manzano, M. a, Tanaka, T., Dee, D.R., Yada, R.Y., 2010. Structure-function characterization of the recombinant aspartic proteinase A1 from *Arabidopsis thaliana*. *Phytochemistry* 71, 515–23.
- Mazorra-Manzano, M. a, Yada, R.Y., 2008. Expression and characterization of the recombinant aspartic proteinase A1 from *Arabidopsis thaliana*. *Phytochemistry* 69, 2439–48.
- Mendieta, J.R., Fimognari, C., Daleo, G.R., Hrelia, P., Guevara, M.G., 2010. Cytotoxic effect of potato aspartic proteases (StAPs) on Jurkat T cells. *Fitoterapia* 81, 329–35.
- Mendieta, J.R., Pagano, M.R., Muñoz, F.F., Daleo, G.R., Guevara, M.G., 2006. Antimicrobial activity of potato aspartic proteases (StAPs) involves membrane permeabilization. *Microbiology (Reading, England)* 152, 2039–47.
- Moore, S.A., Sielecki, A.R., Chernai, M.M., Tarasova, N.I., James, M.N.G., 1995. Crystal and molecular structures of human progastricsin at 1.62 Å resolution. *Journal of molecular biology* 247, 466–485.
- Mullis, K.B., 1990. Target amplification for DNA analysis by the polymerase chain reaction. *Annales de Biologie Clinique (Paris)* 48, 579–582.
- Munford, R.S., Sheppard, P.O., O'Hara, P.J., 1995. Saposin-like proteins (SAPLIP) carry out diverse functions on a common backbone structure. *Journal of lipid research* 36, 1653–63.
- Muñoz, F.F., Mendieta, J.R., Pagano, M.R., Paggi, R. a, Daleo, G.R., Guevara, M.G., 2010. The swaposin-like domain of potato aspartic protease (StAsp-PSI) exerts antimicrobial activity on plant and human pathogens. *Peptides* 31, 777–85.
- Muñoz-Pinedo, C., 2012. Signaling pathways that regulate life and cell death: evolution of apoptosis in the context of self-defense. *Adv Exp Med Biol.* 738, 124–43.
- Mutlu, A., Gal, S., 1999. Plant aspartic proteinases: enzymes on the way to a function. *Physiologia Plantarum* 105, 569–576.
- O'Brien, J., Kishimoto, Y., 1991. proteins : and role in human lysosomal storage disorders. *The Journal of the Federation of American Societies for Experimental Biology* 5, 301–308.

- O'Connell, K.L., Stults, J.T., 1997. Identification of mouse liver proteins on two-dimensional electrophoresis gels by matrix-assisted laser desorption/ionization mass spectrometry of in situ enzymatic digests. *ELECTROPHORESIS* 18, 349–359.
- Okamoto, T., 2006. Transport of Proteases to the Vacuole: ER Export Bypassing Golgi? *Plant Cell*.
- Pagano, M.R., Mendieta, J.R., Muñoz, F.F., Daleo, G.R., Guevara, M.G., 2007. Roles of glycosylation on the antifungal activity and apoplast accumulation of StAPs (*Solanum tuberosum* aspartic proteases). *International journal of biological macromolecules* 41, 512–20.
- Park, H., Kusakabe, I., Sakakibara, Y., Kobayashi, H., 2001. Autoproteolytic processing of aspartic proteinase from sunflower seeds. *Bioscience, Biotechnology and Biochemistry*. 65.
- Park, H., Yamanaka, N., Mikkonen, A., Kusakabe, I., Kobayashi, H., 2000. Purification and characterization of aspartic proteinase from sunflower seeds. *Bioscience, biotechnology, and biochemistry* 64, 931–9.
- Payie, K.G., Tanaka, T., Gal, S., Yada, R.Y., 2003. Construction, expression and characterization of a chimaeric mammalian-plant aspartic proteinase. *The Biochemical journal* 372, 671–8.
- Pereira, C.S., da Costa, D.S., Pereira, S., de Moura Nogueira, F., Albuquerque, P.M., Teixeira, J., Faro, C., Pissarra, J., 2008. Cardosins in postembryonic development of cardoon: towards an elucidation of the biological function of plant aspartic proteinases. *Protoplasma* 232, 203–213.
