

**UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE QUÍMICA E BIOQUÍMICA**

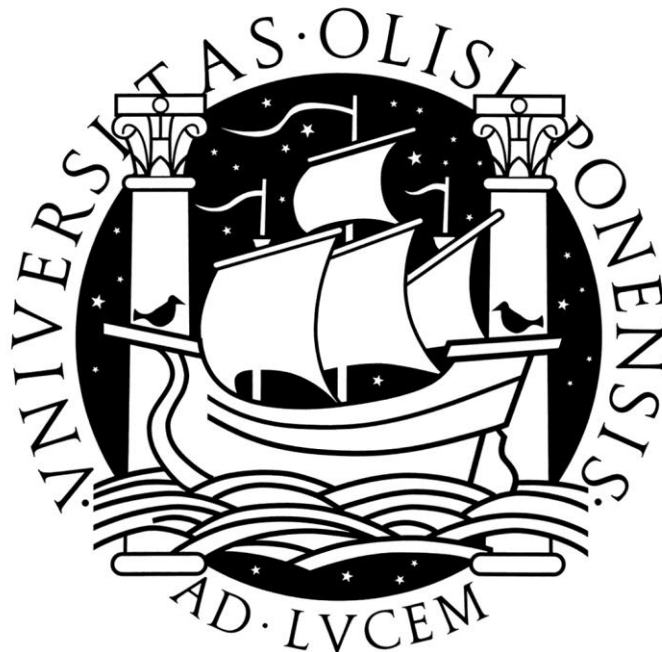


**PRODUÇÃO E CARACTERIZAÇÃO DE
QUITINASES ÁCIDAS DE MACIEIRA EM
*PICHIA PASTORIS***

SARA HOMEM SOUSA

**Mestrado em Bioquímica
(Bioquímica Médica)**

**UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE QUÍMICA E BIOQUÍMICA**



**PRODUÇÃO E CARACTERIZAÇÃO DE
QUITINASES ÁCIDAS DE MACIEIRA EM
*PICHIA PASTORIS***

SARA HOMEM SOUSA

**Mestrado em Bioquímica
(Bioquímica Médica)**

Dissertação orientada por:

Doutora Mariana da Silva Gomes Mota, INSTITUTO SUPERIOR DE AGRONOMIA

Professora Doutora Maria Luísa Mourato de Oliveira Marques Serralheiro,
FACULDADE DE CIÊNCIAS DA UNIVERSIDADE DE LISBOA

AGRADECIMENTOS

Não posso começar este trabalho sem antes agradecer às pessoas que foram importantes para a sua execução.

Em primeiro lugar, obrigado à Dra. Mariana Mota por ter apostado em mim e me ter aceite no seu projecto, pela ajuda incansável, pela total disponibilidade em apoiar-me, quer profissional quer pessoalmente. Sem a Mariana, nunca seria o mesmo.

Ao Pedro, por tudo o que significa para mim, pelo apoio, pelo companheirismo e compreensão.

A toda a minha família e amigos, principalmente aos meus pais, Paula e Eduardo, que financiaram e apoiaram a minha educação ao longo deste largos anos, e ao Filipe Santos pela ajuda preciosa no final desta etapa.

Agradeço também às minhas amigas e colegas Sandra Silva e Sofia Araújo pela ajuda, pelos tempos passados no laboratório, pelos lanches e almoços que partilhámos. Sem vocês, teria sido bem mais difícil. Ao meu colega Jorge Anjos que muita ajuda me prestou durante o tempo em que trabalhámos juntos.

Um obrigado ainda à Professora Cristina Oliveira por me ter recebido no seu laboratório e proporcionado um excelente local de trabalho e à Professora Maria Luísa Serralheiro por ter orientado a minha tese e o meu trabalho.

Este trabalho foi financiado por Fundos FEDER através do Programa Operacional Factores de Competitividade – COMPETE e por Fundos Nacionais através da FCT – Fundação para a Ciência e a Tecnologia, no âmbito do projecto PTDC/AGR-AAM/75506/2006 "Avaliação das propriedades de defesa de quitinases da macieira".

Muito obrigada a todos!

RESUMO

Este trabalho visou a clonagem e expressão de genes codificadores de quitinases ácidos de classe III de macieira no sistema heterólogo *Pichia pastoris*, nas estirpes GS115 e KM71H. Prepararam-se três construções com três genes já anteriormente identificados (C525, C209 e C203) e o sinal de secreção *a-factor* de *S. cerevisiae*, e duas construções com os genes C525 e C209 com o sinal de secreção nativo. Estas construções foram clonadas no vector pPICZaA e introduzidas na levedura. Obtiveram-se colónias positivas para todas as construções, excepto para a C209 com o sinal de secreção nativo, sugerindo uma análise de *Southern* a presença de apenas uma cópia da construção em cada recombinante. Após indução da expressão com metanol, mediu-se, em vários momentos, o crescimento e o teor total de proteína das culturas, bem como a actividade quitinolítica do secretoma das colónias recombinantes, usando glicolquitina como substrato. Seleccionaram-se as construções que apresentaram, pelo menos numa estirpe, actividade superior à dos controlos negativos correspondentes (estirpes WT e WT+pPICZaA), e os momentos que permitiram maior actividade enzimática para cada construção. Os melhores resultados foram obtidos para as isoformas C525 com o sinal de secreção nativo e C209 com o sinal de secreção *a-factor*, clonadas na estirpe GS115 (1,100U/mg de proteína às 24h e 0,233U/mg de proteína às 120h, respectivamente), e a isoforma C525 com o *a-factor*, clonada na estirpe KM71H (0,366U/mg de proteína às 139h). Destas três, a primeira foi a que expressou melhor e mais rapidamente a proteína. A análise de SDS-PAGE dos precipitados e dos secretomas respeitantes aos diferentes momentos não se revelou conclusiva, provavelmente devido à baixa concentração proteica das amostras. Propõe-se a produção em larga escala e purificação destas três proteínas recombinantes para caracterizar a sua actividade quitinolítica, para depois averiguar o seu potencial biofungicida contra fitopatogénios de macieira e pereira.

Palavras-chave

Actividade quitinolítica, expressão heteróloga, macieira, *Pichia pastoris*, quitinase.

ABSTRACT

This work focused on cloning and expressing three different genes encoding class III acidic chitinases from apple in *Pichia pastoris* heterologous expression system, using GS115 and KM71H strains. Three constructs were prepared with the genes C525, C209 e C203 and the α -factor secretion signal from *S. cerevisiae*, and two constructs with the genes C525 and C209 and the native secretion signal. The constructions were cloned in pPICZ α A vector and introduced in the yeast. We obtained positive clones for all constructs, except for construct C209 with the native secretion signal, and a Southern analysis suggested the presence of only one copy of the construct in each recombinant. After inducing the expression with methanol, cultures' growth and total protein amount, at different moments, were evaluated, and the chitinolytic activity of the cell-free's medium was determined, using glycolchitin as substrate. The constructs that showed, at least in one strain, higher activity than the corresponding negative controls (WT strain and WT+pPICZ α A) were chosen, as well as the moments with better enzymatic activity, for each construct. The best results derived from C525 gene with the native secretion signal and C209 gene with the α -factor signal, cloned in the GS115 strain (1,100U/mg protein at 24h and 0,233U/mg protein at 120h, respectively), and C525 gene with the α -factor signal cloned in the KM71H strain (0,366U/mg protein at 139h). Among these three, the first is the one that expressed the protein better and faster. A SDS-PAGE analysis of the pellets and supernatants from the different moments didn't show the expected results, probably due to the samples' low protein concentration. We propose to produce in large scale and purify these three recombinant proteins to characterize their chitinolytic activity and further analyze their potential as natural fungicide against phytopathogens that strongly affect apple and pear.

Keywords

Apple, chitinase, chitinolytic activity, heterologous expression, *Pichia pastoris*.

ÍNDICE

AGRADECIMENTOS	I
RESUMO	II
PALAVRAS-CHAVE	II
ABSTRACT	III
KEYWORDS	III
ÍNDICE	IV
ÍNDICE DE FIGURAS	VI
ÍNDICE DE QUADROS	VIII
LISTA DE ABREVIATURAS	X
INTRODUÇÃO	1
REVISÃO BIBLIOGRÁFICA	3
I. EXPRESSÃO HETERÓLOGA	4
II. A LEVEDURA <i>PICHIA PASTORIS</i>	7
i. <i>Metabolismo</i>	8
ii. <i>Um sistema de expressão</i>	10
iii. <i>Purificação – um sistema secretor</i>	15
iv. <i>Modificações pós-traducionais - glicosilação</i>	16
III. OS QUITINASES ÁCIDOS DE MACIEIRA	17
MATERIAIS E MÉTODOS	22
MATERIAL	23
i. <i>Microrganismos</i>	23
ii. <i>Reagentes</i>	23
iii. <i>Vectores</i>	24
iv. <i>Oligonucleótidos</i>	24
v. <i>Marcadores Moleculares</i>	25
vi. <i>Enzimas</i>	25
vii. <i>Meios de Cultura</i>	26
viii. <i>Kits utilizados</i>	26
ix. <i>Análise de proteínas</i>	27
MÉTODOS	27
i. <i>Biologia Molecular</i>	27
<i>Técnicas gerais de biologia molecular</i>	27
<i>Clonagem heteróloga em Escherichia coli</i>	27
<i>Clonagem de expressão em Pichia pastoris</i>	31
ii. <i>Bioquímica</i>	35
<i>Análise do enzima quitinase</i>	35
RESULTADOS E DISCUSSÃO	37
CONCLUSÃO E PERSPECTIVAS	56
BIBLIOGRAFIA	I
ANEXOS	III
ANEXO 1: COMPOSIÇÃO DOS MEIOS DE CULTURA	IV
ANEXO 2: COMPOSIÇÃO DE SOLUÇÕES E TAMPÕES	VI
ANEXO 3: PADRÃO DE SEPARAÇÃO DAS BANDAS APÓS ELECTROFORESE DOS MARCADORES UTILIZADOS	X
ANEXO 4: SEQUÊNCIAS ISOLADAS DE DNA GENÓMICO	XI
ANEXO 5: SEQUÊNCIA NUCLEOTÍDICA PARCIAL DOS VECTORES UTILIZADOS	XII
A) <i>pPICZaA</i>	<i>xii</i>

<i>B) pCRII</i>	<i>xiii</i>
ANEXO 6: SEQUÊNCIAS DE DNA DOS FRAGMENTOS INTRODUZIDOS EM <i>PICHA PASTORIS</i>	XIV
ANEXO 7: RESULTADOS OBTIDOS	XVII
<i>I. Quantificação proteica pelo método de Bradford (1976) e crescimento celular</i>	<i>xvii</i>
<i>II. Determinação da actividade quitinolítica</i>	<i>xviii</i>
<i>III. O-glicosilação</i>	<i>xix</i>

ÍNDICE DE FIGURAS

Figura 1 – Leveduras <i>Saccharomyces cerevisiae</i> (http://www.scientistlive.com) e <i>Pichia pastoris</i> (http://www.diark.org).	7
Figura 2 – A via metabólica do metanol na levedura <i>Pichia pastoris</i> . 1- álcool oxidase, 2- catalase, 3- formaldeído desidrogenase, 4- formato desidrogenase, 5- DHAS, 6- dihidroxiacetona cinase, 7- frutose 1,6-bifosfato aldolase e 8- frutose 1,6-bisfosfatofosfatase. Adaptado de (Cereghino 2000).	9
Figura 3 – Vectores para <i>Pichia pastoris</i> que contêm as características atrás descritas. Vectores comercializados pela empresa <i>Invitrogen</i> . (<i>Invitrogen</i>). ...	11
Figura 4 – Representação esquemática do processo de permuta de gene em <i>Pichia pastoris</i> . A) DNA derivado do plasmídeo pHIL-D1 digerido com BglII. B) Locus AOX da levedura. C) Permuta do locus AOX1 com o DNA de A) por recombinação homóloga. (Domínguez 1998).	14
Figura 5 – Modelação por homologia do enzima quitinase ácido da classe III de macieira, codificado pelo gene <i>Chi1</i> . Estrutura em <i>ribbon</i> . Programa SPDB Viewer.	21
Figura 6 – Levedura <i>Pichia pastoris</i> : estirpe KM71H utilizada e cultivada em laboratório.	23
Figura 7 - Representação esquemática do plasmídeo pPICZαA utilizado neste trabalho. 5'AOX1 – promotor, α-factor – sinal de secreção, c-myc – epítipo, 6xHis – cauda hexahistidinas, AOX1 TT – terminador da transcrição, Zeocin – resistência à zeocina, pUC ori – origem de replicação <i>high copy</i>	38
Figura 8 – Porção da sequência nucleotídica do plasmídeo pPICZαA correspondente ao MCS. Locais de restrição utilizados no trabalho: EcoRI (verde), XhoI (castanho) e NotI (vermelho). Local de processamento do factor-α (laranja) pelo endopeptidase Kex2 (azul).	39
Figura 9 – Representação esquemática das construções realizadas, com respectivos locais de restrição: Eco+Not e Xho+Not.	39
Figura 10 – Ligações realizadas com o plasmídeo pPICZαA para as várias construções: EcoRI+NotI ou XhoI+NotI.	40
Figura 11 – Exemplo de uma transfecção efectuada com a sequência que codifica para uma isoforma de quitinase, com o sinal de secreção nativo, inserida no vector pPICZαA. A sequência a clonar foi inserida no genoma da levedura <i>Pichia pastoris</i> após linearização da molécula recombinante. É possível a inserção de múltiplas cópias desta molécula (cassete de expressão).	41
Figura 12 - Construção C525XhoNot.	42

- Figura 13 – Gel de agarose 1,2%. Resultados das transfecções das duas estirpes de *Pichia pastoris* com as quatro construções. Amplificação por PCR com *primers* do vector (5'AOX e 3'AOX) do DNA genómico extraído. Marcador lambda DNA/EcoRI+HindIII (poços 1 e 16); KM71H e GS115 WT (2 e 3), +C525XhoNot (4 e 5), +C525EcoNot (7 e 8), +C209EcoNot (10 e 11) e +C203EcoNot (13 e 14); assim como pDNA extraído de *E. coli* transformada com as mesmas construções (controlo positivo – poços 6, 9, 12 e 15).....44
- Figura 14 – Placas MDH e MMH referentes ao estudo de confirmação do fenótipo: com as quatro recombinantes (C5252EcoNot, C5252XhoNot, C209EcoNot e C203EcoNot), com plasmídeo vazio e as duas estirpes WT de levedura.45
- Figura 15 - Resultados do *Southern blot* realizado com as construções inseridas na estirpe KM71H: WT (poço 1), +pPICZαA (poço 2), +C525XhoNot (poço 5), +C525EcoNot (poço 6), +C209EcoNot (poço 7) e +C203EcoNot (poço 8), digeridos com NheI. Marcador lambda com digoxigenina (poço 4) e pPICZαA+C525EcoNot extraído de *E. Coli*, digerido com SacI (poço 3).....47
- Figura 16 – Variação da concentração proteica ao longo do tempo para as amostras seleccionadas das quatro construções nas duas estirpes (GS115 e KM71H) e para as estirpes WT e transfectadas com vector nulo.49
- Figura 17 – Variação do crescimento celular ao longo do tempo para as amostras seleccionadas das quatro construções nas duas estirpes (GS115 e KM71H) e para as estirpes WT e transfectadas com vector nulo.49
- Figura 18 – Variação da actividade quitinolítica ao longo do tempo para as amostras seleccionadas das quatro construções na estirpe KM71H e para a estirpe WT e transfectada com vector nulo.51
- Figura 19 - A) Variação da actividade quitinolítica ao longo do tempo para as amostras seleccionadas das quatro construções na estirpe GS115 e para a estirpe WT e transfectada com vector nulo. B) Pormenor do gráfico de A) entre as 48 e as 139 horas.52
- Figura 20 – Variação da actividade enzimática ao longo do tempo para as construções seleccionadas das duas estirpes GS115 e KM71H.55
- Figura 21 - Output do software *NetOGlyc 3.1 Server* que detecta os resíduos de Ser e Thr passíveis de serem O-glicosilados para as proteínas recombinantes A) C203EcoNot e B) C209EcoNot. xix
- Figura 22 - Output do software *NetOGlyc 3.1 Server* que detecta os resíduos de Ser e Thr passíveis de serem O-glicosilados para as proteínas recombinantes A) C525EcoNot e B) C525XhoNot..... xx

ÍNDICE DE QUADROS

Quadro 1 - Comparação de factores importantes na expressão heteróloga entre sistemas de expressão: bactéria, levedura, insecto e mamífero. Nas figuras, <i>E. coli</i> , <i>Pichia pastoris</i> , <i>Drosophila</i> e <i>Hela-cells</i> . Adaptado de www.genwaybio.com . PTMs – Modificações Pós-tradução.	6
Quadro 2 – Características de alguns vectores disponíveis para transfectar <i>Pichia pastoris</i> (Invitrogen).	11
Quadro 3 - Características de algumas estirpes de <i>Pichia pastoris</i> desenvolvidas pela empresa <i>Invitrogen</i> . (<i>Invitrogen</i> , <i>Pichia pastoris</i> Expression System)	13
Quadro 4 – Características dos microrganismos utilizados: organismos, estirpes, genótipos e fenótipos.	23
Quadro 5 – Características dos vectores plasmídicos utilizados.	24
Quadro 6 – <i>Primers</i> utilizados para amplificar e detectar sequências específicas com respectivas sequências e temperaturas de desnaturação (T _m).	25
Quadro 7 – Enzimas de restrição utilizados, <i>NotI</i> , <i>EcoRI</i> , <i>XhoI</i> , <i>SacI</i> e <i>NheI</i> , respectivo local de restrição, tampão e temperatura de incubação (T _i). Todos estes enzimas foram adquiridos à <i>Fermentas</i> ®.	26
Quadro 8 – Resultados obtidos na clonagem das cinco construções (C525EcoNot, C525XhoNot, C209EcoNot, C209XhoNot e C203EcoNot) nas duas estirpes de <i>Pichia pastoris</i> , assim como transfectadas com o vector nulo.	42
Quadro 9 – Tamanhos dos fragmentos amplificados pelos <i>primers</i> específicos do MCS do vector pPICZαA, para as quatro construções.	43
Quadro 10 – Resultados obtidos para o <i>screening</i> das colónias obtidas, tendo como controlo positivo o pDNA extraído de <i>E. coli</i> com as construções.	44
Quadro 11 – Resultados obtidos nos ensaios de confirmação de fenótipo e de viabilidade dos <i>stocks</i> de glicerol.	45
Quadro 12 – Valores de concentração de proteína em mg/mL das amostras seleccionadas para as quatro construções nas duas estirpes (GS115 e KM71H) e para as estirpes WT e transfectadas com vector nulo.	48
Quadro 13 – Valores de actividade enzimática em U e U/mg de proteína total das amostras seleccionadas para as quatro construções nas duas estirpes (GS115 e	

KM71H) e para as estirpes WT e transfectadas com vector nulo. A negrito evidenciam-se os valores mais elevados para cada isoforma.	50
Quadro 14 – Valores de actividade quitinolítica em U/mg de proteína total para as amostras que apresentam maior actividade e controlos respectivos.	52
Quadro 15 – Valores de U/mL de sobrenadante e mg de GlnNAc/mL de sobrenadante para as amostras que apresentam maior actividade e controlos respectivos.	53
Quadro 16 – Composição e concentrações das soluções utilizadas ao longo do trabalho.	vi
Quadro 17 – Valores de absorvância a 595nm e 600nm das amostras em estudo para as quatro construções nas duas estirpes (GS115 e KM71H) e para as estirpes WT e transfectadas com vector nulo. Concentração proteica total para as mesmas amostras, em mg/mL.	xvii
Quadro 18 – Média das três absorvâncias medidas a 420nm das amostras seleccionadas para as quatro construções nas duas estirpes (GS115 e KM71H) e para as estirpes WT e transfectadas com vector nulo.....	xviii