- Pesquet, E., 2012. Plant proteases – from detection to function. *Physiologia Plantarum* 145, 1–4.
- Petersen, T.N., Brunak, S., Von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* 8, 785–786.
- Piatigorsky, J., Norman, B., Dishaw, L.J., Kos, L., Horwitz, J., Steinbach, P.J., Kozmik, Z., 2001. J3-crystallin of the jellyfish lens: similarity to saposins. *Proceedings of the National Academy of Sciences of the United States of America* 98, 12362–12367.
- Pimentel, C., Van Der Straeten, D., Pires, E., Faro, C., Rodrigues-Pousada, C., 2007. Characterization and expression analysis of the aspartic protease gene family of *Cynara cardunculus* L. *The FEBS journal* 274, 2523–39.
- Pitarresi, T.M., Rubattu, S., Heinrikson, R., Sealey, J.E., 1992. Reversible cryoactivation of recombinant human prorenin. *The Journal of biological chemistry* 267, 11753–9.
- Polanowski, A., Wilusz, T., Kolaczowska, M., Wieczorek, M., Wilimowska-Pelc, A., 1985. Purification and characterisation of aspartic proteinases from *Cucumis sativus* and *Curcubita maxima* seeds., in: Kosta, V., Societies, F. of E.B. (Eds.), *Aspartic Proteinases and Their Inhibitors*. Walter de Gruyter, Berlin, pp. 49–52.

- Ponting, C.P., Russell, R.B., 1995. Swaposins: circular permutations within genes encoding saposin homologues. *Trends in Biochemical Sciences* 20.
- Puente, X.S., Sánchez, L.M., Overall, C.M., López-Otín, C., 2003. Human and mouse proteases: a comparative genomic approach. *Nature reviews. Genetics* 4, 544–58.
- Radlowski, M., Kalinowski, A., Adamczyk, J., Krolikowski, Z., Bartkowiak, S., 1996. Proteolytic activity in the maize pollen wall. *Physiologia Plantarum* 98, 172–178.
- Ramalho-Santos, M., Veríssimo, P., Cortes, L., Samyn, B., Van Beeumen, J., Pires, E., Faro, C., 1998. Identification and proteolytic processing of procardosin A. *European journal of biochemistry / FEBS* 255, 133–8.
- Raposo, S., Domingos, a, 2008. Purification and characterization milk-clotting aspartic proteinases from *Centaurea calcitrapa* cell suspension cultures. *Process Biochemistry* 43, 139–144.
- Rawlings, N.D., Barrett, A.J., 2004. Introduction: aspartic peptidases and their clans, in: Barrett, A.J., Rawlings, N.D., Woessner, J.F. (Eds.), *Handbook of Proteolytic Enzymes*. Elsevier Academic Press, London, pp. 3–12.
- Rawlings, N.D., Bateman, A., Barrett, A.J., 2012. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* 40, D343–50.
- Reis, P.J.M., Malcata, F.X., 2011. Current state of Portuguese dairy products from ovine and caprine milks. *Small Ruminant Research* 101, 122–133.
- Richter, C., Tanaka, T., Koseki, T., Yada, R.Y., 1999. Contribution of a prosegment lysine residue to the function and structure of porcine pepsinogen A and its active form pepsin A. *European journal of biochemistry / FEBS* 261, 746–52.
- Richter, C., Tanaka, T., Yada, R.Y., 1998. Mechanism of activation of the gastric aspartic proteinases: pepsinogen, progastricsin and prochymosin. *The Biochemical journal* 335 (Pt 3, 481–90.
- Rijnboutt, S., Kal, A., Geuze, H., Aerts, H., Strous, G., 1991. Mannose 6-phosphate-independent targeting of cathepsin D to lysosomes in HepG2 cells. *The Journal of Biological Chemistry* 266, 23586–23592.
- Robbins, A., Dunn, B., Agbandje-McKenna, M McKenna, R., 2009. Crystallographic evidence for noncoplanar catalytic aspartic acids in plasmepsin II resides in the Protein Data Bank. *Acta Crystallographica. Section D, Biological Crystallography* 65, 294–296.
- Rodrigo, I., Van Loon, L.C., Vera, P., Conejero, V., 1991. Degradation of Tobacco Pathogenesis-Related Proteins1 Proteins in Plants. *Plant physiology* 95, 616–622.
- Rodrigo, I., Vera, P., Conejero, V., 1989. Degradation of tomato pathogenesis-related proteins by an endogenous 37-kDa aspartyl endoproteinase. *European journal of biochemistry / FEBS* 184, 663–9.

- Roseiro, L.B., Barbosa, M., Ames, J.M., Wilbey, R.A., 2003. Cheesemaking with vegetable coagulants-the use of *Cynara L.* for the production of ovine milk cheeses. *International Journal of Dairy Technology* 56, 76–85.
- Runeberg-Roos, P., Kervinen, J., Kovaleva, V., Raikhel, N. V, Gal, S., 1994. The aspartic proteinase of barley is a vacuolar enzyme that processes probarley lectin in vitro. *Plant physiology* 105, 321–9.
- Runeberg-Roos, P., Tormakangas, K., Ostman, A., 1991. Primary structure of a barley-grain aspartic proteinase. A plant aspartic proteinase resembling mammalian cathepsin D. *European Journal of Biochemistry* 202, 1021–1027.
- Sali, A., Blundell, T.L., 1993. Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology* 234, 779–815.
- Salvesen, G.S., Nagase, H., 2001. Inhibition of proteolytic enzymes., in: Beynon, R., Bond, J.S. (Eds.), *Proteolytic Enzymes. A Practical Approach*. Oxford University Press, Oxford, pp. 105–130.
- Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Preparation of plasmid DNA by alkaline lysis with SDS: Minipreparation., in: *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York, pp. 1.32–1.34.
- Sampaio, P.N., Fortes, A.M., Cabral, J.M.S., Pais, M.S., Fonseca, L.P., 2008. Production and Characterization of Recombinant Cyprosin B in *Saccharomyces cerevisiae* (W303-1A) Strain. *Journal of Bioscience and Bioengineering* 105, 305–312.
- Sarmiento, A.C., Lopes, H., Oliveira, C.S., Vitorino, R., Samyn, B., Sergeant, K., Debyser, G., Van Beeumen, J., Domingues, P., Amado, F., Pires, E., Domingues, M.R.M., Barros, M.T., 2009. Multiplicity of aspartic proteinases from *Cynara cardunculus L.* *Planta* 230, 429–39.
- Schaller, A., 2004. A cut above the rest: the regulatory function of plant proteases. *Planta* 220, 183–97.
- Shi, X.P., Chen, E., Yin, K.C., Na, S., Garsky, V.M., Lai, M.T., Li, Y.M., Platchek, M., Register, R.B., Sardana, M.K., Tang, M.J., Thiebeau, J., Wood, T., Shafer, J.A., Gardell, S.J., 2001. The pro domain of beta-secretase does not confer strict zymogen-like properties but does assist proper folding of the protease domain. *The Journal of biological chemistry* 276, 10366–73.
- Sidrach, L., García-Cánovas, F., Tudela, J., Rodríguez-López, J., 2006. Purification of cynarases from artichoke (*Cynara scolymus L.*): enzymatic properties of cynarase A. *Phytochemistry* 66, 41–49.
- Sieleki, A., Fujinaga, M., Read, R., James, M., 1991. Refined structure of porcine pepsinogen at 1.8 Å resolution. *Journal of molecular biology* 219, 671–692.
- Simões, I., Faro, C., 2004. Structure and function of plant aspartic proteinases. *European journal of biochemistry / FEBS* 271, 2067–75.

- Simões, I., Faro, R., Bur, D., Faro, C., 2007. Characterization of recombinant CDR1, an Arabidopsis aspartic proteinase involved in disease resistance. *The Journal of biological chemistry* 282, 31358–65.