LISTA DE ABREVIATURAS

(p/v)	peso/volume
°C	Graus Celsius
3'	Região a juzante do gene
5'	Região a montante do gene
5'UTR	Região 5' não traduzida
A	Adenina
ADE1	Gene que codifica o N-succinil-5-aminoimidazol-4-carboxamida ribonucleótido sintase
Ala	Alanina
AOX	Álcool oxidase
AOX1	Gene que codifica o álcool oxidase I
AOX2	Gene que codifica o álcool oxidase II
Arg	Arginina
ARG4	Gene que codifica o argininosuccinato liase
arg4	Argininosuccinato liase
Asn	Asparagina
ATG	Codão de iniciação
BMGY	Meio tamponado complexo com Glicerol
BMMY	Meio tamponado complexo com Metanol
bp	Pares de base
C	Citosina
Chi1	Gene que codifica o quitinase
c-myc	Epítipo c-myc
CO₂	Dióxido de carbono
DH	<i>Denhardt solution</i>
DHAS	Dihidroxiacetona sintase
DMSO	Dimetilsulfóxido
DNA	<i>DeoxyriboNucleic Acid</i>
dNTP	Desoxiribonucleótidos trifosfatados
DTT	Ditiotreitol
E. coli	<i>Escherichia coli</i>
EC	<i>Enzyme Commission</i>
EDTA	Ácido Etilenodiamino tetra-acético
FAD	Dinucleótido de flavina e adenina
FDA	<i>American Food and Drug Administration</i>
G	Guanina
GAP	Gliceraldeído-3-fosfato desidrogenase
Glc	Glucosamina
Glu	Glutamina
GRAS	<i>Generally Regarded As Safe</i>
HB1	Tampão de Hibridação 1
HIS4	Gene que codifica o histidinol desidrogenase
his4	Histidinol desidrogenase
His₆	Cauda de hexahistidinas
IUBMB	<i>International Union of Biochemistry and Molecular Biology</i>
kDa	Quilo-Dalton
Kex2	Endopeptidase <i>Kexin</i>
K_i	Constante de inibição enzimática
Km	Constante de <i>Henri-Michaelis-Menton</i>
Lys	Lisina
Man	Manose
MCS	<i>Molecular cloning site</i>
MD	Meio Mínimo com Dextrose
MDH	Meio Mínimo com Dextrose e Histidina
MM	Meio Mínimo com Metanol
MMH	Meio Mínimo com Metanol e Histidina
mRNA	<i>Ribonucleic acid messenger</i>
Mut⁻	<i>Methanol Utilization Minus</i>
Mut⁺	<i>Methanol Utilization Plus</i>

Mut^S	<i>Methanol Utilization Slow</i>
NAC	N-acetil
nm	Nanómetros
OD	<i>Optical density</i>
ORF	<i>Open reading frame</i>
P. pastoris	<i>Pichia pastoris</i>
PCR	<i>Polymerase Chain Reaction</i>
PDB	<i>Protein Data Bank</i>
pDNA	DNA plasmídico
PEG 3350	Polietilenoglicol 3350
pep4	Proteinase A
PEP4	Gene que codifica o proteinase A
pH	Potencial do hidrogenio iónico
PHO1	Gene que codifica a fosfatase acídica
pI	Ponto isoelectrico
PMSF	Fluoreto de Fenilmetilsulfonil
pp	página
PR	<i>Pathogenesis-related</i>
prb1	Proteinase B
PRB1	Gene que codifica o proteinase B
PSA	Ácido Fenolsulfónico
PTMs	<i>Post-translational modifications</i>
RNA	<i>Ribonucleic acid</i>
RNA	<i>RiboNucleic Acid</i>
rpm	Rotações por minuto
S. cerevisiae	<i>Saccharomyces cerevisiae</i>
SDS	Dodecil Sulfato de Sódio
Ser	Serina
SOB	<i>Super optimal broth</i>
SOC	<i>Super optimal broth catabolite repression</i>
SPDB Viewer	<i>Swiss-PDB viewer</i>
SSC	Tampão Citrato de Sódio Salino
Ste13	Dipeptidil aminopeptidase
T	Timina
TAE	Tampão Tris-Acetato-EDTA
Tan	Temperatura de <i>annealing</i>
Taq	DNA polimerase da bactéria <i>Thermus aquaticus</i>
TE	Tampão Tris-EDTA
Thr	Treonina
Tm	Temperatura de emparelhamento
U	Unidades enzimáticas
URA3	Gene que codifica o orotidina 5-fosfato descarboxilase
ura3	Orotidina 5-fosfato descarboxilase
UV	Ultravioleta
WT	<i>Wild type</i>
YNB	<i>Yeast Nitrogen Base</i>
YPD	<i>Yeast Extract Peptone Dextrose Medium</i>
YPDS	<i>Yeast Extract Peptone Dextrose Medium with Sorbitol</i>
Zeo	Zeocina

INTRODUÇÃO

A diversidade do solo português e as condições climáticas do país permitem o grande desenvolvimento da cultura da macieira em Portugal. De facto, a produção de maçã representa uma parte significativa da produção total de fruta em território nacional.

O grande desafio na cultura da macieira prende-se com os fitopatogénios que afectam a qualidade do fruto e o rendimento da sua produção. Tal pode apenas ser controlado pelo uso repetido de pesticidas, compostos que originam problemas ecológicos ou mesmo na saúde do consumidor. Assim, urge a procura de fungicidas que sejam menos prejudiciais ao ambiente e à saúde humana, como biopesticidas derivados de plantas.

Neste seguimento, proteínas que tenham uma função defensiva na macieira podem ser indicadas para funcionar como fungicidas naturais. Poderá ser o caso dos quitinases ácidos de classe III da macieira. Dois genes que codificam estas enzimas mostraram ter sequências semelhantes às de quitinases isoladas de *Rhizopus oligosporus*, tendo já sido demonstrada actividade antifúngica nestas últimas.

No âmbito de um projecto visando a avaliação do potencial antifúngico dos quitinases ácidos de classe III de macieira, procurou-se, neste trabalho, clonar num vector apropriado as sequências do DNA genómico da macieira 'Royal Gala' que codificam estas enzimas e expressar as proteínas correspondentes em *Pichia pastoris*, caracterizando-as quanto à sua actividade quitinolítica.

REVISÃO BIBLIOGRÁFICA

I. Expressão Heteróloga

O desenvolvimento de proteínas funcionais em larga escala é um campo de extensa investigação na área da Biotecnologia. De facto, estas podem actuar como agentes terapêuticos, constituir uma ferramenta na investigação de problemas científicos ou ainda ser usadas na indústria. Para estes fins é necessário, então, produzir quantidades suficientes de uma proteína pura.

A produção *in vivo* de uma proteína é um processo muito complexo e regulado que envolve muitos enzimas e co-factores, modificações pós-traducionais da proteína e a secreção da mesma, se for o caso (Dessai 2010). Todos estes factores são necessários para a estabilidade e função da proteína e são difíceis de reproduzir em laboratório.

Foi a necessidade de criar métodos de expressão reprodutíveis, aliada ao potencial biotecnológico de microrganismos unicelulares e de células eucariotas, que abriu portas à produção heteróloga de proteínas. A expressão heteróloga não é mais do que a produção de uma proteína fora do seu hospedeiro original, e com os avanços genéticos na área da biologia molecular, foi possível ultrapassar a barreira da espécie e produzir proteínas de interesse, na forma de proteínas recombinantes (Dessai 2010).

Assim, um sistema de expressão válido deve ser capaz de reproduzir a proteína desejada com a conformação correcta, ser fácil de manipular e de manter, ser seguro e economicamente viável e ter uma boa produtividade (Dessai 2010).

As bactérias, como organismos procariotas simples e fáceis de manipular, tornaram-se a opção mais viável para a expressão heteróloga de proteínas. A bactéria *Escherichia coli* foi o sistema de expressão pioneiro e é o organismo mais utilizado na expressão de proteínas, dado o vasto conhecimento da sua genética e fisiologia, e detém o maior número de sucessos (Rai 2001; Dessai 2010). Apesar dos grandes avanços na tecnologia do DNA recombinante, a expressão génica é ainda um processo empírico, pois cada mRNA e cada proteína codificada apresenta os seus desafios (Balbás 2001). Não há um sistema de expressão universal para proteínas heterólogas e todos os disponíveis têm vantagens e desvantagens, sendo então necessário avaliar questões como o rendimento, a glicosilação, o *fold*ing ou a viabilidade económica antes de escolher um sistema para expressar uma proteína.

Por exemplo, a utilização de sistemas de expressão Procariotas é limitada quando se quer produzir proteínas eucariotas, apesar de serem o hospedeiro mais económico e mais fácil de manter e manipular. Esta limitação prende-se com o

facto de os organismos procariotas serem incapazes de realizar modificações pós-traducionais para o *fold*ing correcto, como a glicosilação, a fosforilação ou a formação de pontes persulfureto, entre outras, que os organismos Eucariotas realizam (Rai 2001). O problema centra-se principalmente na ausência de O- e N-glicosilação e, apesar de já se terem desenvolvido bactérias capazes de realizar O-glicosilação, o trissacárido adicionado à proteína para o *fold*ing correcto não é igual ao dos eucariotas (Rai 2001).

Outros factores como a precipitação de proteínas em corpos de inclusão, a diferente frequência de utilização de codão, a possível existência de pirogénios tóxicos na parede celular, como endotoxinas (Domínguez 1998), e limites em mecanismos de secreção extracelular (Dessai 2010) fazem das bactérias uma opção desapropriada para a expressão de proteínas de organismos superiores.

Com o passar dos anos, outros sistemas de expressão foram sendo desenvolvidos, como células de mamífero, leveduras, insectos, plantas transgénicas e animais transgénicos (Dessai 2010). Mais uma vez, as vantagens e desvantagens de cada um destes sistemas variam consoante a proteína a sintetizar.





Os organismos Eucariotas inferiores, como as leveduras, são capazes de realizar modificações pós-traducionais em proteínas recombinantes mas têm custos mais elevados que as bactérias e fases de crescimento mais lentas, apesar de serem mais económicos do que células de mamífero, por exemplo (Dessai 2010). O facto de as células de mamífero poderem conter DNA oncogénico ou viral é um factor que desaconselha a expressão de proteínas heterólogas neste sistema (Domínguez 1998).

As leveduras são a alternativa mais favorável às bactérias para a expressão de proteínas Eucariotas. Têm fases de crescimento mais rápidas do que outros sistemas Eucariotas, em meio simples e para densidades elevadas, e a sua genética é melhor conhecida do que a de qualquer outro organismo Eucariota, o que as torna fáceis de manipular. São também os organismos mais indicados para a expressão de proteínas secretadas, o que facilita muito o processo de purificação da proteína recombinante, como será discutido posteriormente (Rai 2001; Dessai 2010).

No entanto, e apesar de estes organismos realizarem glicosilação, foi demonstrado que os oligossacáridos da N- e da O-glicosilação não são idênticos aos dos mamíferos, dificultando a expressão de certas proteínas de Eucariotas superiores em levedura (Domínguez 1998; Rai 2001).

O Quadro 1 compara os aspectos mais importantes a ter em conta na expressão heteróloga de proteínas, para vários sistemas de expressão.

Quadro 1 - Comparação de factores importantes na expressão heteróloga entre sistemas de expressão: bactéria, levedura, insecto e mamífero. Nas figuras, *E. coli*, *Pichia pastoris*, *Drosophila* e *Hela-cells*. Adaptado de www.genwaybio.com. PTMs – Modificações Pós-tradução.

Características	 <i>E. coli</i>	 Leveduras	 Insectos	 Mamíferos
<ul style="list-style-type: none"> •Crescimento celular •Complexidade do meio de crescimento •Custos de crescimento •Nível de expressão •Expressão extracelular •<i>Folding</i> proteico •N-Glicosilação •O-Glicosilação •Fosforilação •Acetilação e Acilação •Rendimento (mg/L cultura) •% Sucesso •Custo do projecto •Uso recomendado •Vantagens •Desvantagens 	<ul style="list-style-type: none"> •Rápido •Mínimo •Baixo •Elevado •Secreção para o periplasma •Necessário <i>refold</i> muitas vezes •Não •Não •Não •Não •50-500 •40-60 •Baixo •Antigénios, <i>standards</i> proteicos, proteínas funcionais •Simples, robusto, custo baixo, rendimento alto •Poucas PTMs 	<ul style="list-style-type: none"> •Rápido •Mínimo •Baixo •Baixo-Elevado •Secreção para o meio •<i>Refold</i> pode ser necessário •Elevada manosiilação •Sim •Sim •Sim •10-200 •50-70 •Baixo •Proteínas glicosiladas, vacinas, formas secretadas •Simples, baixo custo, bom para algumas proteínas •Mais longo, algumas PTMs 	<ul style="list-style-type: none"> •Lento •Complexo •Elevado •Baixo-Elevado •Secreção para o meio •<i>Folding</i> correcto •Simples, sem ácido siálico •Sim •Sim •Sim •10-200 •50-70 •Médio •Proteínas glicosiladas, secretadas, alternativa às leveduras •Rendimento elevado, melhores PTMs •Mais longo, custo elevado 	<ul style="list-style-type: none"> •Lento •Complexo •Elevado •Baixo-Elevado •Secreção para o meio •<i>Folding</i> correcto •Complexo •Sim •Sim •Sim •0,1-100 •80-95 •Elevado •Estudos funcionais, estudos PTMs, caracterização •Configuração proteica, óptimas PTMs •Custo elevado, menor rendimento

A expressão heteróloga em leveduras revela-se assim a mais indicada para expressar proteínas puras de organismos superiores, em grandes quantidades, pois conjuga as vantagens dos sistemas Procarióticos (fácil manipulação, crescimento rápido, genética bem caracterizada) com as vantagens dos sistemas eucarióticos (capacidade de realizar processamento e modificações pós-tradução), criando um bom compromisso entre as bactérias e as células de Eucariotas superiores.

De todas as leveduras utilizadas para expressão heteróloga, são de salientar as leveduras *Saccharomyces cerevisiae* e *Pichia pastoris* (Figura 1). Destas, a primeira tem a desvantagem de hipermanosilar algumas proteínas secretadas (Higgins 1998), factor ultrapassável pela utilização da *Pichia pastoris*, como será detalhado mais à frente. Uma vantagem desta levedura, uma levedura metilotrófica, é a extensa informação relativa à sua segurança, nomeadamente o seu *status* GRAS (*Generally Regarded As Safe* – FDA) (Domínguez 1998).

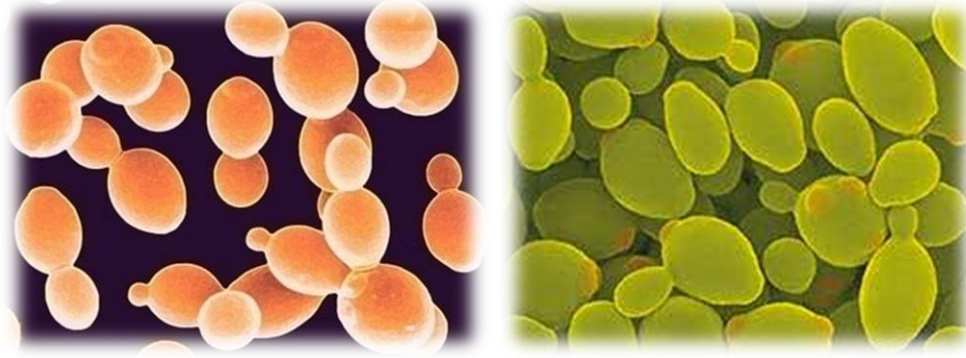


Figura 1 – Leveduras *Saccharomyces cerevisiae* (<http://www.scientistlive.com>) e *Pichia pastoris* (<http://www.diark.org>).

Outras limitações da levedura mais utilizada, a *S. cerevisiae*, são o baixo rendimento e a secreção ineficiente, já que pode ocorrer retenção da proteína heteróloga sintetizada no periplasma ou ainda integração desta na membrana plasmática da célula (Domínguez 1998).

II. A Levedura *Pichia pastoris*

Há cerca de 40 anos, Ogata percebeu a capacidade de algumas espécies de leveduras, incluindo a *Pichia pastoris*, em utilizar metanol como única fonte de carbono e energia. Desde então, muitas empresas (*Phillips Petroleum Company* ou *Salk Institute Biotechnology*) têm desenvolvido métodos e meios de cultura para otimizar a cultura contínua em larga escala da *P. pastoris*. O facto de, em 1993, a *Phillips Petroleum Company* ter decidido colocar à disponibilidade de todos os laboratórios de investigação académica o sistema de expressão que desenvolveu, permitiu o aumento significativo do conhecimento sobre esta levedura. Passou então a desenvolver-se esta levedura como sistema de expressão heteróloga de genes, conseguindo isolar-se o gene *AOX1* e o seu promotor. Desenvolveram-se ainda vectores e estirpes da levedura assim como métodos de biologia molecular mais indicados para a *Pichia* (Cereghino 2000).

Para além de ser considerada um sistema modelo para estudos celulares, a levedura *Pichia pastoris* tornou-se um sistema de expressão otimizado, bem sucedido na expressão de centenas de proteínas heterólogas (Higgins 1998). Vários são os factores que levaram à aceitação deste Eucariota como sistema de expressão e podem salientar-se:

- ✎ Como levedura, a facilidade de manipulação, já que os métodos aplicados no seu manuseamento são em tudo semelhantes aos da levedura *Saccharomyces cerevisiae*, melhor caracterizada (Higgins 1998; Cereghino 2000), mais rápidos, mais acessíveis e menos dispendiosos do que os utilizados nos sistemas de expressão de Eucariotas superiores;

- ✂ Como levedura metilotrófica, a preferência pelo crescimento respiratório com metanol, em oposição à fermentação. A fermentação de leveduras como *S. cerevisiae* forma etanol que intoxica o meio, impedindo o crescimento e a produção de proteínas heterólogas. No caso das leveduras metilotróficas, como *P. pastoris*, obtêm-se níveis de expressão proteica cerca de 10 a 100 vezes mais elevados do que com *S. cerevisiae* (Higgins 1998; Cereghino 2000);
- ✂ O desenvolvimento de um promotor forte derivado do gene do álcool oxidase I (*AOX1*) da levedura, uma ferramenta de regulação da expressão de genes heterólogos (Higgins 1998; Cereghino 2000);
- ✂ A baixa secreção de proteínas endógenas, que permite que as proteínas heterólogas secretadas constituam a maior parte da fracção proteica do meio, facilitando a sua purificação (Higgins 1998; Cereghino 2000);
- ✂ Como eucariota, a capacidade de realizar modificações pós-tradução associadas a Eucariotas superiores, como a proteólise, a formação de pontes persulfureto, a adição de lípidos e a N- e O-glicosilação, com um padrão de glicosilação equiparável ao dos Eucariotas superiores (Domínguez 1998; *cit. in* Arruda 2008), de forma a expressar proteínas enroladas correctamente e activas (Higgins 1998; Cereghino 2000).

Assim, um sistema de expressão baseado neste microrganismo unicelular, fácil de manipular e de cultivar, rápido e menos dispendioso, é indicado para a expressão heteróloga de proteínas Eucariotas e, geralmente, obtêm-se níveis de expressão destas mais elevados, quando comparados com *E. coli* (Higgins 1998). A seu favor tem, ainda, a semelhança das técnicas de manipulação genética que requer com as da levedura *S. cerevisiae* (Cereghino 2000).

Actualmente, esta levedura é utilizada não só como sistema de expressão mas também como modelo para estudos da biogénese de peroxissomas, análise genética de processos de autofagia ou mesmo para estudos de organização entre o Retículo Endoplasmático e o aparelho de Golgi (*cit. in* Arruda 2008).

i. Metabolismo

Como levedura metilotrófica, a *Pichia pastoris* é capaz de metabolizar metanol. Um facto curioso, e que permitiu grandemente a aceitação desta levedura como sistema de expressão heterólogo, é a presença, em concentrações elevadas, dos enzimas necessários ao metabolismo do metanol apenas quando as células crescem em meio metanólico (Cereghino 2000).