- Sinha, S., Anderson, J., Barbour, R., Basi, G., Caccavello, R., Davis, D., Doan, M., Dovey, H., Frigon, N., Hong, J., Jacobson-Croak, K., Jewett, N., Keim, P., Knops, J., Lieberburg, I., Power, M., Tan, H., Tatsuno, G., Tung, J., Schenk, D., Seubert, P., Suomensari, S., Wang, S., Walker, D., Zhao, J., McConlogue, L., John, V., 1999. Purification and cloning of amyloid precursor protein beta-secretase from human brain. *Natural Structural & Molecular Biology* 402, 537–540.
- Soding, J., Remmert, M., Biegert, A., Lupas, A.N., 2006. HHSenser: exhaustive transitive profile search using HMM-HMM comparison. *Nucleic Acids Res* 34(Web Server issue).
- St. Angelo, A., Ory, R., Hansen, H., 1970. Properties of a purified proteinase from hempseed. *Phytochemistry* 9, 1933–1938.
- Staes, A., Van Damme, P., Helsen, K., Demol, H., Vandekerckhove, J., Gevaert, K., 2008. Improved recovery of proteome-informative, protein N-terminal peptides by combined fractional diagonal chromatography (COFRADIC). *Proteomics* 8, 1362–70.
- Stenger, S., Hanson, D., Teitelbaum, R., Dewan, P., Niazi, K., Froelich, C., Ganz, T., Thoma-Uszynski, S., Melián, A., Bogdan, C., Porcelli, S., Bloom, B., Krensky, A., Modlin, R., 1998. An Antimicrobial Activity of Cytolytic T Cells Mediated by Granulysin. *Science* 282, 121–125.
- Tamer, M., 1993. Identification and partial purification of a novel milk clotting enzyme from *Onopordum turcicum*. *Biotechnology Letters* 15, 427–32.
- Tamura, T., Terauchi, K., Kiyosaki, T., Asakura, T., Funaki, J., Matsumoto, I., Misaka, T., Abe, K., 2007. Differential expression of wheat aspartic proteinases, WAP1 and WAP2, in germinating and maturing seeds. *Journal of plant physiology* 164, 470–477.
- Tang, J., James, M.N., Hsu, I.N., Jenkins, J.A., Blundell, T.L., 1978. Structural evidence for gene duplication in the evolution of the acid proteases. *Nature* 271, 618–621.
- Tang, J., Wong, R.N., 1987. Evolution in the structure and function of aspartic proteases. *Journal of Cell Biochemistry* 33, 53–63.
- Tatnell, P.J., Powell, D.J., Hill, J., Smith, T.S., Tew, D.G., Kay, J., 1998. Napsins: new human aspartic proteinases. Distinction between two closely related genes. *FEBS letters* 441, 43–8.
- Terauchi, K., Asakura, T., Ueda, H., Tamura, T., Tamura, K., Matsumoto, I., Misaka, T., Hara-Nishimura, I., Abe, K., 2006. Plant-specific insertions in the soybean aspartic proteinases, soyAP1 and soyAP2, perform different functions of vacuolar targeting. *Journal of plant physiology* 163, 856–62.
- Timotijević, G.S., Radović, S.R., Maksimović, V.R., 2006. ASPARTIC PROTEINASES FROM BUCKWHEAT (FAGOPYRUM ESCULENTUM MOENCH) SEEDS - PURIFICATION AND PROPERTIES OF THE 47 kDa ENZYME. *Pharmacia* 58, 171–177.

- Tormakangas, K., Hadlington, J.L., Pimpl, P., Hillmer, S., Brandizzi, F., Teeri, T.H., Denecke, J., 2001. A vacuolar sorting domain may also influence the way in which proteins leave the endoplasmic reticulum. *Plant Cell* 13, 2021–2032.
- Tsiatsiani, L., Gevaert, K., Van Breusegem, F., 2011. Natural substrates of plant proteases: how can protease degradomics extend our knowledge? *Physiologia plantarum* 28–40.
- Tsitsigiannis, I., Antoniou, P., Tjamos, C., 2012. Biological control strategies of mycotoxigenic fungi and associated mycotoxins in Mediterranean basin crops.