A via metabólica do metanol desta levedura (Figura 2) envolve um conjunto de enzimas únicos e específicos. O primeiro passo da via é a oxidação do metanol a formaldeído, com produção de peróxido de hidrogénio, pelo enzima álcool oxidase (AOX). O enzima tem pouca afinidade para o oxigénio e, necessitando deste para a oxidação, a célula compensa este facto com a síntese de grandes quantidades de AOX (Higgins 1998; Cereghino 2000). Uma parte do formaldeído formado é oxidada por dois desidrogenases citosólicos, com formação de formato e CO_2 . É desta oxidação que a levedura retira parte da sua energia para o crescimento. A restante parte do formaldeído produzido é utilizada numa via de produção de gliceraldeído 3-fosfato, catalisada pelo enzima dihidroxiacetona sintase (DHAS). Esta via é um ciclo, como mostra a figura, produzindo uma molécula de gliceraldeído 3-fosfato a cada três ciclos. Este servirá à célula para sintetizar constituintes celulares (Cereghino 2000).

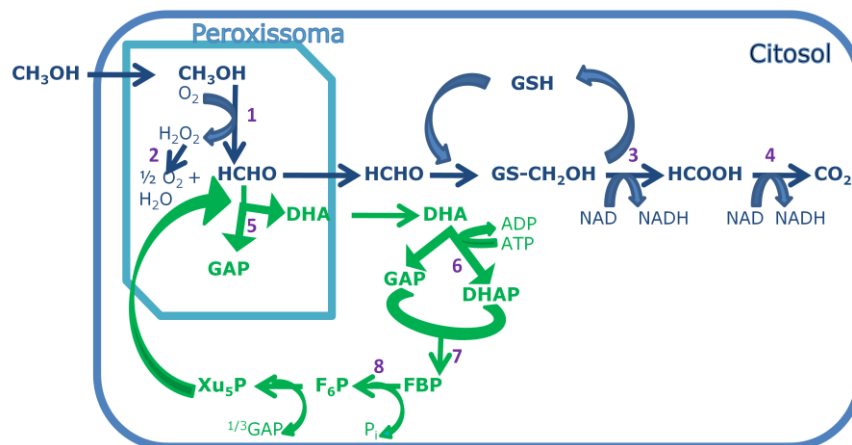


Figura 2 – A via metabólica do metanol na levedura *Pichia pastoris*. 1- álcool oxidase, 2- catalase, 3- formaldeído desidrogenase, 4- formato desidrogenase, 5- DHAS, 6- dihidroxiacetona cinase, 7- frutose 1,6-bisfosfato aldolase e 8- frutose 1,6-bisfosfatofosfatase. Adaptado de (Cereghino 2000).

São estes dois enzimas, o AOX e o DHAS, que estão presentes em quantidades vestigiais quando a célula cresce num meio com outra fonte de carbono que não o metanol (glucose, glicerol ou etanol), e que se encontram altamente expressos quando a célula cresce em meio metanólico. O enzima AOX chega a compreender 30% das proteínas totais solúveis de células que crescem em meio metanólico (Cereghino 2000).

Nesta levedura, o enzima AOX é codificado por dois genes, o *AOX1* e o *AOX2*, apesar de o *AOX1* expressar a maior parte do enzima activo da célula. Este gene é regulado e induzido pelo metanol e, na célula em meio metanólico, o nível de mRNA poliadenilado do gene *AOX1* é aproximadamente 5%, enquanto que noutra fonte de carbono é quase indetectável (Higgins 1998). A sua regulação dá-se ao nível da transcrição, através de um processo que envolve dois mecanismos: um mecanismo de repressão/não-repressão e outro de indução. O crescimento em

glucose reprime a transcrição mesmo quando é adicionado metanol, enquanto que o crescimento em glicerol permite a indução com metanol. O crescimento em glicerol apenas (condições de não-repressão) não resulta na transcrição elevada do gene *AOX1*, sendo a presença de metanol essencial para induzir níveis elevados de transcrição do gene (Cereghino 2000).

ii. Um sistema de expressão

A produção heteróloga de proteínas em levedura é um processo trabalhoso e envolve vários passos. São três os passos fundamentais: inserir o gene a clonar num vector de expressão, introduzir o vector recombinante na levedura e examinar o produto genómico heterólogo em várias estirpes.

Inicialmente, é necessário identificar e isolar o DNA que codifica a proteína de interesse. Clonar em *E. coli* e verificar, por sequenciação, a existência do fragmento correcto pode mostrar-se importante para evitar erros em fases avançadas.

Após a obtenção do DNA a clonar, deve seleccionar-se um vector apropriado para a clonagem. Os vectores plasmídicos desenhados para expressão heteróloga em *Pichia pastoris* (Figura 3, Quadro 2) têm várias características em comum:

- ✘ O promotor *AOX1* da *Pichia*, altamente regulado por um mecanismo dependente de metanol, permitindo a regulação da expressão do gene de interesse;
- ✘ Uma sequência de terminação da transcrição do gene *AOX1* da *Pichia*, para certificar que o RNA polimerase desliga o DNA e não continua a transcrever genes a jusante direccionando o processamento em 3' e a poliadenilação do mRNA;
- ✘ Um *polylinker* com múltiplos locais de restrição para clonagem, que permite introduzir mais facilmente genes heterólogos na orientação correcta (Nicholl 2002), entre o promotor e o terminador da transcrição (Cereghino 2000);
- ✘ A sequência do factor- α de *S. cerevisiae* para secreção de proteínas recombinantes;
- ✘ Marcadores de selecção que permitam o crescimento apenas de células portadoras da molécula recombinante.

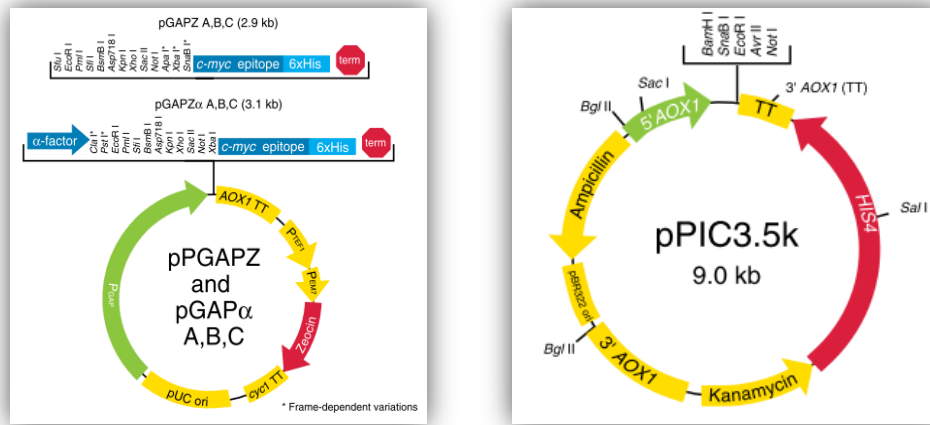


Figura 3 – Vetores para *Pichia pastoris* que contêm as características atrás descritas. Vetores comercializados pela empresa Invitrogen. (Invitrogen).

De entre estas características, é de salientar o promotor AOX1. Este é um fragmento do sistema genético AOX1 da *Pichia*. Com 900 pares de base, contém a sequência 5' do promotor, um fragmento do gene AOX1 mais curto e as sequências necessárias à terminação da transcrição (Cereghino 2000). Como referido anteriormente, este promotor é regulado por um mecanismo de repressão/indução dependente de metanol (Domínguez 1998). Quando a sequência a expressar é inserida com o seu ATG (codão de iniciação) o mais perto possível do ATG do AOX1, obtêm-se melhores resultados de clonagem de expressão. Muitas vezes, esta posição coincide com o primeiro local de restrição do MCS (Cereghino 2000).

Quadro 2 – Características de alguns vectores disponíveis para transfectar *Pichia pastoris* (Invitrogen).

<p>PGAPZ, PGAPZα A, B ou C - para expressão sem metanol</p> <ul style="list-style-type: none"> • Promotor GAP, resistência à Zeocina e epítopo c-myc e polihistidina, pode ter factor-α para secreção
<p>PIC6, PIC6α A, B ou C - para expressão com selecção pela blasticidina</p> <ul style="list-style-type: none"> • Promotor AOX1, resistência à blasticidina, epítopo c-myc e polihistidina e gene 5' AOX1 para integração no genoma, pode ter factor-α para secreção
<p>PIC3.5K, PIC9K e AO815 - para aumentar o número de cópias do gene a expressar</p> <ul style="list-style-type: none"> • Promotor AOX1, gene HIS4 para identificação dos transformantes, gene 3'AOX1 para integração no genoma, resistência à canamicina (PIC3.5K e PIC9K), resistência à geneticina para identificar locais de inserção múltipla (AO815), pode ter factor-α para secreção (PIC9K)
<p>PICZ, PICZα A, B ou C - selecção com Zeocina, utilização em <i>E. coli</i></p> <ul style="list-style-type: none"> • Promotor AOX1, epítopo c-myc e polihistidina e gene 5'AOX1 para integração no genoma, pode ter factor-α para secreção

Algumas construções possibilitam, ainda, a fusão *in frame* da sequência que codifica a proteína com um sinal para secreção da proteína expressa. A sequência que codifica para este sinal encontra-se logo após o promotor AOX1 e a mais frequentemente utilizada é a do factor-α de *Saccharomyces cerevisiae*, embora

também se utilize uma sequência de sinal derivada do gene da fosfatase acídica da *Pichia* (*PHO1*) ou mesmo sequências de sinal nativas da proteína heteróloga (Higgins 1998; Cereghino 2000), como será mencionado no ponto iii.

Alguns vectores permitem ainda a complementação de outros plasmídeos já que contêm sequências 3' flanqueantes do AOX1 capazes de recombinar (Higgins 1998). Sequências que codificam os *tags* His₆ ou c-myc também podem fazer parte dos vectores, de modo a simplificar a purificação das proteínas expressas (Higgins 1998). Os vectores pPICZa e pPICZ são vectores *shuttle* que apresentam ainda uma origem de replicação e marcas de selecção para expressão em *E. coli*.

Como marcadores de selecção, foram implementados nestes vectores genes de resistência a antibióticos: resistência à zeocina pelo gene *Sh ble* de *Streptoalloteichus hindustanus*; ou marcas auxotróficas: HIS4, ADE1, ARG4 e URA3 de *P. pastoris* ou ARG4 e HIS4 de *S. cerevisiae* (Cereghino 2000). A zeocina é um antibiótico semelhante à bleomicina que actua na dupla-hélice da cadeia de DNA cromossomal. A proteína codificada pelo gene *Sh ble* liga-se ao antibiótico inactivando-o (Higgins 1998).

Foram também desenvolvidos vectores com promotores alternativos que não necessitam da adição do metanol ao meio de cultura. Estes permitem expressar genes cujos produtos não sejam tóxicos e permitem uma activação da expressão que não por metanol (Higgins 1998; Cereghino 2000). Um desses casos é o promotor derivado do gene do gliceraldeído-3-fosfato desidrogenase de *Pichia*, o promotor GAP, que proporciona uma expressão constitutiva forte, em meio rico em glucose, a um nível comparável com o do promotor AOX1. Mas, sendo um promotor constitutivo, não deve ser utilizado para a produção de proteínas tóxicas para a levedura (Cereghino 2000).

Após a construção do vector recombinante, que contém a sequência da proteína a ser expressa, é necessário escolher um hospedeiro e a estirpe a usar. A escolha da estirpe é um processo empírico que deve ser realizado experimentalmente de forma a poder testar-se o efeito de vários *backgrounds* genéticos na síntese da proteína heteróloga (Domínguez 1998).

Todas as estirpes de *P. pastoris* disponíveis derivam da estirpe selvagem NRRL-Y 11430 (Northern Regional Research Laboratories). Podem ter mutações em genes auxotróficos para selecção (como no gene do histidinol desidrogenase, *his4*) mas podem também ter mutações nos genes AOX, o que varia a capacidade de utilizar o metanol. Estas estirpes crescem em meio completo mas necessitam de suplementos nutricionais para crescer em meio mínimo (Higgins 1998).

Por vezes, estirpes com os genes *AOX* deletados são melhores para expressão heteróloga de proteínas do que estirpes WT, pois requerem menos metanol para induzir a expressão proteica, que fica assim mais lenta. Daí a classificação das estirpes pelo fenótipo Mut:

- ✂ Mut⁺ - fenótipo utilização de metanol +,
- ✂ Mut⁻ - fenótipo utilização de metanol -,
- ✂ Mut^s - fenótipo utilização de metanol lento.

As estirpes Mut⁺ crescem em metanol à velocidade da estirpe WT, as Mut⁻ não crescem em metanol e as Mut^s crescem em metanol a uma velocidade lenta (Higgins 1998). Tal deve-se à deleção de um (Mut^s) ou de ambos (Mut⁻) os genes *AOX*, o que permite um crescimento sem a necessidade de altas quantidades de metanol, como na estirpe Mut⁺. Todas estas estirpes retêm a capacidade de induzir a expressão elevada pelo promotor *AOX1*, aquando da sua reintrodução (Cereghino 2000).

Quadro 3 - Características de algumas estirpes de *Pichia pastoris* desenvolvidas pela empresa Invitrogen. (Invitrogen, *Pichia pastoris* Expression System)

GS115 (<i>his4</i>)	•Seleção de vectores de expressão com <i>HIS4</i>
X-33 WT	•Seleção de vectores resistentes à Zeocina
KM71 (<i>his4 arg4 aox1Δ::SARG4</i>)	•Deleção parcial do gene <i>AOX1</i> e adição do gene <i>ARG4</i> de <i>S. cerevisiae</i> Seleção de vectores com <i>HIS4</i> - estirpes com fenótipo Mut ^s
KM71H (<i>arg4 aox1Δ::SARG4</i>)	•Seleção de vectores resistentes à Zeocina - estirpes com fenótipo Mut ^s
SMD1168 (<i>his4 ura3 Δpep4::URA3</i>)	•Seleção de vectores com <i>HIS4</i> - estirpes sem actividade de protease A
SMD1168H (<i>pep4</i>)	•Seleção de vectores resistentes à Zeocina - estirpes sem actividade de protease A
MC100-3 (<i>his4 arg4 aox1Δ::SARG4 aox2Δ::Phis4</i>)	•Deleção de ambos os genes <i>AOX</i> - estirpes com fenótipo Mut ⁻

A existência de proteases em meio de cultura pode dificultar a estabilidade da proteína final. Este problema de degradação no meio de cultura pode ser ultrapassado pela expressão num organelo ou órgão específico, pela co-expressão com inibidores de proteases ou com co-factores e subunidades (Balbás 2001), mas também pelo do uso de estirpes desenvolvidas sem proteases (como por exemplo

estirpes sem o gene *PEP4* que codifica para uma proteinase A). As estirpes SMD1163 (*his4 pep4prb1*), SMD1165 (*his4 prb1*) ou SMD1168 (*his4 pep4*) mostraram-se eficazes na diminuição da degradação de proteínas heterólogas (Cereghino 2000). No entanto, as estirpes deficientes em proteases são menos viáveis do que as WT, principalmente quando a deleção é no gene *PEP4*, e têm uma velocidade de crescimento menor, sendo mais difíceis de transfectar. Assim, só é recomendado o uso destas estirpes quando outras tentativas não se mostrarem satisfatórias (Cereghino 2000).

Posto isto, existem várias formas de introduzir DNA-alvo num hospedeiro. Na levedura *P. pastoris*, uma molécula recombinante pode ser introduzida por electroporação, por formação de esferoplastos ou ainda através de métodos de *whole-cell* como os que envolvem cloreto de lítio ou polietilenoglicol₁₀₀. Esta levedura prefere a integração como modo de transfecção, o que maximiza a sua estabilidade e permite obter transformantes com múltiplas cópias do DNA (construções multicópia estáveis). Uma forma de integração passa pela restrição única do vector num local partilhado pelo hospedeiro e pelo vector, como o gene de um marcador (*HIS4*, por exemplo) ou o promotor *AOX1*, seguida da transfecção no mutante auxotrófico correcto. Assim, o terminal livre do DNA estimula a recombinação homóloga que resulta no *crossing-over* nesse *locus*, com frequências de transfecção elevadas (50-80% de transformantes) (Domínguez 1998; Cereghino 2000).

Outra forma é digerir o plasmídeo de modo a que a cassette de expressão e o marcador sejam libertados, flanqueados pelas sequências 5' e 3' *AOX* (Figura 4). Tal resulta num evento de permuta do gene *AOX1* com a cassette de expressão e o marcador, com uma eficiência de transfecção de 10-20%, mais baixa do que para *S. cerevisiae*. Assim, estas estirpes são estirpes que contam apenas com o gene mais fraco *AOX2* para crescimento em metanol (estirpe Mut^S, como explicado mais à frente) (Cereghino 2000).

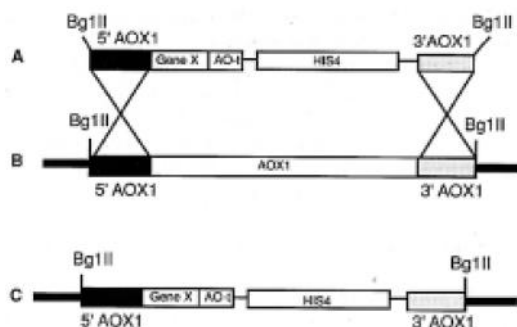


Figura 4 – Representação esquemática do processo de permuta de gene em *Pichia pastoris*. A) DNA derivado do plasmídeo pHIL-D1 digerido com BglII. B) Locus AOX da levedura. C) Permuta do locus AOX1 com o DNA de A) por recombinação homóloga. (Domínguez 1998).

Assim, aquando da expressão heteróloga de uma proteína, são muitos os factores a ter em conta para que esta percorra todos os passos necessários ao *folding* correcto e, por conseguinte, seja activa, se for o caso. Ter o DNA-alvo num vector apropriado e inserido no sistema de expressão escolhido é o primeiro passo para otimizar a produção de proteínas heterólogas, mas os processos de transcrição e tradução compreendem reacções complexas que envolvem inúmeros enzimas e co-factores, às quais deve também ser dada importância (Rai 2001).

Por fim, o último passo na expressão de uma proteína heteróloga é a obtenção de clones que a expressem em larga escala e a selecção da proteína de interesse de modo a purificá-la.

iii. Purificação – um sistema secretor

A levedura *Pichia pastoris* não secreta muitas proteínas endógenas (Higgins 1998; Cereghino 2000). Este facto permite que uma proteína heteróloga que seja expressa e secretada no meio de crescimento seja facilmente purificada. Tal explica-se por o meio de crescimento desta levedura não ter proteínas adicionadas e, assim, a proteína desejada ser grande parte da fracção proteica total do meio (Cereghino 2000).

Este passo de secreção funciona como primeiro passo de purificação pois separa a proteína em estudo das proteínas celulares. Contudo, é necessário ser cauteloso aquando da secreção proteica nesta levedura. O processo de secreção necessita de um sinal peptídico no terminal amina da proteína, mediando a translocação co-traducional para o Retículo Endoplasmático, para que esta seja inserida na via secretora.

Devido a eventuais problemas de *folding* e estabilidade da proteína, este passo é reservado à expressão heteróloga de proteínas normalmente secretadas no seu hospedeiro original. Quando a proteína não é, normalmente, secretada é necessário adicionar uma sequência sinalizadora, com base na proteína a secretar e no seu tamanho.

Geralmente, utilizam-se sequências de sinal incorporadas na cassette de expressão dos plasmídeos e a mais utilizada, e que detém maior número de sucessos, é a sequência que codifica para o factor- α da levedura *S. cerevisiae* (Domínguez 1998; Higgins 1998). Este gene codifica um pré-pro-péptido de 165 resíduos de aminoácidos (Domínguez 1998; Cereghino 2000) e o seu processamento envolve três passos que passam pela remoção do sinal pré-, seguida de clivagens por parte das proteínas Kex2 endopeptidase e Ste13, em locais específicos do péptido (Cereghino 2000). Outra sequência de sinal que pode ser utilizada e que também se encontra disponível em variados plasmídeos é o sinal

da fosfatase ácida (*PHO1*) da levedura *Pichia pastoris* (Cereghino 2000). Estas sequências de sinal podem ser substituídas por sequências de sinal nativas da proteína heteróloga, o que pode otimizar a secreção, em alguns casos (Higgins 1998).