- Vaccaro, A.M., Salvioli, R., Tatti, M., Ciaffoni, F., 1999. Saposins and their interaction with lipids. *Neurochemical Research* 24, 307–314.
- Vairo Cavalli, S.E., 2005. Aspartilendopeptidasas de *Silybum marianum* (L.) Gaertn. con potencial aplicación en la industria láctea.
- Vairo Cavalli, S.E., Lufrano, D., Cimino, C. V, Faro, C., Veríssimo, P., Pires, E., 2009. Proteasas de Asteraceae, in: Caffini, N.O. (Ed.), *Enzimas Proteolíticas De Vegetales Superiores. Aplicaciones Industriales. Ciencia y Tecnología para el Desarrollo* (Cytel), Buenos Aires, pp. 25–40.
- van der Hoorn, R.A.L., 2008. Plant Proteases: From Phenotypes to Molecular Mechanisms. *The Annual Review of Plant Biology* 59, 191–223.
- Varón, R., Minaya-Pacheco, M.A., García-Molina, F., Arribas, E., Arias, E., Masiá, J., García-Sevilla, F., 2006. Competitive and uncompetitive inhibitors simultaneously acting on an autocatalytic zymogen activation reaction. *Journal of enzyme inhibition and medicinal chemistry* 21, 635–45.
- Vassar, R., Bennett, B., Babu-Khan, S., Kahn, S., Mendiaz, E., Denis, P., Teplow, D., Ross, S., Amarante, P., Loeloff, R., Luo, Y., Fisher, S., Fuller, J., Edenson, S., Lile, J., Jarosinski, M., Biere, A., Curran, E., Burgess, T., Louis, J., Collins, F., Treanor, J., Rogers, G., Citron, M., 1999. Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science* 286, 735–741.
- Veerapandian, B., Cooper, J.B., Sali, a, Blundell, T.L., Rosati, R.L., Dominy, B.W., Damon, D.B., Hoover, D.J., 1992. Direct observation by X-ray analysis of the tetrahedral "intermediate" of aspartic proteinases. *Protein science : a publication of the Protein Society* 1, 322–8.
- Veríssimo, P., Esteves, C.L., Faro, C.J., Pires, E., 1995. The vegetable rennet of *Cynara cardunculus* L. contains two proteinases with chymosin and pepsin-like specificities. *Biotechnology Letters* 17, 621–626.
- Veríssimo, P., Faro, C., Moir, a J., Lin, Y., Tang, J., Pires, E., 1996. Purification, characterization and partial amino acid sequencing of two new aspartic proteinases from fresh flowers of *Cynara cardunculus* L. *European journal of biochemistry / FEBS* 235, 762–8.
- Vieira, M., Pissarr, J., Veríssimo, P., Castanheira, P., Costa, Y., Pires, E., Faro, C., 2001. Molecular cloning and characterization of cDNA encoding cardosin B, an aspartic proteinase accumulating extracellularly in the transmitting tissue of *Cynara cardunculus* L. *Plant molecular biology* 45, 529–39.

- Walstra, P., Geurts, T., Noomen, A., Jellema, A., van Boekel, M., 1999. Dairy technology: principles of milk properties and processes. Marcel Dekker Inc., New York.
- Wang, L.-S., Xia, L., Shen, S.-M., Zheng, Y., Yu, Y., Chen, G.-Q., 2012. Dissecting cell death with proteomic scalpels. *Proteomics* 12, 597–606.
- Weber, I.T., Wang, Y.-F., 2010. HIV-1 Protease: Role in Viral Replication, Protein-Lig and X-Ray Crystal Structures and Inhibitor Design, in: Ghosh, A.K. (Ed.), *Aspartic Acid Proteases as Therapeutic Targets*. Wiley-VCH Verlag & Co., Weinheim.
- White, P.C., Cordeiro, M.C., Arnold, D., Brodelius, P.E., Kay, J., 1999. Processing, activity, and inhibition of recombinant cyprosin, an aspartic proteinase from cardoon (*Cynara cardunculus*). *The Journal of biological chemistry* 274, 16685–93.
- Wiederstein, M., Sippl, M.J., 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 35(Web Server issue).
- Wilimowska-Pelc, A Polanowski, A., Kołaczowska, M., Wieczorek, M., Wilusz, T., 1983. Aspartyl proteinase from cucumber (*Cucumis sativus*) seeds. Preparation and characteristics. *Acta Biochimica Polonia* 30, 23–31.
- Yan, R., Bienkowski, M., Shuck, M., Miao, H., Tory, M., Pauley, A., Brashier, J., Stratman, N., Mathews, W., Buhl, A., Carter, D., Tomasselli, A., Parodi, L., Heinrikson, R., Gurney, M., 1999. Membrane-anchored aspartyl protease with Alzheimer's disease beta-secretase activity. *Nature* 402, 533–537.
- Zhu, Y., Conner, G., 1994. Intermolecular association of lysosomal protein precursors during biosynthesis. *The journal of biological chemistry* 269, 3846–51.

Anexos



Alineamiento diferentes clones de pGEM-preproCV obtenidos

```

clon8 (*)      MGTSIKASLLALFLFFLLSPTAISVSN DGLIRVGLKRRKVDQINQLSGHGASMEGKARKD 60
clon5         MGTSIKASLLALFLFFLLSPTAFSVSN DGLIRVGLKRRKVDQINQLSGHGASMEGKARKD 60
clon12       MGTSIKASLLALFLFFLLSPTAFSVSN DGLIRVGLKRRKVDQINQLSGHGASMEGKARKD 60
clon2        MGTSIKASLLALFLFFLLSPTAFSVSN DGLIRVGLKRRKVDQINQLSGHGASMEGKARKD 60
clon11       -GYLNQASLLALFLFFLLSPTAFSVSN DGLIRVGLKRRKVDQINQLSGHGASMEGKARKD 59
               * :*****:*****:*****:*****:*****:*****:*****:*****

clon8 (*)      FGGGTLRDSDDIIALKNYMDAQY YGEIGIGAPPQKFTVIFDTGSSNLWVPSAKCYFSV 120
clon5         FGI GGTLRDSDDIIALKNYMDAQY YGEIGIGAPPQKFTVIFDTGSSNLWVPSAKCYFSV 120
clon12       FGI GGTLRDSDDIIALKNYMDAQY YGEIGIGAPPQKFTVIFDTGSSNLWVPSAKCYFSV 120
clon2        FGI GGTLRDSDDIIALKNYMDAQY YGEIGIGAPPQKFTVIFDTGSSNLWVPSAKCYFSV 120
clon11       FGI GGTLRDSDDIIALKNYMDAQY YGEIGIGAPPQKFTVIFDTGSSNLWVPSAKCYFSV 119
               **.:*****:*****:*****:*****:*****:*****:*****:*****

clon8 (*)      ACLFH SKYKSSHSSTYKKNGTSAAIQYGTGSI SGFVSQDSVQLGDLVVKEQDFIEATKEP 180
clon5         ACLFH SKYKSSKSSTYKKNGTSAAIQYGTGSI SGFVSQDSVQLGDLVVKEQDFIEATKEP 180
clon12       ACLFH SKYKSSKSSTYKKNGTSAAIQYGTGSI SGFVSQDSVQLGDLVVKEQDFIEATKEP 180
clon2        ACLFH SKYKSSKSSTYKKNGTSAAIQYGTGSI SGFVSQDSVQLGDLVVKEQDFIEATKEP 180