É, no entanto, de frisar que a utilização de um sinal de secreção nativo da proteína de interesse nem sempre é indicativo de sucesso na secreção, tendo havido já casos de melhor sucesso com o factor- α em relação à sequência de sinal nativa. Também já foram indicados casos em que ocorreu uma variabilidade no terminal amina da proteína expressa, aquando da secreção com o factor- α . Noutros casos, tanto o factor- α como o sinal *PHO1* não resultaram na expressão correcta da proteína heteróloga e desenvolveram-se sequências com mutações no sinal nativo a fim de construir um motivo de ligação da Kex2 mais eficiente, ou chegou mesmo a desenvolver-se um sinal pré-pro totalmente sintético (Cereghino 2000).

Por outro lado, para conseguir uma purificação da proteína heteróloga secretada, é necessário ter especial atenção à degradação por proteases no meio de cultura, já que o ambiente de elevada densidade celular em conjunto com a lise de alguma percentagem de células, pode contribuir para a degradação das proteínas sintetizadas e secretadas, como já referido (Arruda 2008).

Ainda assim, as sequências que codificam os *tags* His₆ ou c-myc presentes em muitos vectores desenvolvidos para *Pichia* permitem a fácil purificação das proteínas expressas, por cromatografia.

iv. Modificações pós-traducionais - glicosilação

A levedura *Pichia pastoris* tem o potencial de realizar modificações pós-tradução associadas a Eucariotas superiores. São exemplos o processamento de sequências de sinal, o *fold*ing, a formação de ligações persulfureto ou a O- e N-glicosilação (Higgins 1998), como já referido.

Esta levedura é capaz de adicionar resíduos de hidrocarbonetos aos locais N- e O- de proteínas secretadas, apesar de a O-glicosilação ser praticamente inexistente. Os Eucariotas superiores adicionam vários monossacáridos, ligados por ligações O-glicosídicas, aos grupos hidroxilo dos aminoácidos Serina e Treonina, enquanto que os Eucariotas inferiores, como a *Pichia pastoris*, adicionam apenas O-oligosacáridos de Manose, parecendo não existir um consenso em relação ao aminoácido a ligar. Não deve assumir-se que os resíduos de Ser e Thr escolhidos pela *Pichia* para glicosilar sejam os mesmos que no hospedeiro natural. Também não deve assumir-se que a levedura não glicosilará uma proteína que não é glicosilada no seu hospedeiro nativo.

Relativamente à N-glicosilação, todos os Eucariotas a iniciam no Retículo Endoplasmático, com a transferência de uma unidade oligossacárida ligada a um lípido, $\text{Glc}_2\text{Man}_9\text{GlcNAc}_2$, para o resíduo de asparagina de uma sequência Asn-X-Ser/Thr. Este *core* do oligossacárido é clivado a $\text{Man}_8\text{GlcNAc}_2$ e a glicosilação de Eucariotas superiores e inferiores começa a divergir: em *S. cerevisiae* é ligada uma série de cadeias de manose ao *core* já existente, no aparelho de Golgi. Geralmente, estas cadeias estão hipermanosiladas, ou seja, têm entre 50 a 150 resíduos de manose. Na levedura *P. pastoris* tal não acontece já que a taxa de manosilação é mais baixa e as cadeias exteriores típicas desta levedura são compostas por unidades de $\text{Man}_8\text{GlcNAc}_2$ ou $\text{Man}_9\text{GlcNAc}_2$. O grande problema da hipermanosilação de proteínas heterólogas é a possível interferência com o *fold*ing ou função destas. Nas plantas, uma diferença significativa no que toca à N-glicosilação de proteínas é a existência de resíduos de xilose, nunca presentes noutros organismos Eucariotas (Cereghino 2000; Gomord 2010).

Em conclusão, as variadas combinações marcador/hospedeiro, os promotores alternativos e a facilidade de manipulação tornam possível a expressão de proteínas heteroligoméricas e de cofactores na levedura *Pichia pastoris*. Pouco se sabe, ainda, sobre a regulação do promotor AOX1 a nível molecular, e são também necessários mais estudos sobre problemas associados com a secreção, nomeadamente, sinais de secreção, assim como sobre a glicosilação e os proteases endógenos da levedura.

III. Os quitinases ácidos de Macieira

As plantas não apresentam um sistema imunitário estando por isso mais vulneráveis a patogénios (Huynh 1992). Respondem a ataques de invasores pela indução da expressão de genes específicos que codificam muitas proteínas defensivas (Collinge 1993).

Uma das respostas defensivas contra patogénios mais estudada é a expressão de proteínas relacionadas com patogénese (proteínas-PR, de *pathogenesis-related*). Assume-se que uma proteína nativa que seja expressa *de novo* após indução por qualquer tipo de agente infeccioso, ou condição comparável, é uma proteína PR (Datta 1999). Este tipo de proteínas é não só induzido por vírus mas também por bactérias e fungos (Datta 1999), e podem ser induzidas em condições de stress (Kasprzewska 2003).

Com base nas propriedades biológicas, actividades enzimáticas e semelhança genética, agruparam-se as proteínas PR em 14 classes, sendo identificadas quatro famílias de quitinases, as famílias PR-3, 4, 8 e 11 (Datta 1999). As proteínas PR induzidas por ácido salicílico são consideradas proteínas PR ácidas e as induzidas por etileno ou ácido jasmónico são consideradas proteínas PR básicas (Kasprzewska 2003).

Os quitinases, de nome sistemático (1→4)-2-acetamido-2-deoxi-β-D-glucano glicanohidrolase, são enzimas hidrolíticos ubíquos em muitos organismos. Hidrolisam ligações β-1,4 O-glicosídicas de N-acetilglucosamina (GlcNAc), um constituinte estrutural de inúmeros patogénios (Kramer 1997; Datta 1999).

O principal substrato destes enzimas é a quitina, um polissacárido estrutural da parede celular de muitos fungos, algas e bactérias e do exoesqueleto de invertebrados (Datta 1999; Kasprzewska 2003). A quitina é um homopolímero linear longo de N-acetilglucosamina, com ligações β1→4, não ramificado e com uma configuração em hélice (Kramer 1997; Kasprzewska 2003). Dado que até à data não foi identificada a presença de quitina em plantas, e tendo em conta a função digestiva dos quitinases, crê-se que estes enzimas podem actuar como biopesticidas em plantas (Collinge 1993; Kramer 1997).

Presentes em muitos organismos como em bactérias, fungos, insectos ou vertebrados, para além de degradar quitina, os quitinases têm outras funções: em bactérias são fundamentais no catabolismo da quitina para questões nutricionais e em fungos participam em processos de morfogénese. Mas, em organismos sem quitina, como plantas e animais, é a actividade quitinolítica que se destaca (Kasprzewska 2003).

O seu número da *Enzyme Commission* é o EC 3.2.1.14 que corresponde à classe dos hidrolases, à subclasse dos glicosilases e à sub-subclasse dos glicosidasas (IUBMB). Distinguem-se dois grandes tipos de quitinases:

- ✂ Os endoquitinases, que hidrolisam ao acaso ligações β-1,4 internas de GlcNAc, formando múltímeros de GlcNAc de baixa massa molecular. Estes, principalmente enzimas que apresentem um pI elevado, podem apresentar actividade de lisozima (Collinge 1993; Kasprzewska 2003);
- ✂ Os exoquitinases, que hidrolisam as ligações β-1,4 apenas nos terminais não redutores de GlcNAc da quitina (Collinge 1993; Kasprzewska 2003).

Dado o grande interesse dos quitinases como proteínas PR, e dado que foi a primeira proteína induzida por patogénios cuja função foi identificada, muitos quitinases de muitas plantas foram purificados e/ou clonados (Collinge 1993).

As plantas regulam a expressão dos quitinases conforme o seu estado de desenvolvimento mas também de acordo com o tecido, podendo estar localizados tanto no apoplasto como no vacúolo. Estes enzimas são muito variados, diferindo também na sua estrutura molecular, no gene que os expressa, na especificidade para o substrato e ainda na sensibilidade a inibidores e no mecanismo de catálise (Kasprzewska 2003). Assim, foi necessário classificar os quitinases em seis classes (I-VI), tendo sido ainda classificados de acordo com a sua função de glicosilhidrolases, a nomenclatura PR e a família génica, surgindo onze classes distintas (Datta 1999).

Em plantas saudáveis ocorre a síntese constitutiva de alguns quitinases, tanto no vacúolo como no apoplasto, sendo que a expressão constitutiva destas proteínas aumenta com a idade da planta. Neste caso, os quitinases participam na regulação de processos de crescimento e desenvolvimento, nomeadamente na produção ou degradação de moléculas sinalizadoras (Kasprzewska 2003).

A expressão de quitinases em plantas é estimulada pelo ataque patogénico, podendo ser o aumento da expressão local ou sistémico. Um ataque por parte de um patógeno induz um aumento endógeno de ácido salicílico e de ácido jasmónico que induzem a expressão de vários quitinases. Já quando o ataque é viral, a expressão dos quitinases é menos extensa e mais lenta e podem não ser induzidos. Então, este processo varia com a interacção planta-patógeno e com a isoforma de quitinase (Kasprzewska 2003). Mas não são apenas estes os factores que induzem a expressão destas proteínas de defesa. Factores como ferimento, seca, frio, ozono, metais pesados, salinidade extensa ou mesmo luz UV podem desencadear um aumento na transcrição dos genes dos quitinases (Kasprzewska 2003). Assim, a hipótese de que os quitinases podem funcionar como biopesticidas é reforçada.

Quando induzidos por fungos patogénicos, os quitinases funcionam de duas formas. Os quitinases do apoplasto actuam em fases iniciais da patogénese, libertam moléculas indutoras que transportam o sinal de infecção das hifas ao espaço intercelular. Estes sinais ligam-se a receptores celulares próprios activando os mecanismos de defesa activa, ou seja, são estimuladas a síntese e secreção de mais quitinases do apoplasto, assim como de quitinases do vacúolo. A formação de mais quitinases do apoplasto intensifica a produção de moléculas sinalizadoras e a transmissão do sinal de infecção. Quando as hifas penetram e destroem a célula, os quitinases do vacúolo degradam as novas cadeias de quitina do fungo e reprimem o seu crescimento. Assim, esta hipótese postula que os quitinases do apoplasto actuam como sinalizadores da infecção por parte do fungo, enquanto que os quitinases do vacúolo inibem o crescimento do patógeno (Kasprzewska 2003).

Das várias classes de quitinases, os da classe III são os menos estudados e conhecidos, embora se saiba que não têm qualquer semelhança estrutural com os outros tipos de quitinases. Estes enzimas apresentam sim uma região algo semelhante aos quitinases de Procariotas (Nagasaki 1997). Por comparação com quitinases de classe III isolados de algumas plantas, podem reunir-se características gerais destes enzimas.

Os quitinases de classe III de plantas apresentam uma média de 300 aminoácidos (Nagasaki 1997) com uma massa molecular que varia entre os 30 e os 35 kDa (Koga 1999; Singh 2007). Estas proteínas têm seis resíduos de cisteína conservados, importantes para o estabelecer e para a manutenção da sua estrutura tridimensional (Nagasaki 1997), sendo a sua estrutura não catalítica a menos conservada e o centro catalítico o mais conservado (Kramer 1997). Este último tem apenas duas zonas conservadas com um resíduo de ácido glutâmico que se pensa ser essencial para a actividade de quitinase (Nagasaki 1997).

Do ponto de vista estrutural, os quitinases da classe III apresentam oito hélices- α e oito folhas- β (Fukamizo 2003) ou seja, têm uma estrutura em barril $(\alpha/\beta)_8$ com um sulco no terminal-C das folhas β , semelhante aos quitinases bacterianos da família 18 (Kramer 1997).

O ponto isoeléctrico dos quitinases varia com a sua localização, sendo conhecidos por terem pIs muito distintos. Os quitinases de classe III básicos, os vacuolares, têm pIs elevados (Kasprzewska 2003), enquanto que os ácidos, secretados para o apoplasto, têm pIs na ordem das 5-7 unidades de pH (Kramer 1997). Estes enzimas estão activos a um pH entre 4 e 9, embora o pH óptimo varie com o substrato utilizado (Kramer 1997; Koga 1999). São estáveis a temperaturas elevadas, a mais de 80°C (Koga 1999).

A estrutura dos genes que expressam quitinases de classe III é pouco conhecida e, embora tenham sido determinadas estruturas de genes que codificam quitinases de *Arabidopsis* e de outras plantas, as suas estruturas são muito divergentes, não sendo possível determinar uma estrutura geral para estes genes (Datta 1999).

Foram isolados, pelo grupo de trabalho, dois genes distintos que codificam quitinases ácidos da classe III, em macieira. São expressos com um sinal peptídico que os localiza no apoplasto para poderem, então, sinalizar a invasão de um patogénio e contribuir, desta forma, para a protecção da planta. Um destes genes, denominado *Chi1* (Figura 5, accession number: AF309514_1), tem cerca de 900 pares de base, e codifica para uma proteína com 299 aminoácidos e 31,6 kDa.

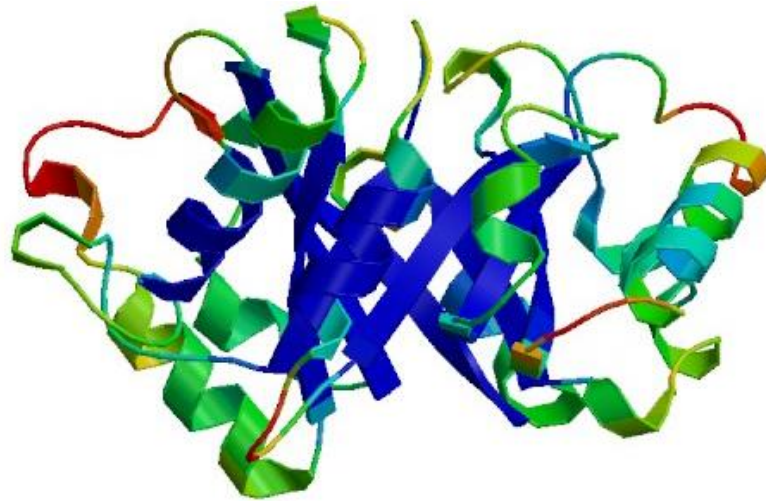


Figura 5 – Modelação por homologia do enzima quitinase ácido da classe III de macieira, codificado pelo gene *Chi1*. Estrutura em *ribbon*. Programa SPDB Viewer.

Estes genes apresentam características semelhantes a genes de *Rhizopus oligosporus*, que já provaram ter actividade antifúngica (Yanai 1992), sendo por isso considerados bons candidatos a biopesticidas.

MATERIAIS E MÉTODOS

Material

i. Microrganismos

Os diferentes microrganismos usados neste trabalho encontram-se listados no quadro 5. Estirpes de *P. pastoris his4* (ou *arg4*) têm uma mutação no gene que codifica a histidinol desidrogenase (ou arginosuccinato liase) que permite a selecção de transformantes por complementação auxotrófica em meio mínimo sem histidina (ou arginina). No caso de o vector da transfecção não conter o gene *HIS4* para complementação, é necessário adicionar histidina ao meio de cultura. As estirpes *aox1::ARG4* têm uma substituição do gene *aox1* pelo gene que codifica para a arginosuccinato liase *ARG4*, não sendo necessário adicionar arginina ao meio (a menos que o plasmídeo introduzido obrigue à remoção desta sequência). Todos os microrganismos foram adquiridos à Invitrogen.

Quadro 4 – Características dos microrganismos utilizados: organismos, estirpes, genótipos e fenótipos.

Organismo	Estirpe	Genótipo	Fenótipo
<i>Escherichia coli</i>	DH5 α	F' supE44, Δ lacU169(Φ 80lacZ Δ M15), thi-1, (r $_{k}^{-}$, m $_{k}^{-}$), <i>recA1</i> , <i>relA1</i> , <i>endA1</i> , <i>gyrA96</i>	-
<i>Pichia pastoris</i>	GS115 KM71H	<i>his4</i> <i>arg4 aox1::ARG4</i>	Mut $^{+}$, His $^{-}$ Mut S , Arg $^{+}$

A estirpe de *E. coli* foi utilizada para testar a correcta inserção dos genes a clonar no *MCS*, nomeadamente, a inserção *in frame* das sequências. As estirpes de *P. pastoris* foram utilizadas para expressar as proteínas em estudo.



Figura 6 – Levedura *Pichia pastoris*: estirpe KM71H utilizada e cultivada em laboratório.



As composições dos meios de cultura utilizados encontram-se listadas no ANEXO 1.

ii. Reagentes

Todos os reagentes utilizados na preparação de tampões e soluções foram adquiridos à Sigma, excepto quando mencionado em contrário. As composições dos tampões e soluções utilizados encontram-se listados no ANEXO 2. Os géis de agarose utilizados neste trabalho foram preparados com a adição do reagente

fluorescente *Greensafe* (2,5µL/100mL de gel) da nzytech[®] que permite visualizar o DNA corado sobre uma lâmpada de luz UV, já que se liga à molécula.

iii. Vectores

Os vectores usados neste trabalho encontram-se listados no Quadro 5, tendo sido adquiridos à Invitrogen.

Quadro 5 – Características dos vectores plasmídicos utilizados.

Vector	Características principais	Tamanho
pCRII	<i>Plac</i> , <i>lacZα</i> , MCS, Kan, Amp, pUC ori, f1 ori, <i>priming sites</i> Sp6, T7, M13rev e M13fwd, promotores Sp6 e T7	3971pb
pPICZαA	5'AOX1 (com PAOX1), factor-α, MCS, tag <i>c-myc</i> e <i>c-polihis</i> , TTAOX1, promotor <i>TEF1</i> e EM7, gene <i>Sh ble</i> , <i>TTCY1</i> , pUC ori	3593pb

O vector pCRII de *E. coli* foi utilizado para testar a correcta inserção dos locais de restrição nas sequências dos genes em estudo, necessária à clonagem *in frame* do AOX1 com o gene codificador da quitinase. O vector pPICZαA de *P. pastoris* foi utilizado para introduzir as sequências-alvo na levedura e expressar as proteínas em estudo.

iv. Oligonucleótidos

Os *primers* utilizados ao longo do trabalho foram os listados no Quadro 6. Todos os *primers* foram desenhados pelo grupo de trabalho, à excepção dos *primers* 5'AOX, 3'AOX e α-factor, que acompanham o *kit* adquirido para a transfecção da levedura. Os *primers* liofilizados foram dissolvidos em TE pH8,0 para uma concentração de 1 nmol/µL de modo a conseguir uma solução *stock*. Foram, depois, diluídos em água destilada esterilizada para uma concentração de 10 pmol/µL sendo esta a solução de trabalho. As temperaturas de desnaturação (T_m) foram calculadas pelo conteúdo em adenina (A), timina (T), citosina (C) e guanina (G), atribuindo 2°C a cada A ou T e 4°C a cada C ou G. Todos os *primers* referidos foram adquiridos à StabVida.

Os *primers* C525Eco-F, C209Eco-F, C203Eco-F, C525Xho-F, C209Xho-F, C525Not-R e C290Not-R foram utilizados para amplificar sequências específicas que codificam os genes de interesse de forma a colocar locais de restrição nas extremidades dessas sequências, que permitissem a sua subsequente clonagem no vector pPICZαA. Os *primers* T7Seq-F e M13puc48-R foram utilizados para ampliar a sequência completa inserida no vector pCRII para sequenciação. Os *primers* 5'AOX, 3'AOX e α-factor foram utilizados para ampliar a sequência inserida no vector pPICZαA já que complementam locais deste vector que flanqueiam a inserção. Os *primers* C525-3', C209-3' e C203-3' foram utilizados na construção de sondas marcadas com digoxigenina para análise de DNA específico

Quadro 6 – Primers utilizados para amplificar e detectar sequências específicas com respectivas sequências e temperaturas de desnaturação (T_m).

Designação	Sequência	T _m
C525Eco-F	5' NNNGAATTCCTCGGGGCTAATGCCGGTGGGA 3'	70°C
C525Xho-F	5' NNNCTCGAGAAAAGAATGGCTTCAAAGTCCACAGC 3'	54°C
C209Eco-F	5' NNNGAATTCGGTGGGAATAGCTATATATTGGG 3'	62°C
C209Xho-F	5' NNNCTCGAGAAAAGAATGGCTTCAAAGTCCACAATG 3'	74°C
C203Eco-F	5' NNNGAATTCCCAGAAATCAAGTCATGCCAAG 3'	64°C
C525Not-R	5' NNNGCGGCCGCGACATCATTCTTGATGGAGGAGC 3'	68°C
C209Not-R	5' NNNGCGGCCGCGACATGACTCTTGATGGAGGAG 3'	66°C
T7Seq-F	5' CGTAATACGACTCACTATAGG 3'	60°C
M13puc48-R	5' AGCGGATAACAATTCACACAGGA 3'	68°C
5' AOX	5' GACTGGTTCCAATTGACAAGC 3'	62°C
3' AOX	5' GCAAATGGCATTCTGACATCC 3'	62°C
α-factor	5' TACTATTGCCAGCATTGCTGC 3'	62°C
C525-3'	5' CCTGAGGCACATGTTTCTGCTA 3'	66°C
C209-3'	5' TCCTCCAATGGAAAGTATGACC 3'	64°C
C203-3'	5' AGTTGTTCCACAAGTCAGTTGC 3'	64°C

v. Marcadores Moleculares

Foram utilizados diferentes marcadores moleculares ao longo deste trabalho, o marcador *O'GeneRuler™ 100bp DNA Ladder* da Fermentas[®], para fragmentos até 1000pb, e o marcador *Lambda DNA/EcoRI+HindIII, 3* da Fermentas[®] e o marcador *NZYDNA Ladder II* da nzytech[®], para fragmentos entre 21000pb e 550pb, aproximadamente. Os padrões de separação das bandas após electroforese destes marcadores encontram-se no Anexo 3.

vi. Enzimas

Os enzimas de restrição utilizados neste trabalho encontram-se listados no Quadro 7. Outros enzimas foram também utilizados como o enzima *Taq* DNA polimerase e *Taq* DNA polimerase "*long range*" (nzytech[®]), nas reacções de PCR para amplificar uma dada sequência; o enzima T4 DNA ligase de modo a poderem ligar-se as sequências a clonar aos vectores utilizados; o enzima agarase para poder extrair fragmentos de DNA separados em géis de agarose; o enzima RNase para poder eliminar quaisquer contaminações de moléculas de RNA aquando da extracção de DNA plasmídico ou genómico; o enzima liticase que cliva a parede celular das leveduras permitindo extrair o seu DNA genómico; e o enzima Chitinase 18A (GH18) de *Clostridium thermocellum* (nzytech[®]) que serviu como termo de comparação e controlo positivo para os ensaios de actividade quitinolítica realizados no estudo.

Quadro 7 – Enzimas de restrição utilizados, NotI, EcoRI, XhoI, SacI e NheI, respectivo local de restrição, tampão e temperatura de incubação (Ti). Todos estes enzimas foram adquiridos à *Fermentas*®.

Enzima	Local de Restrição	Tampão	Ti
NotI	5' GC'GGCCGC 3'	Orange	37°C
EcoRI	5' G'AATTC 3'	EcoRI	37°C
XhoI	5' C'TCGAG 3'	Red	37°C
SacI	5' GAGCT'C 3'	Ecl136II, SacI	37°C
NheI	5' G'CTAGC 3'	Tango	37°C

vii. Meios de Cultura

O modo de preparação dos meios utilizados encontra-se no Anexo 1. Os meios de cultura utilizados foram:

- LB (*Luria-Bertani*) com canamicina (50 µg/mL), para selecção dos recombinantes bacterianos com pCRII;
- LB_{lowsalt} (*Luria-Bertani Low Salt*) com zeocina (25 µg/mL), para selecção dos recombinantes bacterianos com pPICZaA;
- LB_{comp} (adaptado de *Luria-Bertani*) para crescimento das bactérias competentes;
- SOC para crescimento das bactérias recombinantes aquando do processo de transformação;
- YPD para crescimento das leveduras *Pichia pastoris*;
- YPDS com zeocina (100 µg/mL), para selecção dos recombinantes de levedura com pPICZaA;
- MDH e MMH para confirmação do fenótipo dos recombinantes;
- BMGY e BMMY para crescimento das leveduras recombinantes e expressão da proteína, através da activação do promotor AOX com metanol.

viii. Kits utilizados

Os kits utilizados ao longo deste trabalho foram os listados abaixo, tendo sido seguidas as instruções dos fabricantes, salvo indicação em contrário.

- *TA Cloning*® *Dual Promoter (pCR*®*II)* da Invitrogen para obtenção de moléculas recombinantes com pCRII em *E. coli*;
- *EasySelect*™ *Pichia Expression Kit* da Invitrogen para expressão de proteínas recombinantes com pPICZaA em *Pichia pastoris*;
- *QIAquick PCR Purification Kit*® da QIAGEN para purificar produtos de PCR, para sequenciação;
- *QIAGEN Plasmid Midi Kit*® da QIAGEN para preparar DNA plasmidial em larga escala;
- *PCR DIG Probe Synthesis kit* da Roche para preparar sondas marcadas com digoxigenina.

ix. Análise de proteínas

Os reagentes utilizados na análise do conteúdo proteico (quantificação e SDS-PAGE) e também no ensaio de actividade quitinolítica encontram-se descritos no anexo 2. Para a leitura das absorvâncias a 595nm utilizou-se um espectrofotómetro *GeneQuant pro* e as leituras a 420nm foram efectuadas num espectrofotómetro *Shimadzu UV-2100*. O liofilizador utilizado para concentrar as amostras foi um *Heto powerdry LL3000* e o aparelho utilizado para a electroforese SDS-PAGE foi um *Mini Protean* da *BioRad*.

Métodos

i. Biologia Molecular

TÉCNICAS GERAIS DE BIOLOGIA MOLECULAR

As manipulações de DNA efectuadas neste trabalho, p.e. digestão com enzimas de restrição ou modificação de sequências de DNA, foram realizadas como indicado em Sambrook *et al.* (1989) ou segundo indicação do fabricante.

A análise dos fragmentos de DNA foi realizada por electroforese em gel não desnaturante de agarose, em tampão TAE 1x. Às amostras a analisar foi adicionado 1/5 do volume total de tampão de aplicação (composto por azul de bromofenol e glicerol) de modo a conferir cor e densidade. Aquando da análise de fragmentos menores que 1000pb, foram utilizados géis de 2% (p/v) de agarose. Já para fragmentos maiores, foram utilizados géis de 1,2% (p/v) de agarose. Ao fim de 1 hora de corrida do gel, a 120V, a visualização foi possível sobre uma lâmpada de UV, já que os géis contêm o reagente *Greensafe*. As amostras a analisar foram sempre acompanhadas, no mesmo gel e em paralelo, por um marcador (100bp *ladder* ou lambda *DNA/EcoRI+HindIII* ou *NZYDNA Ladder II*, conforme o tamanho do fragmento a analisar) submetido aos mesmos tratamentos.

CLONAGEM HETERÓLOGA EM *ESCHERICHIA COLI*

i. Obtenção do DNA

No caso da clonagem em pCRII, os fragmentos de DNA codificadores dos quitinases de classe III de macieira já inseridos num vector (trabalho anterior do grupo), foram submetidos a uma Reacção em Cadeia do Polimerase (PCR – de *Polymerase Chain Reaction*) com os *primers* específicos de cada uma das cinco isoformas do gene. Estes *primers* contêm sequências específicas de enzimas de

restrição na extremidade 5', que são adicionadas à sequência amplificada a fim de facilitar a clonagem no local adequado do vector pPICZαA.

Para a clonagem no pPICZαA, é necessário purificar o plasmídeo pCRII com a sequência clonada para depois retirar o fragmento pretendido. Para tal, prepara-se uma cultura bacteriana líquida de 25mL e com o kit *QIAGEN Plasmid Midi Kit*[®] da QIAGEN, purifica-se o DNA plasmídico que é, então, digerido sequencialmente com os enzimas de restrição apropriados. A digestão foi realizada na estufa a 37°C por 2 horas, realizando-se depois uma extracção com PCI para remover o enzima. O tampão foi complementado de modo a adequá-lo ao segundo enzima de restrição, e realizou-se a segunda reacção de digestão com este enzima. Por último, o construto foi separado do vector por electroforese em gel de agarose 1,2% sendo a banda correspondente ao construto cortada do gel. A massa de gel é calculada, o gel é derretido a 70°C e digerido com o enzima agarase por 30 min a 42°C, ficando o DNA em solução. O DNA é, então, precipitado com isopropanol, lavado com álcool a 70% e dissolvido em água esterilizada.

ii. Transformação em *E. coli*

a. Preparação de bactérias competentes

Fizeram-se crescer bactérias *Escherichia coli* DH5α em 2,5mL de meio LB, a 37°C e 170rpm por 18 horas. Após o tempo de incubação transferiu-se o inóculo para um frasco com 250mL de meio LB e fez-se crescer mais 3 horas a 37°C e 170rpm, até uma OD_{600nm} entre 0,3 e 0,4. Ao fim deste tempo o inóculo foi centrifugado a 4°C e 3000rpm por 10 minutos e o precipitado recuperado de cada 50mL de inóculo foi ressuspenso em 5mL de uma solução contendo meio LB e meio LB_{comp}, em partes iguais. Após a centrifugação, trabalhou-se sempre em gelo. Esta solução de bactérias competentes foi depois aliquoteada em 200μL para microtubos pré-refrigerados, que foram submetidos a um choque térmico com azoto líquido. Armazenaram-se, por último, no congelador -80°C.

b. Obtenção da molécula recombinante e transformação

A ligação do produto de PCR obtido ao vector pCRII foi conseguida pelo método TA usando o vector pCRII incluído no *kit TA Cloning*[®] *Dual Promoter (pCR*[®]*II)* da Invitrogen. Utilizou-se para a reacção de ligação, um enzima T4 DNA ligase (*Fermentas*) que não vinha fornecido no kit pois preferiu-se utilizar a temperatura de ligação de 22°C.

No caso das transformações com o vector pPICZαA, é necessário criar extremidades compatíveis entre o vector e o construto a ligar. Com este intuito,

realizou-se uma digestão do vector com os enzimas apropriados. Assim, já é possível ligar as construções ao vector, em presença do enzima T4 DNA ligase e respectivo tampão. Este *mix* é sujeito a uma temperatura de 22°C por, pelo menos, 16 horas.

As ligações foram desnaturadas a 65°C por 15 minutos e posteriormente introduzidas em bactérias *E. coli* DH5α competentes por um método de choque térmico (42°C por 40 segundos), juntando-se seguidamente SOC e crescendo a cultura 1 hora a 37°C e 170rpm. Duas alíquotas de 125µL desta transformação foram plaqueados em placas de meio LB (*Luria-Bertani*) com canamicina (50 µg/mL) e X-Gal (1,75mg), no primeiro caso, ou em placas de meio LB_{is} (*Luria-Bertani low salt*) com zeocina (25 µg/mL), no segundo caso.

c. Cultivo dos microrganismos e inoculação

As placas com a transformação das bactérias DH5α foram cultivadas em estufa a 37°C por 18 horas. Foram seleccionadas cerca de 10 colónias isoladas de forma a realizar inóculos líquidos de 2mL (para controlo) ou 15-20-25mL (para multiplicação). Estes cresceram a 37°C, a 170rpm e nos respectivos meios por, pelo menos, 16 horas. Ao fim deste tempo, a cultura estava densa, sendo centrifugada para subsequente purificação do plasmídeo.

iii. Extracções de ácidos nucleicos

a. Extracção e precipitação do DNA plasmidial

Para extrair em pequena escala o plasmídeo das culturas líquidas que cresceram anteriormente, seguiu-se o protocolo descrito por Sambrook *et al.* (1998) com o nome "Small-scale preparations of plasmid DNA" pp 1.25. Este protocolo foi alterado no sentido de realizar uma digestão de RNA seguida de uma extracção com clorofórmio antes da precipitação do DNA. Estas extracções foram monitorizadas por PCR *standard* e, nos casos em que a transformação foi bem sucedida, foi realizado um PCR em larga escala para obter uma quantidade de produto amplificado elevada.

De modo a obter uma quantidade suficiente de DNA para a clonagem no vector pPICZαA, realizou-se uma purificação em média escala do vector pCRII de bactéria. Para tal utilizou-se o kit *QIAGEN Plasmid Midi Kit*[®] da QIAGEN para recuperar o DNA plasmidial em quantidades elevadas.

b. Purificação do DNA clonado

Utilizou-se o kit *QIAquick PCR Purification Kit*[®] da QIAGEN para purificar produtos de PCR do DNA plasmidial extraído de bactérias recombinantes. O

resultado desta purificação é monitorizado em gel de agarose, como descrito anteriormente. Neste caso, é necessária uma quantificação mais cuidada a fim de poder estimar a concentração de DNA obtida e poder reportá-la para fins de sequenciação.

c. Quantificação de ácidos nucleicos

A quantidade de DNA foi detectada por espectrofotometria do UV pela medição da absorvância a 260nm. A concentração foi calculada tendo em conta que uma unidade de absorvância a 260nm de DNA em dupla hélice corresponde a 50µg de DNA por mL. A quantidade foi, em alguns casos, adicionalmente estimada por comparação do sinal de fluorescência com o reagente Greensafe em gel de agarose, com o sinal emitido por fragmentos de quantidades conhecidas.

iv. Condições gerais de uma reacção de PCR

a. Monitorização das extracções do DNA plasmidial

Para monitorizar e analisar os fragmentos de DNA obtidos foi necessário amplificá-los através de uma PCR com *Taq* DNA polimerase. Uma reacção *standard* é levada a cabo num volume de 25µL com 1x de tampão próprio com Mg²⁺ adicionado, 200µM de cada dNTP, 0,2µM de cada *primer* e 0,5U de enzima. O termociclador utilizado para as amplificações foi o *UNO-Thermoblock* da Biometra e o programa utilizado para a maioria das reacções foi:

- desnaturação inicial: 94°C, 2 min;
- 35 ciclos de:
 - desnaturação: 94°C, 45 sec;
 - emparelhamento: à T_{an} , 1 min, sendo $T_{an} = T_m - 2^\circ\text{C}$;
 - extensão: 72°C, 1 min e 30 sec
- extensão final: 72°C, 5 min.

Aquando do cálculo da temperatura de emparelhamento, a temperatura de desnaturação escolhida foi a menor de entre os dois *primers* presentes, retirando-se a essa 2°C. A amplificação foi monitorizada em gel de agarose, como referido anteriormente.

Após a confirmação da obtenção dos fragmentos desejados, realizou-se uma amplificação por PCR em larga escala (100µL) a fim de se obter uma quantidade significativa de produto de PCR para purificar e prosseguir com a sequenciação. O programa de PCR é o mesmo que foi utilizado para a monitorização das extracções de DNA plasmídico. Esta amplificação foi, também, monitorizada em gel de agarose, como já referido.

b. PCR de colónia

Este PCR tem como objectivo o *screening* rápido de colónias para uma determinada sequência. Realizar este PCR após crescimento bacteriano em placa permite identificar a presença ou não da molécula recombinante sem ser necessário recorrer à sua extracção. Contudo, é um método menos fiável e não quantitativo. O programa de PCR é o mesmo que é utilizado para a monitorização das extracções de DNA plasmídico. Esta amplificação foi, também, monitorizada em gel de agarose.

v. Sequenciação e análise de sequências

Após purificação do DNA amplificado e quantificação do mesmo, uma alíquota de, aproximadamente, 400ng do DNA purificado é enviada para a empresa STABVIDA a fim de sequenciar o DNA. Pela análise dos resultados, nomeadamente, pela identificação dos locais de restrição adicionados e dos locais de emparelhamento dos *primers*, pode confirmar-se a clonagem ou não da sequência alvo.

Para esta identificação, utilizou-se o programa *Geneious* versão 4.8.5. Este programa permite, entre outras funcionalidades, procurar sequências de *primers* conhecidos na sequência alvo, assim como o alinhamento de duas sequências. Permite ainda extrair uma sequência *consensus* entre duas sequências e realizar a sequência complementar invertida da sequência alvo.

vi. Stocks de glicerol

Com o intuito de armazenar exemplares da clonagem realizada fez-se crescer 2mL de cultura no meio correspondente, a 37°C e com agitação, durante a noite. Retirou-se um volume de 500µL de cultura à qual se adicionaram 500µL de glicerol 50%. Esta mistura foi submetida a um choque térmico com azoto líquido de modo a realizar um congelamento rápido e homogéneo, sendo os exemplares armazenados no congelador a uma temperatura de -80°C.

CLONAGEM DE EXPRESSÃO EM *PICHIA PASTORIS*

i. Obtenção do DNA

Realizou-se uma cultura bacteriana em meio líquido para a extracção do DNA plasmidial em média escala, com o kit *QIAGEN Plasmid Midi Kit*[®] da QIAGEN, de modo a obter uma quantidade suficiente dos clones de pPICZaA com os construtos. Recuperou-se, então, o DNA plasmidial em quantidades elevadas e quantificou-se por espectrofotometria.

ii. Transfecção em *Pichia pastoris*

a. Digestão do plasmídeo PICZαA

De modo a poder inserir o DNA plasmidial na levedura *Pichia pastoris*, é necessário linearizá-lo. Para tal, digere-se o DNA com um enzima que o clive só uma vez e fora da região do gene, preferencialmente na zona do promotor 5'AOX. Neste caso, escolheu-se o enzima SacI para o efeito. São necessários 3µg de clone para transfectar uma alíquota de levedura e, assim, digere-se esse material com o enzima, por 2 horas a 37°C.

b. Activação das leveduras *Pichia pastoris* e preparação de células competentes

As leveduras devem ser activadas antes de qualquer outro processo. Para tal, espalhou-se com uma ansa esterilizada a levedura em meio YPD solidificado na diagonal de tubos de vidro esterilizados. As leveduras das duas estirpes cresceram a 30°C em 21 horas. Depois desse processo, as leveduras foram espalhadas em placa de YPD e cresceram nas mesmas condições em 16 ou 18 horas.

As células de levedura *Pichia pastoris* GS115 e KM71H foram tornadas competentes e transformadas como indica o protocolo *EasyComp™ Transformation* do manual do kit *EasySelect™ Pichia Expression Kit* da Invitrogen no qual, a uma alíquota de 50µL de leveduras congeladas lentamente no frigorífico -80°C, se adicionam 3µg do digerido. A transfecção foi, depois, plaqueada em duas placas YPBS+Zeo (100µg/mL) e não em apenas uma como indica o protocolo original. Antes de definir qual a concentração de zeocina que iria ser utilizada ao longo do trabalho, plaquearam-se células das duas estirpes de levedura em estudo transfectadas com o plasmídeo pPICZαA em meio YPDS com concentrações de antibiótico de 100 e 200 µg/mL.

iii. Análise dos transformantes

a. Extracções do DNA e análise por PCR

As colónias que cresceram em placa YPDS+Zeo foram inoculadas em meio líquido MD, ou MDH, crescendo a 30°C durante a noite e com agitação. O DNA total das leveduras foi extraído como indica o protocolo *Total DNA Isolation from Pichia* do manual do kit *EasySelect™ Pichia Expression Kit* da Invitrogen. Neste procedimento, a lise das células é conseguida com o tampão SCED e o enzima Liticase (E.C. 3.2.1.58) que digere a parede celular. Após obtenção do DNA por precipitação em etanol absoluto, o DNA é dissolvido e diluído (1/5) e realizou-se um PCR de modo a poder verificar se se trata de uma colónia recombinante. A reacção

desenvolveu-se como no ponto *IIIa* da secção anterior, tendo utilizado os *primers* do kit que amplificam locais flanqueantes do gene *AOX1*. A visualização dos produtos amplificados ocorreu como anteriormente descrito.

b. PCR long range

O DNA de uma estirpe utilizada, quando amplificado por PCR com os *primers* 5'AOX e 3'AOX, dá origem a uma banda com cerca de 3300pb. Assim, e não tendo sido detectada esta banda com o PCR descrito anteriormente, recorreu-se a um enzima *Taq* DNA polimerase "*long range*" de modo a tentar amplificá-la. A reacção é levada a cabo num volume de 25µL com 1x de tampão próprio, 200µM de cada dNTP, 0,2µM de cada *primer* e 2,5U de enzima. O termociclador utilizado para as amplificações é o *UNO-Thermoblock* da Biometra e o programa utilizado para a maioria das reacções foi:

- desnaturação inicial: 94°C, 3 min;
- 35 ciclos de:
 - desnaturação: 94°C, 45 sec;
 - emparelhamento: à T_{an} , 1 min, sendo $T_{an}=60^{\circ}\text{C}$;
 - extensão: 68°C, 5min
- extensão final: 68°C, 10 min.