clon11       ACLFH SKYKSSKSSTYKKNGTSAAIQYGTGSI SGFVSQDSVQLGDLVVKEQDFIEATKEP 179
               *****:*****:*****:*****:*****:*****:*****:*****

clon8 (*)      GITFLAAKFDG-ILGLGFQEISV GKSVPVWYNMVNQGLVQEPVFSFWFN RNANEEEGGEL 239
clon5         GITFLAGKFDWYTWPLDFRRSLLGNLVPVWYNMVNQGLVQEPVFSFWFN RNANEEEGGEL 240
clon12       GITFLAAKFDG-ILGLGFQEISV GKSVPVWYNMVNQGLVQEPVFSFWFN RNANEEEGGEL 239
clon2        GITFLAAKFDG-ILGLGFQEISV GKSVPVWYNMVNQGLVQEPVFSFWFN RNANEEEGGEL 239
clon11       GITFLAAKFDG-ILGLGFQEISV GKSVPVWYNMFVK-VLSRTCVSFWLN RNANEEEGGEL 237
               *****.***      *.:. :*: *****: :.:..._**:*****

clon8 (*)      VFGGVDPNHFKGKHTYVPVTEKGYWQFNMGDVLI EDKTTGFCSDGCAAIADSGTSLLAGP 299
clon5         VFGGVDPNHFNKGKHTYVPVTEKGYWQFNMGDVLI EDKTTGFCSDGCAAIADSGTSLLAGP 300
clon12       VFGGVDPNHFNKGKHTYVPVTEKGYWQFKYDVSIEDKTTGFCSDGCAAIADSGTSLLAGP 299
clon2        VFGGVDPNHFNKGKHTYVPVTEKGYWQFNMGDVLI EDKTTGFCSDGCAAIADSGTSLLAGP 299
clon11       VFGGVDPNHFNKGKHTYVPVTEKGYWQFNMGDVLI EDKTTGFCSDGCAAIADSGTSLLAGP 297
               *****:*****:*****:*****:*****:*****:*****:*****

clon8 (*)      TAIITEINHASGAKGVMSQQCKTLVSQY GKSIIEMLLSEAQPDKIC SQMKLCTFDGARDV 359
clon5         TTIITEINHAIGAKGVMSQQCKTLVSEY GKSIIEMLLSEAQPDKIC SQMKLCTFDGARDV 360
clon12       TTIITEINHAIGAKGVMSQQCKTLVSEY GKSIIEMLLSEAQPDKIC SQMKLCTFDGARDV 359
clon2        TTIITEINHAIGAKGVMSQQCKTLVSEY GKSIIEMLLSEAQPDKIC SQMKLCTFDGARDV 359
clon11       TTIITEINHAIGAKGVMSQQCKTLVSEY GKSIIEMLLSEAQPDKIC SQMKLCTLDGARDV 357
               *:***** *****:*****:*****:*****:*****:*****

clon8 (*)      SSIIESVVDKNNKSSGGANDEMCTFC EMAVVWMQNQIKRNQTEDNI INYVNELCDRLPS 419
clon5         SSIIESVVDKNNKSSGGANDEMCTFC EMAVVWMQNQIKRNQTEDNI INYVNELCDRLPS 420
clon12       SSIIESVADKNNKSSGGANDEMCTFC EMAVVWMQNQIKRNQTEDNI INYVNELCDRLPS 419
clon2        SSIIESVVDKNNKSSGGANDEMCTFC EMAVVWMQNQIKRNQTEDNI INYVNELCDRLPS 419
clon11       SSIIESVVDKNNKSSGGANDEMCTFC EMAVVWMQNQIKRNQTEDNI INYVNELCDRLPS 417
               *****:*****:*****:*****:*****:*****:*****:*****