Esta amplificação foi, também, monitorizada em gel de agarose, como já referido.

c. Determinação do fenótipo Mut

Para confirmar o fenótipo das colónias positivas repicaram-se as colónias para placas de MMH e MDH e fizeram-se crescer a 30°C, como indica o manual do kit *EasySelect™ Pichia Expression Kit* da Invitrogen.

d. Determinação do número de cópias integradas

A determinação do número de cópias integradas em cada colónia é realizada através de uma hibridação de um *Southern Blot* com uma sonda específica marcada com digoxigenina. Para tal, recorreu-se ao *kit PCR DIG Probe Synthesis* a fim de produzir a sonda por reacção de PCR. Para cada isoforma foi construída uma sonda (C525, C203 e C209) e confirmada a sua síntese. De seguida realizou-se o *Southern blot* com o DNA genómico extraído das colónias positivas, segundo *Sambrooks et al.* (1989). Para tal digeriram-se quantidades aproximadamente semelhantes de DNA de cada construção com o enzima de restrição *NheI*. Separaram-se os vários fragmentos por electroforese em gel de agarose 1%, 1xTAE, incluindo um marcador molecular (*Lambda DNA/EcoRI+HindIII* marcado com digoxigenina) para estimar o tamanho dos fragmentos. Após despurinação leve

(HCl 0,25N, 10min), desnaturação (15 e 30min) e neutralização (2x 15min) do DNA contido no gel, transferiu-se, por capilaridade e durante 18h, o DNA do gel para uma membrana de nylon (*Boehringer Mannheim*), na presença de 10xSSC. Depois da transferência, o DNA foi fixado à membrana (120°C, 2h). Posteriormente, a membrana foi sujeita a uma pré-hibridação num saco de plástico selado (HB1, 3h, 65°C). A hibridação deu-se durante 18h e a 65°C com as sondas diluídas em HB1 (4ng/mL), após desnaturação (10min, 100°C) e arrefecimento em gelo. De seguida, a membrana foi lavada a 65°C com 2xSSC+0,1%SDS 15 e 30min e com 0,2xSSC+0,1%SDS, os mesmos tempos. Já à temperatura ambiente, o DNA foi bloqueado (1h, 1xTampão Bloqueio) e foi adicionada uma diluição 1/20000 do anticorpo anti-digoxigenina (30min). A membrana foi lavada para retirar o anticorpo em excesso (3x 20min) antes da detecção (5min). As sequências desejadas foram detectadas após incubação com CSPD diluído ½ em tampão de detecção e após serem expostas 18h a um filme de raio-X. O filme foi revelado com uma solução de revelação (*Ilford*), lavado com água e a fixação deu-se por imersão de 3min num fixador (*Ilford*).

iv. Stocks de glicerol

Uma vez confirmada a inserção da molécula recombinante no genoma das leveduras e o seu fenótipo, realizaram-se *stocks* de glicerol a 25%, como anteriormente, e armazenaram-se a -80°C após choque com azoto líquido. A viabilidade destes *stocks* foi confirmada pelo plaqueamento em YPDS com zeocina (100 µg/mL).

v. Expressão de estirpes recombinantes

Tendo confirmado, por PCR, que as colónias recombinantes contêm o inserto, é necessário determinar as condições e métodos óptimos para a expressão do gene clonado. Seguindo o protocolo *Expression of recombinant Pichia strains* (p. 31) do manual do kit *EasySelect™ Pichia Expression Kit* da Invitrogen, fizeram-se crescer os *stocks* de glicerol em meio BMGY até uma OD_{600nm} de 2, aproximadamente e diluíram-se as células em meio BMMY, no qual cresceram por sete dias, após indução da expressão com metanol. A cada 18 e 24 horas foi retirada uma alíquota de 1,070 mL para realizar várias análises. Após centrifugação de 1 mL de amostra tanto o sobrenadante como o precipitado são sujeitos a choque térmico com azoto líquido e congelados a -80°C, para posterior análise por SDS-PAGE. Dos restantes 70µL de amostra, 20µL foram usados para doseamento proteico pelo método de Bradford (1976) e 50µL para medição da OD₆₀₀. A cada 24 horas a quantidade de metanol no meio é repostada de modo a poder prosseguir a expressão da proteína, sendo a concentração final de 0,5%.

a. Monitorização do crescimento celular

O crescimento celular das culturas em análise foi monitorizado pela medição da OD_{600nm}. Tal realizou-se com o intuito de perceber se se dava o crescimento ou a morte celular, ao longo do ensaio. As amostras, retiradas ao longo do tempo, foram diluídas de 1/20 em meio BMMY e lidas no espectrofotómetro a 600nm.

ii. **Bioquímica**

ANÁLISE DO ENZIMA QUITINASE

i. Quantificação proteica pelo método de Bradford (1976)

Esta análise permite quantificar a proteína presente nas amostras retiradas ao longo do tempo (ponto v) por medição da absorvância a 595 nm. A 20µL de amostra é adicionado 1mL de reagente de *Bradford* previamente diluído (1/5) e filtrado. Este reagente contém o corante *Coomassie Blue* que é estabilizado quando liga a proteínas: há uma mudança clara na cor do corante e o seu máximo de absorção varia de 465nm para 595nm. Assim, é realizada uma curva padrão com várias concentrações conhecidas de uma solução de albumina de soro bovino (BSA), que permite calcular a quantidade proteica de cada amostra, por extrapolação.

ii. Análise proteica por SDS-PAGE

As amostras obtidas previamente no ponto v foram analisadas por SDS-PAGE descontínuo. O tratamento das amostras foi realizado conforme indicado no manual do kit *EasySelect™ Pichia Expression Kit* da Invitrogen (p. 35). O aparelho *Mini Protean* da *BioRad* foi utilizado para montar a electroforese e todos os procedimentos que dizem respeito à preparação dos géis e das soluções foram descritos por *Sambrooks et al.* (1989). Tanto as amostras de sobrenadante como de precipitados foram analisadas, mas nem todos os pontos foram sujeitos a análise por SDS-PAGE, apenas as amostras que mostraram ter maior quantidade de proteína pelo método de Bradford foram analisadas.

Tendo em conta que os resultados iniciais da análise da proteína em estudo em colónias recombinantes não foi o esperado nem satisfatório, foi necessário concentrar as amostras com um liofilizador *Heto powerdry LL3000* e realizar novamente os estudos para os sobrenadantes amostrados.

iii. Medição da actividade quitinolítica – p-DMAB

A actividade quitinolítica foi avaliada recorrendo a dois métodos.

a. Teste com p-DMAB

A actividade enzimática das amostras retiradas previamente foi testada inicialmente por um ensaio colorimétrico (Yanai 1992; Ano 2002). A 50µL das amostras adicionou-se 200µL de substrato (0,1% de glicolquitina) e incubou-se a mistura a 37°C por 1 hora, de modo a que se desse a hidrólise dos oligómeros de glicolquitina em N-acetilglucosamina. A reacção foi parada com a adição de 400µL de ferricianeto de potássio e, para promover a ligação deste agente aos oligómeros de N-acetilglucosamina, foi fervida a 100°C por 3 minutos. Foi então adicionado o agente colorimétrico (600µL): p-dimetilaminobenzaldeído, seguindo-se nova incubação a 37°C por 20 minutos, de modo a estabilizar a coloração. A quantidade de N-acetilglucosamina formada foi medida pela absorção a 585nm, usando como referência água que sofreu os mesmos procedimentos.

b. Teste com K₄Fe(CN)₆

Adoptando outro método (Uchibori 2000; Ano 2002), testou-se a actividade enzimática das amostras retiradas previamente. A 100µL das amostras adicionou-se 300µL de substrato (0,1% de glicolquitina) e incubou-se a mistura a 40°C por 1 hora, de modo a dar-se a hidrólise dos oligómeros, como no método anterior. A reacção foi parada com a adição de 600µL de ferricianeto de potássio e fervida a 100°C por 15 minutos. A extensão da reacção (quantidade de GlcNAc formada) foi medida pela absorção a 420nm, usando como referência apenas água. Foi realizada, em paralelo, uma curva de calibração com concentrações conhecidas de N-acetil-D-glucosamina (0,5; 1,0; 1,5; 2,0; 2,5; 3,0; 3,5 e 4,0 mg/mL). A 400µL de soluções com as concentrações escolhidas foi adicionado o ferrocianeto de potássio e medida a Abs_{420nm} , como anteriormente. Sabendo que 1 Unidade enzimática (U) corresponde à quantidade de enzima que catalisa a libertação de 1µmol de GlcNAc do substrato por minuto a 40°C (Uchibori 2000), e que 1µmol de N-acetil-D-glucosamina corresponde a 0,55mg/mL, pode calcular-se a U por hora e, assim, extrapolar os valores para as amostras desconhecidas. A quantidade de enzima relativa de cada amostra (U/mg de proteína) é calculada através dos valores obtidos de U e de concentração de proteína total (ponto i).

iv. Previsão da N- e O-glicosilação

De modo a prever os locais de N-glicosilação das proteínas expressas, recorreu-se ao programa *NetNGlyc 1.0 Server*, que procura na sequência-alvo domínios Asn-X-Ser/Thr (em que X é qualquer aminoácido). Para prever os locais de O-glicosilação, recorreu-se ao programa *NetOGlyc 3.1 Server*, que analisa os resíduos de Ser e Thr passíveis de serem O-glicosilados.

RESULTADOS E DISCUSSÃO

Com o objectivo de produzir e caracterizar três isoformas de quitinases ácidos da classe III de macieira, recorreu-se à clonagem heteróloga. Para tal escolheu-se, como sistema de expressão, a levedura *Pichia pastoris* e o vector pPICZ α A (Figura 7), a fim de introduzir as sequências que codificam estes enzimas na levedura. Este vector confere resistência à Zeocina e possibilita a adição de um sinal de secreção de *S. cerevisiae*, o factor- α , às sequências a expressar, sendo esta uma clonagem traducional visto que o vector fornece os sinais para a tradução.

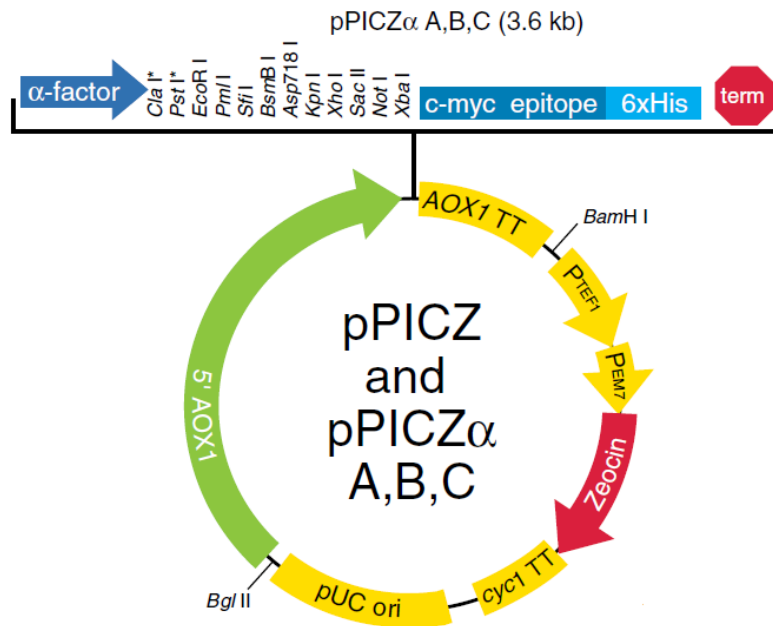


Figura 7 - Representação esquemática do plasmídeo pPICZ α A utilizado neste trabalho. 5'AOX1 – promotor, α -factor – sinal de secreção, c-myc – epítipo, 6xHis – cauda hexahistidinas, AOX1 TT – terminador da transcrição, Zeocin – resistência à zeocina, pUC ori – origem de replicação *high copy*.

Tendo por base sequências previamente obtidas pelo grupo de trabalho (ANEXO 4) e a sequência do vector escolhido (ANEXO 5), realizaram-se cinco construções. As três isoformas da proteína obtidas anteriormente pelo grupo foram denominadas C203, C525 e C209. Destas, a primeira não apresenta sinal peptídico para secreção nem a parte inicial da região codificante (cerca de 200pb em falta, na extremidade 5') e as restantes possuem o sinal peptídico de secreção da macieira (sinal nativo) e a região codificante completa.

Realizaram-se, então, as construções de modo a obter construtos com o sinal de secreção do vector, para as três isoformas e construtos com o sinal de secreção nativo, para as isoformas C209 e C525.

Para clonar com o sinal peptídico do vector, inseriu-se a sequência a expressar no local de restrição EcoRI (a verde na Figura 8). Do mesmo modo, para clonar com o sinal peptídico nativo, inseriu-se a sequência a expressar no local de restrição XhoI (a castanho na Figura 8). Isto porque, ao clonar em XhoI, o local de

sinalização para a clivagem por Kex2 é retirado (a azul na Figura 8), não ocorrendo o processamento correcto da sequência do factor- α . O terminal 3' foi clonado em NotI (a vermelho na Figura 8) de modo a conseguir uma clonagem *in frame* das sequências. Estes enzimas foram seleccionados de entre os presentes no MCS por não clivarem nenhuma das sequências-alvo.

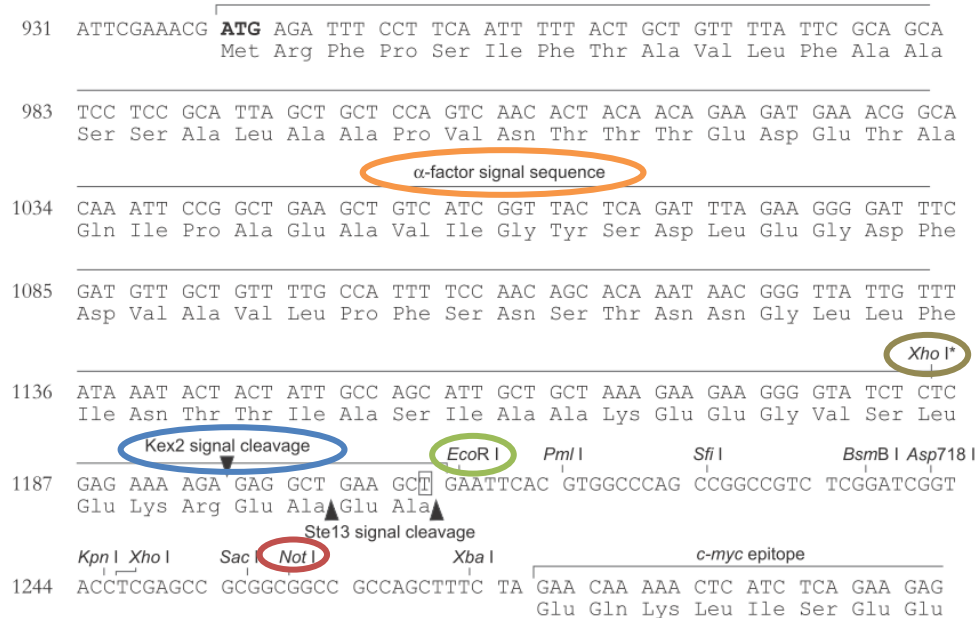


Figura 8 – Porção da sequência nucleotídica do plasmídeo pIZC2 α A correspondente ao MCS. Locais de restrição utilizados no trabalho: EcoRI (verde), XhoI (castanho) e NotI (vermelho). Local de processamento do factor- α (laranja) pelo endopeptidase Kex2 (azul).

Para tal, e antes de se introduzirem as sequências no vector apropriado, adicionaram-se as sequências de restrição dos enzimas XhoI ou EcoRI e NotI ao DNA a clonar, por PCR com *primers* específicos com as sequências de restrição na extremidade 5' (Figura 9). Para a isoforma C203 só se inseriram as sequências de restrição EcoRI e NotI pois pretende-se, com esta construção, estudar apenas o efeito do sinal de secreção do vector, tendo em conta que esta isoforma não tem um sinal de secreção nativo.

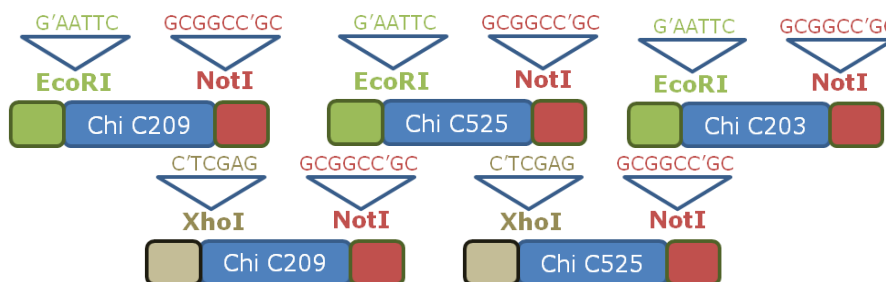


Figura 9 – Representação esquemática das construções realizadas, com respectivos locais de restrição: Eco+Not e Xho+Not.

Uma vez inseridos os locais de restrição nas sequências em estudo, ligaram-se as construções ao vector pCRII (ANEXO 5B), clonaram-se em *E. coli* e, por sequenciação, verificou-se a correcta inserção dos locais de restrição nas

sequências das isoformas, para as várias construções (resultados não apresentados).

As sequências com os locais de restrição adicionados foram retiradas do vector pRCII, por digestão enzimática, e clonadas no vector pPICZ α em *E. coli*, como demonstra a Figura 10. Sendo este um vector *shuttle*, é passível de ser clonado tanto em levedura como em bactéria. Por sequenciação verificou-se a correcta inserção *in frame* das construções no vector, cujos resultados se encontram no ANEXO 6.

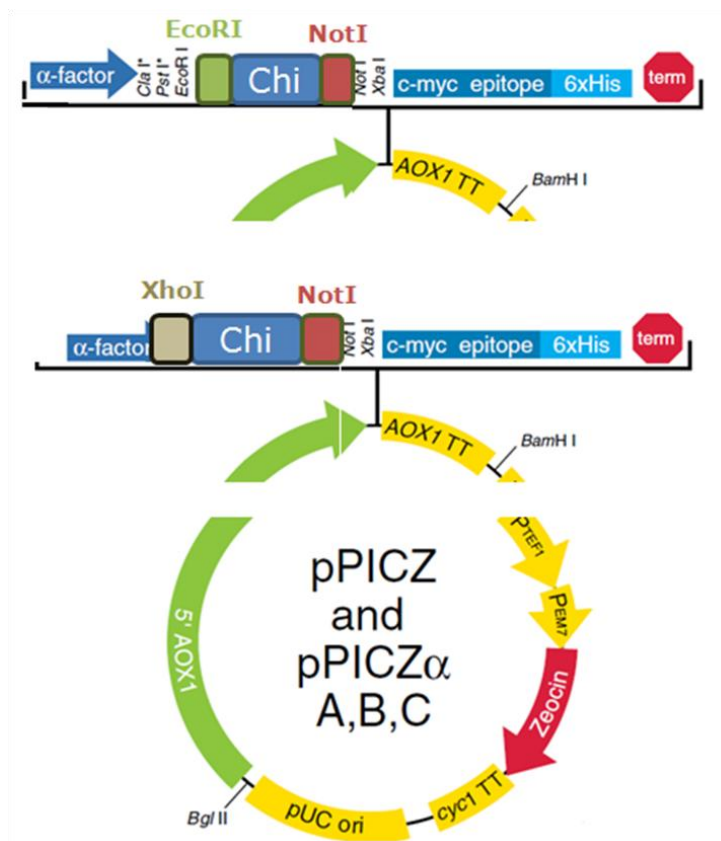


Figura 10 – Ligações realizadas com o plasmídeo pPICZ α A para as várias construções: EcoRI+NotI ou XhoI+NotI.

Após análise das sequências obtidas, e utilizando o software *Geneious* 4.8.5, concluiu-se que as sequências das isoformas estão correctas, assim como os locais de restrição inseridos por PCR, e que as sequências a clonar foram inseridas *in frame*.

Procedeu-se à multiplicação da molécula recombinante por crescimento em meio apropriado (LB₅+Zeo), seguindo-se a purificação do DNA plasmidial em larga escala.

Para a transfecção da levedura *Pichia pastoris*, que ocorre através de moléculas de DNA linear, obrigando a que o vector seja digerido com um enzima antes de ser introduzido na levedura, escolheu-se o enzima SacI para a digestão,

pois cliva no promotor 5'AOX e não cliva nenhuma das isoformas, garantindo que toda a sequência a expressar é integrada (Figura 11).

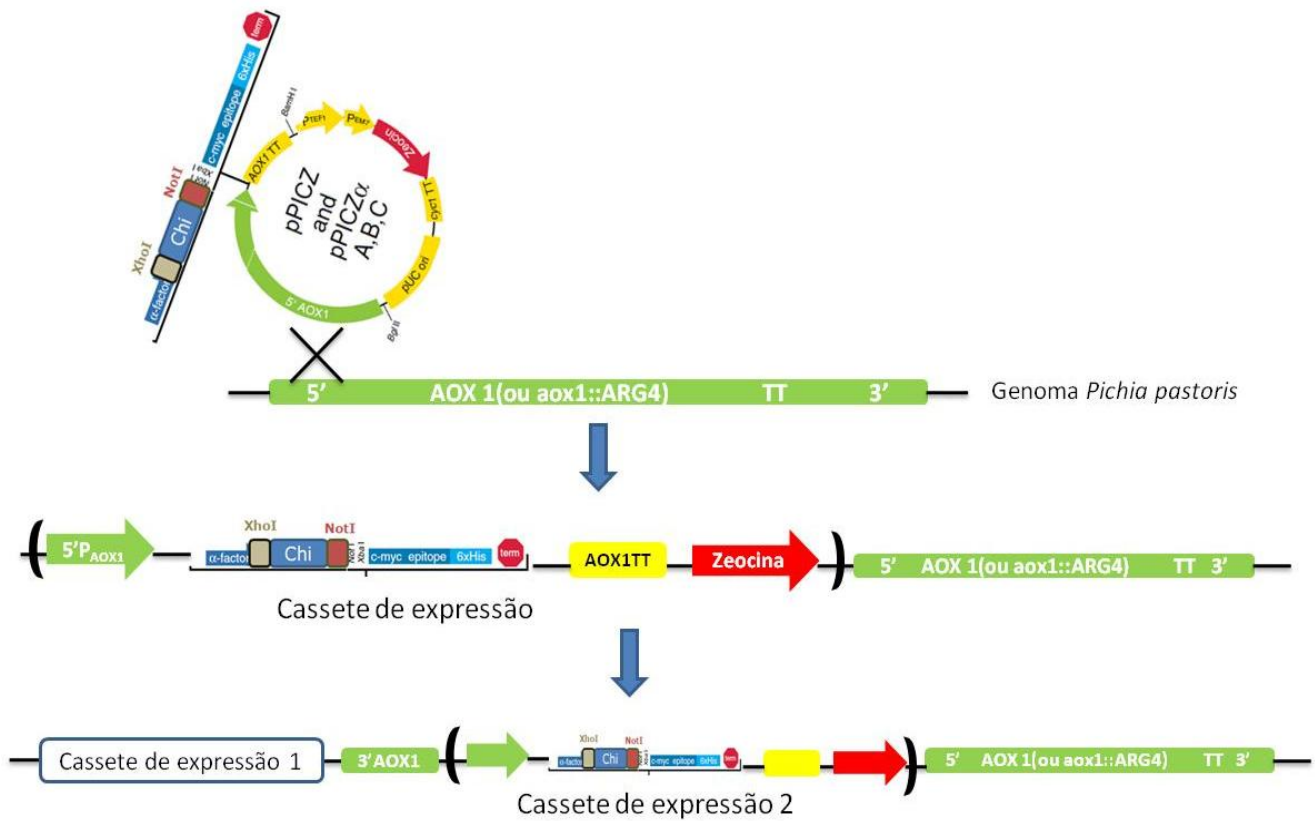


Figura 11 – Exemplo de uma transfeção efectuada com a sequência que codifica para uma isoforma de quitinase, com o sinal de secreção nativo, inserida no vector pPICZαA. A sequência a clonar foi inserida no genoma da levedura *Pichia pastoris* após linearização da molécula recombinante. É possível a inserção de múltiplas cópias desta molécula (cassete de expressão).

Escolheram-se duas estirpes para a transfeção já que diferentes ambientes genéticos podem influenciar a expressão da proteína clonada (Invitrogen 2005). Essas estirpes foram a GS115, que é uma estirpe Mut⁺ e a KM71H que é uma estirpe Mut^S. Antes de ser realizada a transfeção, foi analisada a quantidade de antibiótico necessária para a escolha dos transformantes. Utilizaram-se duas concentrações distintas de Zeocina no meio YPDS, 100µg/mL e 200µg/mL, e inocularam-se células competentes das duas estirpes com vector nulo, mas não foram observadas diferenças significativas no crescimento celular destas. Assim, continuou-se o estudo apenas com a primeira concentração de antibiótico, 100µg/mL.

Neste seguimento, tentou-se transfectar as cinco construções em pPICZαA nas estirpes competentes GS115 (Mut⁺) e KM71H (Mut^S). Os resultados obtidos nas transfecções encontram-se no Quadro 8, que permitiu comparar a quantidade aproximada de DNA plasmídico inserido na levedura e o número de colónias obtido, para cada construção e estirpe.

Quadro 8 – Resultados obtidos na clonagem das cinco construções (C525EcoNot, C5252XhoNot, C209EcoNot, C209XhoNot e C203EcoNot) nas duas estirpes de *Pichia pastoris*, assim como transfectadas com o vector nulo.

	Construção	Quantidade pDNA inserido	Número de colónias
GS115	C525EcoNot	2,7 µg	4
	C525XhoNot	2,0 µg	2
	C209EcoNot	3,3 µg	6
	C209XhoNot	2,0 µg	5
	C203EcoNot	1,64 µg	3
	pPICZαA	3,3 µg	21
KM71H	C525EcoNot	2,7 µg	7
	C525XhoNot	2,0 µg	7
	C209EcoNot	3,3 µg	2
	C209XhoNot	40 ng	7
	C203EcoNot	1,64 µg	14
	pPICZαA	0,4 µg	40

Pela análise do quadro não existe uma relação entre a quantidade de DNA inserido e o número de colónias obtido. Não há uma diferença significativa entre o número de colónias obtidas para as duas estirpes transfectadas com o mesmo construto, em ambos os casos o número de colónia foi bastante menor do que o que seria de esperar (Invitrogen 2005). Para a construção C5252EcoNot, uma transfecção com a mesma quantidade de DNA origina mais colónias na estirpe KM71H do que na estirpe GS115. O mesmo acontece para as construções C525XhoNot e C203EcoNot mas nota-se o oposto para a construção C209EcoNot. Contudo, é de notar que a transfecção com o vector nulo origina um número muito elevado de colónias, quando comparado com as transfecções com as construções. Tal pode dever-se ao tamanho das construções inseridas já que quanto maior a molécula recombinante, menor o sucesso da transfecção. Assim, conseguiram-se transfectar as cinco construções nas duas estirpes escolhidas com sucesso. A Figura 12 mostra, como exemplo, o procedimento para a construção C525XhoNot na estirpe GS115.

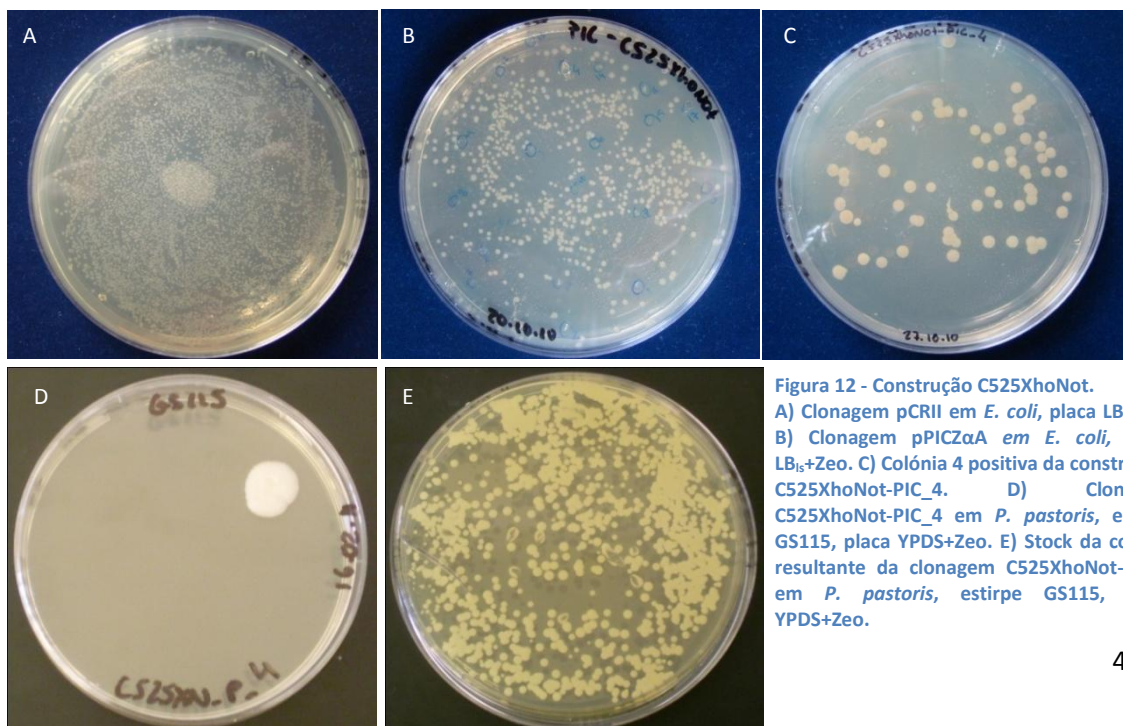


Figura 12 - Construção C525XhoNot. A) Clonagem pCRII em *E. coli*, placa LB+Kan. B) Clonagem pPICZαA em *E. coli*, placa LB₈+Zeo. C) Colónia 4 positiva da construção: C525XhoNot-PIC_4. D) Clonagem C525XhoNot-PIC_4 em *P. pastoris*, estirpe GS115, placa YPDS+Zeo. E) Stock da colónia resultante da clonagem C525XhoNot-PIC_4 em *P. pastoris*, estirpe GS115, placa YPDS+Zeo.

De seguida, as colónias recombinantes obtidas foram sujeitas a um *screening* por PCR com *primers* específicos para confirmar a inserção do plasmídeo com a construção. Para tal, fizeram-se crescer as colónias em meio MD (para estirpes GS115) e MDH (para estirpes KM71H) líquidos e extraiu-se o DNA em pequena escala. Por análise de PCR com *primers* específicos para o vector, foi possível detectar a presença da construção no genoma da levedura. Uma vez que os *primers* utilizados amplificam o gene *AOX1*, tanto o gene nativo como a sequência inserida pelo plasmídeo (que flanqueia a construção) são amplificados. Assim, obtêm-se duas bandas: uma correspondente ao gene *AOX1* da *Pichia pastoris* e outra ao vector recombinante. A banda que diz respeito ao *AOX1* endógeno tem tamanhos diferentes nas duas estirpes utilizadas. No caso da estirpe GS115, o tamanho do produto de PCR é 2200pb mas para a estirpe KM71H, tendo em conta que esta estirpe apresenta a inserção do gene *ARG4* no gene *AOX1*, o tamanho da banda será 3600pb. Neste caso, tentou identificar-se a banda com uma *Taq* DNA polimerase "long range". É a existência da segunda banda que prova a inserção do plasmídeo na levedura e o seu tamanho varia com o construto ligado, como se mostra no Quadro 9.

Quadro 9 – Tamanhos dos fragmentos amplificados pelos *primers* específicos do MCS do vector pPICZαA, para as quatro construções.

Construção	Tamanho da sequência	Tamanho no vector recombinante*
C525XhoNot	923pb	1440pb
C525EcoNot	854pb	1392pb
C209EcoNot	839pb	1377pb
C203EcoNot	656pb	1194pb

*Estes valores foram calculados tendo em conta que o tamanho do produto de PCR com os *primers* 5' e 3'AOX do plasmídeo pPICZαA é 588pb e que, aquando da inserção nos locais de restrição das construções, se perdem alguns nucleótidos da sequência a amplificar. Assim, o tamanho do fragmento, para as construções EcoNot e XhoNot, é 538pb e 517pb, respectivamente, sendo que a porção que se perde corresponde a 50pb e 74pb, respectivamente.

Os resultados das reacções de PCR revelaram que todas as transfectões deram origem a colónias positivas à excepção do clone C209XhoNot, em ambas as estirpes de levedura. Apesar de a sequência a inserir ter sido sequenciada e apresentar-se correcta, as colónias que se formam não são transformantes pois o DNA extraído destas não dá origem a uma banda com o tamanho esperado, quando amplificado com *primers* do vector. Assim sendo, não se seguiram os estudos para esta construção.

Os resultados obtidos nem sempre seguiram o esperado. As bandas amplificadas correspondentes ao gene *AOX1* endógeno nem sempre foram identificadas, o que pode dever-se a constrangimentos na reacção de PCR. A tentativa de identificar a banda de 3600pb referente ao AOX endógeno das estirpes

KM71H com a *Taq* DNA polimerase "long range" resultou positivamente para a estirpe WT. De modo a identificar correctamente a banda referente ao vector recombinante, realizou-se em paralelo uma reacção de PCR com o pDNA extraído de *E. coli*, que contém os insertos clonados em pPICZaA. A Figura 13 mostra as bandas amplificadas com os *primers* do vector pPICZaA (5'AOX e 3'AOX) e os resultados foram listados no Quadro 10.

Quadro 10 – Resultados obtidos para o *screening* das colónias obtidas, tendo como controlo positivo o pDNA extraído de *E. coli* com as construções.

Construção	Estirpe	Bandas	
		<i>AOX1</i> endógeno	pPICZaA
C525XhoNot	KM71H	Não	1440pb
	GS115	Não	1440pb
C525EcoNot	KM71H	Não	1392pb
	GS115	Não	1392pb
C209EcoNot	KM71H	Não	1377pb
	GS115	2200pb	1377pb
C203EcoNot	KM71H	Não	1194pb
	GS115	Não	1194pb

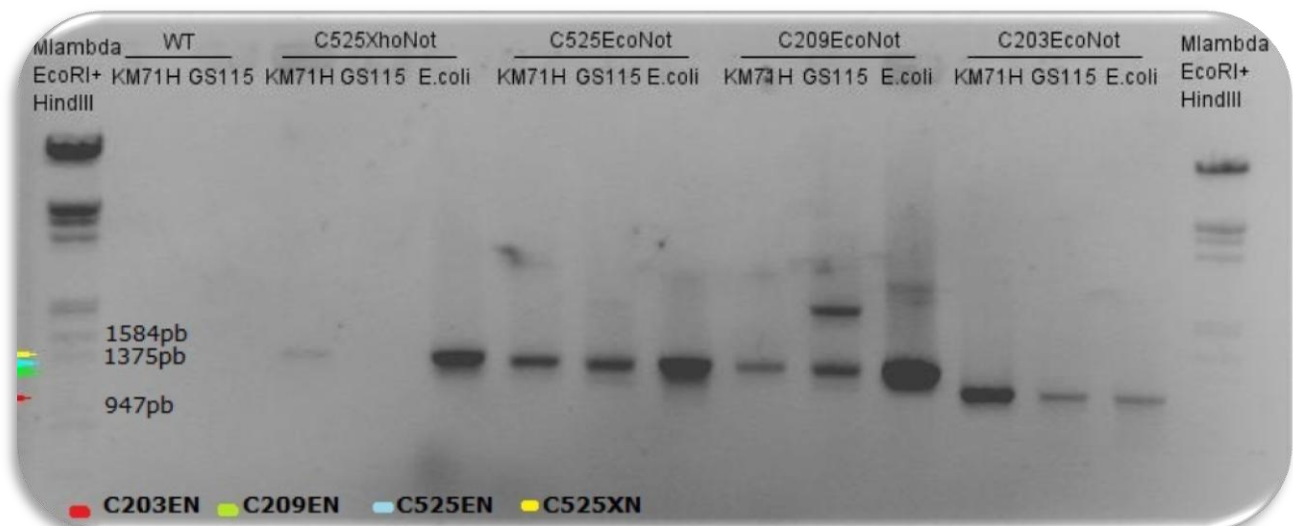


Figura 13 – Gel de agarose 1,2%. Resultados das transfectões das duas estirpes de *Pichia pastoris* com as quatro construções. Amplificação por PCR com *primers* do vector (5'AOX e 3'AOX) do DNA genómico extraído. Marcador lambda DNA/EcoRI+HindIII (poços 1 e 16); KM71H e GS115 WT (2 e 3), +C525XhoNot (4 e 5), +C525EcoNot (7 e 8), +C209EcoNot (10 e 11) e +C203EcoNot (13 e 14); assim como pDNA extraído de *E. coli* transformada com as mesmas construções (controlo positivo – poços 6, 9, 12 e 15).

Como se pode observar na Figura 13, a banda referente à transfecção de GS115 com a construção C525XhoNot não foi amplificada, apesar de ter sido anteriormente detectada noutro gel. É também observável que na transfecção com a construção C209EcoNot em GS115 a banda do gene *AOX1* endógeno foi amplificada. As bandas amplificadas com os *primers* do vector são as bandas esperadas. Ou seja, nas transfectões com a construção C525XhoNot, quer em KM71H quer em GS115 (não apresentado), obteve-se uma banda com cerca de 1440pb, que corresponde ao esperado (Quadro 9 e Figura 8). O mesmo acontece para

as outras transfecções: C525EcoNot, 1392pb; C209EcoNot, 1377pb; C203EcoNot, 1194pb.

Após confirmar a correcta inserção da molécula recombinante nas duas estirpes de levedura foi necessário confirmar-se o fenótipo das transformantes. Para tal, repicaram-se as colónias positivas em placas MDH e MMH para comparar as diferenças no crescimento. A Figura 14 e o Quadro 11 mostram o crescimento destas estirpes, evidenciando o seu diferente comportamento.

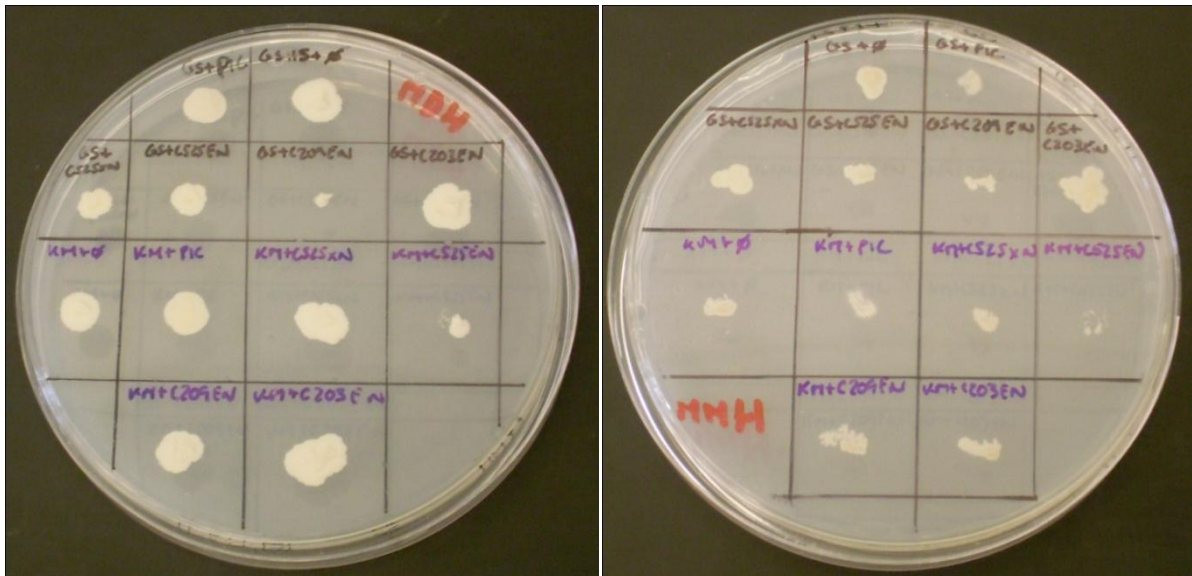


Figura 14 – Placas MDH e MMH referentes ao estudo de confirmação do fenótipo: com as quatro recombinantes (C525EcoNot, C525XhoNot, C209EcoNot e C203EcoNot), com plasmídeo vazio e as duas estirpes WT de levedura.