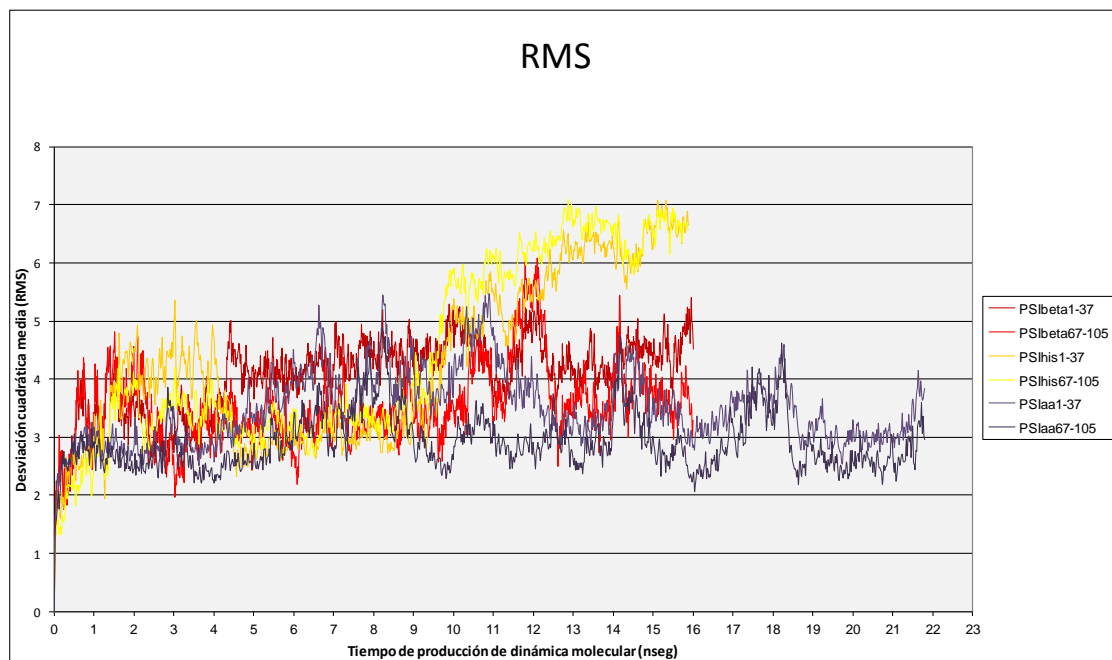
clon8 (*)      PMGESAVDCNSLSSMPNIAFTIGGKV FELCPEQYILKIGEGEAAQCISGFTAMDVAPPRG 479
clon5         PMGESAVDCNSLSSMPNIAFTIGGKV FELSP EQYILKIGEGEAAQCISGFTAMDVAPPRG 480
clon12       PMGESAVDCNSLSSMPNIAFTIGGKV FELSP EQYILKIGEGEAAQCISGFTAMDVAPPRG 479
clon2        PMGESAVDCNSLSSMPNIAFTIGGKV FELSP EQYILKIGEGEAAQCISGFTAMDVAPPRG 479
clon11       PTGESAVDCNSLSSMPNIAFTIGGKV FELSP EQYILKIGEGEAAQCISGFTAMDVAPPRG 477
               * *****:***.*****

clon8 (*)      PLWILGDVFMGRYHTVFDY GKS RVGFAEAA- 509
clon5         PLWILGDVFMGRYHTVFDY GKS RVGFAEAA- 510
clon12       PLWILGDVFMGRYHTVFDY GKS RVGFAEAA- 509
clon2        PLWILGDVFMGRYHTVFDY GKS RVGFAEAA- 509
clon11       PLWILGDVFMGRYHTVFDY GKS RVGFAEAA- 507
               *****:*****
    
```

SeqA	Name	Length	SeqB	Name	Length	Score
1	clon8_*	510	2	clon2	510	97.0
1	clon8_*	510	3	clon5	510	95.0
1	clon8_*	510	4	clon11	508	94.0
1	clon8_*	510	5	clon12	510	96.0
2	clon2	510	3	clon5	510	97.0
2	clon2	510	4	clon11	508	96.0
2	clon2	510	5	clon12	510	99.0
3	clon5	510	4	clon11	508	93.0
3	clon5	510	5	clon12	510	96.0
4	clon11	508	5	clon12	510	95.0

TABLA 22. Calificaciones de los alineamientos de a pares de secuencias

Desviación de la estructura en función del tiempo de la producción de dinámica molecular respecto a la estructura de partida minimizada



Gráfica desviación cuadrática media (RMS, root mean square) por cadenas α vs tiempo de producción de dinámica molecular. En la leyenda se denomina PSibeta a la molécula del dominio PSI del precursor de cirsina sin ningún residuo añadido en su extremo C-terminal; PSIhis, refiere al dominio PSI portando la cola de seis histidinas en su extremo C-terminal, mientras que PSIaa representa al dominio con el agregado de los residuos remanentes del sitio de reconocimiento para la enzima TEV luego del tratamiento con esta enzima (ENLYFQ). Los rangos 1-37 y 65-105, indican los números de los residuos que conforman las dos cadenas α del dominio PSI.