Comprovado o fenótipo das transformantes, realizaram-se *stocks* de glicerol das colónias positivas e testou-se a viabilidade destes por crescimento em meio YPDS com Zeocina, como na Figura 12E, os resultados se encontram no Quadro 11.

Quadro 11 – Resultados obtidos nos ensaios de confirmação de fenótipo e de viabilidade dos *stocks* de glicerol.

Construção	Estirpe	Crescimento MDH	Crescimento MMH	Viabilidade <i>stock</i>
C525XhoNot	GS115	+++	+++	Confirmada
	KM71H	+++	+	Confirmada
C525EcoNot	GS115	++	++	Confirmada
	KM71H	+	-	Confirmada
C209EcoNot	GS115	+	+	Confirmada
	KM71H	++	+	Confirmada
C203EcoNot	GS115	+++	+++	Confirmada
	KM71H	+++	+	Confirmada

É de notar que todas as construções em GS115, incluindo a construção com o vector nulo e a estirpe WT, tiveram um crescimento comparável nos dois meios. Assim, tratam-se de colónias que crescem em meio com dextrose da mesma forma que em meio com metanol, ou seja, estirpes com fenótipo Mut⁺, como seria de esperar. Por outro lado, todas as construções em KM71H, incluindo a construção com o vector nulo, e a estirpe WT cresceram mais em meio com dextrose, comparativamente ao meio com metanol. Assim, tratam-se de colónias que

crecem menos em meio metanólico, ou seja, estirpes com fenótipo Mut^S, como seria de esperar. Estes resultados evidenciam que as leveduras Mut⁺ crescem mais em metanol do que as leveduras Mut^S, ou seja, a estirpe GS115 cresce mais em meio MMH do que a estirpe KM71H, sendo que em meio MDH o seu crescimento é comparável.

Relativamente ao teste de viabilidade dos *stocks* de glicerol, todas as colónias repicadas das várias construções nas duas estirpes cresceram em meio YPDS+Zeo, o que mostra que as colónias estão viáveis. A extracção de DNA destas colónias demonstrou, também, que se tratam de cópias positivas das construções (resultados não apresentados).

Como esta clonagem envolve a inserção de DNA linear no DNA genómico da levedura, há uma elevada probabilidade de inserir mais do que uma cópia do DNA digerido, podendo obter-se colónias positivas com várias cópias inseridas, como mostra atrás a Figura 11. Assim, procurou identificar-se quantas cópias do constructo apresenta cada transformante, através da realização de um *Southern blot*. Para esta análise, o DNA foi clivado com o enzima NheI que corta fora da cassette de expressão. Utilizou-se o DNA genómico purificado das estirpes transfectadas com as quatro construções, com plasmídeo nú e sem plasmídeo (WT) e ainda o pDNA extraído de *E. coli* e digerido com SacI.

Para a hibridação, tentou-se construir sondas que permitam distinguir as várias construções. Assim, e recorrendo a programas informáticos, desenharam-se *primers* que emparelham com uma zona específica de cada construção: o final da sequência sinalizadora e o início da sequência da proteína. Estes *primers* específicos, em conjunto com o *primer* 5'AOX, permitem construir uma sonda marcada com digoxigenina, por PCR. De seguida hibridou-se o *Southern blot* com as sondas construídas, detectando-se com o anticorpo anti-DIG, usando CSPD como substracto, como mostra a Figura 15.

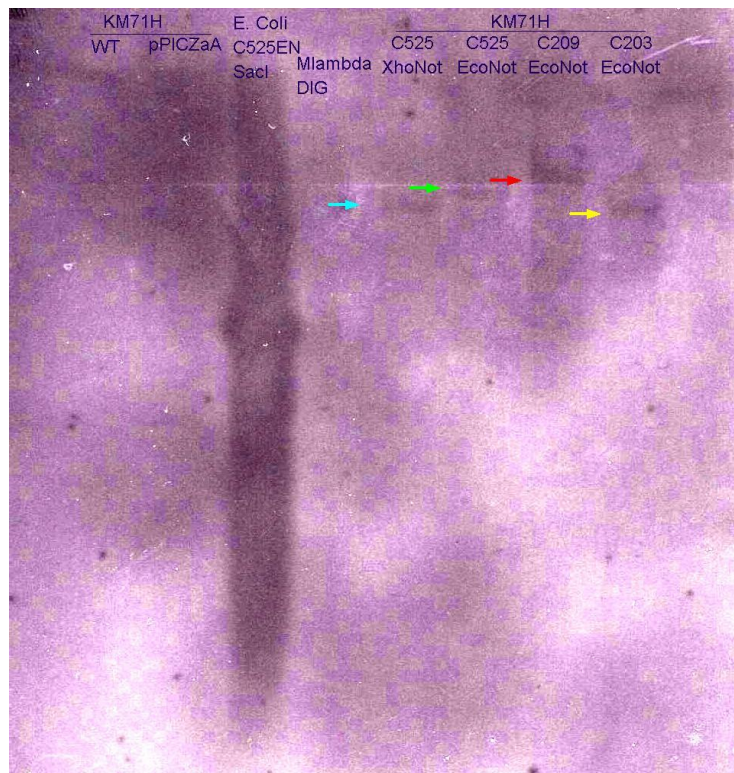


Figura 15 - Resultados do *Southern blot* realizado com as construções inseridas na estirpe KM71H: WT (poço 1), +pPICZαA (poço 2), +C525XhoNot (poço 5), +C525EcoNot (poço 6), +C209EcoNot (poço 7) e +C203EcoNot (poço 8), digeridos com NheI. Marcador lambda com digoxigenina (poço 4) e pPICZαA+C525EcoNot extraído de *E. Coli*, digerido com Sacl (poço 3).

Em análise *dot spot* (resultados não apresentados), verificou-se que as diferentes sondas hibridaram com todas as construções pPICZαA-gene de quitinase, não só com a própria que lhe deu origem, pelo que se revelou desnecessária a utilização de uma sonda específica para cada construção. Este facto pode resultar da presença da mesma sequência sinalizadora nas diferentes construções.

Os resultados obtidos (Figura 15) sugerem que só houve introdução de uma cópia da construção com o gene do quitinase em cada clone, na medida em que não apareceu sinal no WT e no recombinante com vector nulo e apareceu uma banda em todas as outras construções. Se as sondas hibridarem com os genes de quitinases endógenos, facto que não pode ser excluído, ou com o gene AOX1 endógeno, facto que também não pode ser excluído, seria de esperar que aparecesse sinal nas duas primeiras amostras. Assim considerou-se que a hibridação foi específica para o fragmento "AOX1-gene de quitinase introduzido". No entanto, estes resultados são ainda preliminares tornando-se necessário repetir esta análise com maior quantidade de DNA e noutras condições.

Para apurar os métodos e as condições óptimos para a expressão do gene do quitinase introduzido, fizeram-se crescer as colónias recombinantes positivas de cada construção em meio BMMY activando a expressão da proteína pela adição de metanol, a cada 24 horas. Os resultados da quantificação das proteínas totais em solução e das OD_{600nm} das alíquotas, tanto do precipitado (interior celular) como o

sobrenadante (secretoma da levedura), que foram retiradas periodicamente durante 7 dias, encontram-se listados no Anexo 7, Quadro 17.

Após o doseamento proteico, seleccionaram-se, em cada série temporal, as amostras a analisar por SDS-PAGE e para a medição da actividade quitinolítica. Esta escolha recai sobre as amostras que têm mais quantidade de proteína total, pois pensa-se ser nestas que os transformantes mais expressam a proteína clonada. Assim, seleccionaram-se as amostras referentes aos picos de proteína total (negrito) e as que os flanqueiam, como mostra o Quadro 12.

Quadro 12 – Valores de concentração de proteína em mg/mL das amostras seleccionadas para as quatro construções nas duas estirpes (GS115 e KM71H) e para as estirpes WT e transfectadas com vector nulo.

Construção	Estirpe	Horas seleccionadas	[proteína] mg/mL	Construção	Estirpe	Horas seleccionadas	[proteína] mg/mL
WT	GS115	24	0,278	C525EcoNot	GS115	100	0,408
		48	0,359			139	0,401
		72	0,379			163	0,472
		96	0,460			187	0,659
		120	0,491			24	0,606
	KM71H	114	0,299		30	0,636	
		120	0,328		57	0,368	
		138	0,218		100	0,326	
		144	0,271		139	0,123	
		162	0,313		48	0,308	
pPICZαA	GS115	54	0,299	C209EcoNot	GS115	72	0,222
		81	0,514			96	0,319
		124	0,474			120	0,184
		148	0,441			24	2,053
		172	0,502			42	1,751
	KM71H	24	1,078		48	1,555	
		42	1,016		66	1,112	
		48	0,971		72	0,980	
		66	1,202		90	0,750	
		72	1,286		96	0,588	
C525XhoNot	GS115	90	1,447	C203EcoNot	GS115	24	0,230
		96	1,529			48	0,322
		24	0,029			72	0,420
		48	0,182			96	0,349
		72	0,328			120	0,322
	KM71H	96	0,302		24	0,812	
		120	0,367		48	0,789	
		80	0,239		72	0,565	
		114	0,308		96	0,482	
		138	0,292				

Com os dados recolhidos traçaram-se, ainda, dois gráficos para cada estirpe, que mostram como o crescimento celular e a quantidade de proteína variam com o tempo, Figuras 10 e 11.

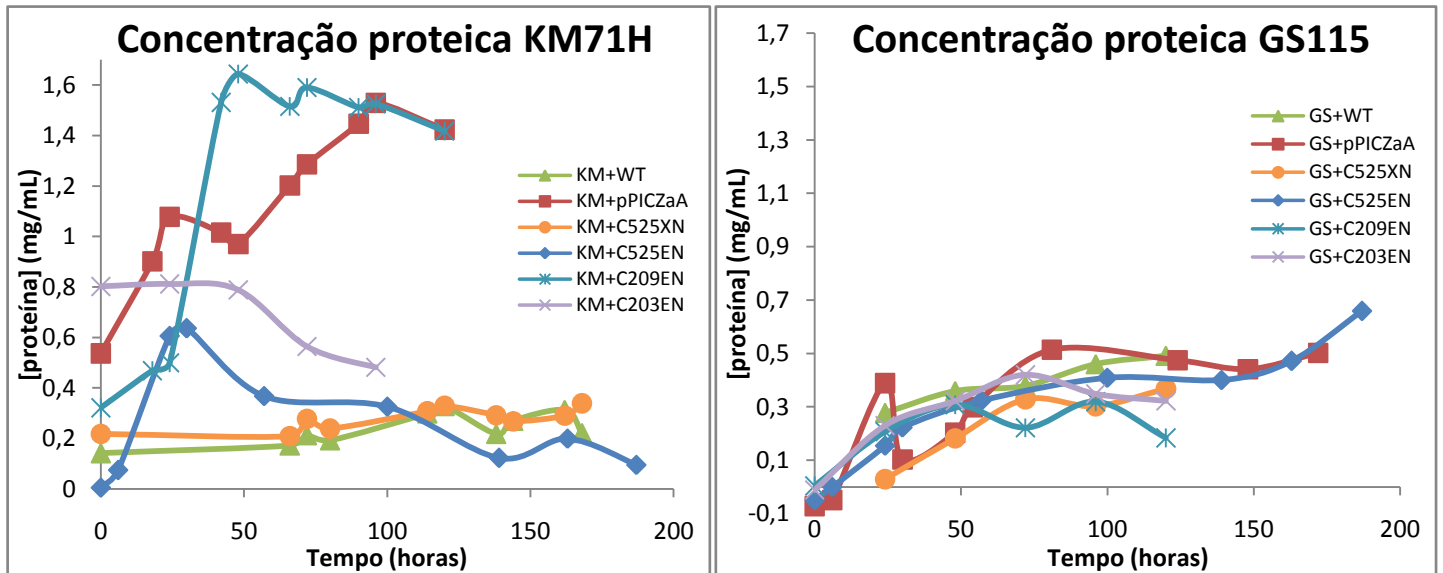


Figura 16 – Variação da concentração proteica ao longo do tempo para as amostras seleccionadas das quatro construções nas duas estirpes (GS115 e KM71H) e para as estirpes WT e transfectadas com vector nulo.

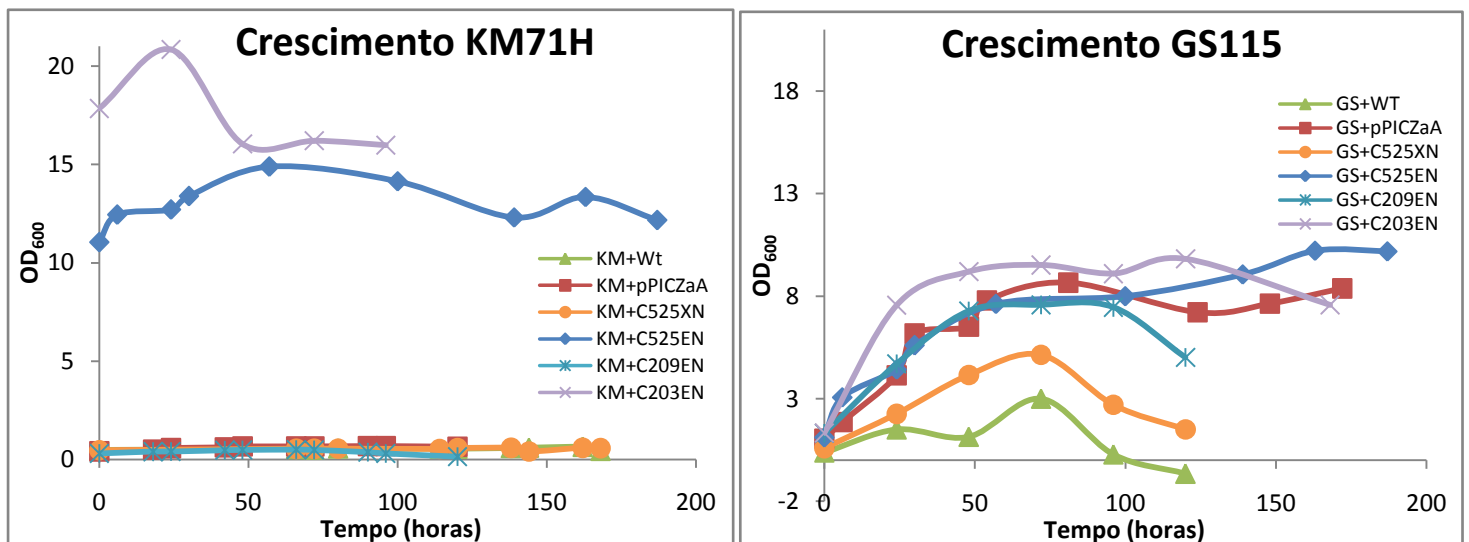


Figura 17 – Variação do crescimento celular ao longo do tempo para as amostras seleccionadas das quatro construções nas duas estirpes (GS115 e KM71H) e para as estirpes WT e transfectadas com vector nulo.

É de notar que as transformações em GS115 apresentam todas uma concentração proteica comparável, o que pode dever-se e está de acordo com o crescimento semelhante que todas mostraram (Figura 16 e 17). Por outro lado, no caso das transformações com KM71H, os resultados são mais díspares, havendo casos em que, apesar de um crescimento baixo, a concentração proteica é elevada (C209EcoNot e pPICZaA), casos em que o crescimento é elevado e a concentração proteica é baixa (C525EcoNot), casos em que tanto o crescimento como a concentração proteica são baixos (C525XhoNot e WT) e ainda casos em que ambos os parâmetros são elevados (C203EcoNot). Tal pode mostrar-se importante aquando da análise da actividade quitinolítica de cada transformante.

Realizaram-se, então, ensaios com o intuito de determinar a quantidade de enzima com actividade quitinolítica presente nas amostras recolhidas e seleccionadas. Numa primeira abordagem, adoptou-se um método colorimético com p-DMAB, mas este não se mostrou robusto já que evidenciava excessiva coloração nos controlos negativos, que dificultava a leitura espectrofotométrica. Assim, alterou-se o método de determinação da actividade quitinolítica com $K_4Fe(CN)_6$, como descrito na secção Materiais e Métodos. Os resultados obtidos para os sobrenadantes recolhidos nos momentos seleccionados encontram-se no Anexo 7 de forma detalhada e o Quadro 13 mostra a actividade quitinolítica calculada para cada amostra seleccionada. Os ensaios enzimáticos realizados com os precipitados (resultados não apresentados) sugeriram actividades quitinolíticas, das diferentes construções, próximas dos valores obtidos para as estirpes WT, apontando para o correcto processamento das sequências introduzidas.

Quadro 13 – Valores de actividade enzimática em U e U/mg de proteína total das amostras seleccionadas para as quatro construções nas duas estirpes (GS115 e KM71H) e para as estirpes WT e transfectadas com vector nulo. A negrito evidenciam-se os valores mais elevados para cada isoforma.

Construção	Estirpe	Horas seleccionadas	U	U/mg	Construção	Estirpe	Horas seleccionadas	U	U/mg
WT	GS115	24	0,030	0,107	C525EcoNot	GS115	100	0,038	0,094
		48	0,031	0,085			139	0,038	0,095
		72	0,030	0,080			163	0,041	0,086
		96	0,030	0,066			187	0,048	0,073
		120	0,032	0,066			24	0,035	0,058
	KM71H	114	0,029	0,096		KM71H	30	0,039	0,062
		120	0,028	0,085			57	0,045	0,122
		138	0,034	0,157			100	0,049	0,150
		144	0,032	0,120			139	0,045	0,366
		162	0,028	0,090			48	0,042	0,137
pPICZaA	GS115	54	0,036	0,120	C209EcoNot	GS115	72	0,041	0,184
		81	0,035	0,069			96	0,045	0,140
		124	0,036	0,075			120	0,043	0,233
		148	0,035	0,079			24	0,031	0,061
		172	0,036	0,071			42	0,033	0,022
	KM71H	24	0,035	0,032		KM71H	48	0,033	0,020
		42	0,035	0,034			66	0,037	0,024
		48	0,034	0,035			72	0,043	0,027
		66	0,035	0,029			90	0,033	0,022
		72	0,034	0,027			96	0,038	0,025
C525XhoNot	GS115	90	0,035	0,024	C203EcoNot	GS115	24	0,031	0,133
		96	0,036	0,024			48	0,031	0,095
		24	0,032	1,100			72	0,031	0,074
		48	0,036	0,195			96	0,034	0,096
		72	0,034	0,103			120	0,027	0,085
	KM71H	96	0,035	0,117		KM71H	24	0,029	0,036
		120	0,034	0,093			48	0,031	0,039
		80	0,033	0,139			72	0,031	0,055
		114	0,032	0,104			96	0,034	0,070
		120	0,032	0,098					
	138	0,032	0,109						

A expressão de todas as isoformas de quitinase foi conseguida, apesar de a actividade enzimática parecer ser baixa para algumas recombinantes. No entanto, é de salientar que estes são ainda resultados preliminares, foram já parcialmente

reconfirmados, mas é necessário efectuar mais ensaios de indução para maior robustez dos valores.

Pelos resultados do Quadro 13, a proteína recombinante derivada da construção C525XhoNot foi expressa em maior quantidade (1,10U/mg de proteína total) e as restantes expressas em quantidades mais baixas: C209EcoNot (0,23U/mg de proteína total), C525EcoNot (0,37U/mg de proteína total) e C203EcoNot (0,13U/mg de proteína total). Com os dados anteriores traçou-se um gráfico para cada estirpe que mostra como a actividade enzimática por mg de proteína total varia com o tempo, Figuras 18 e 19.

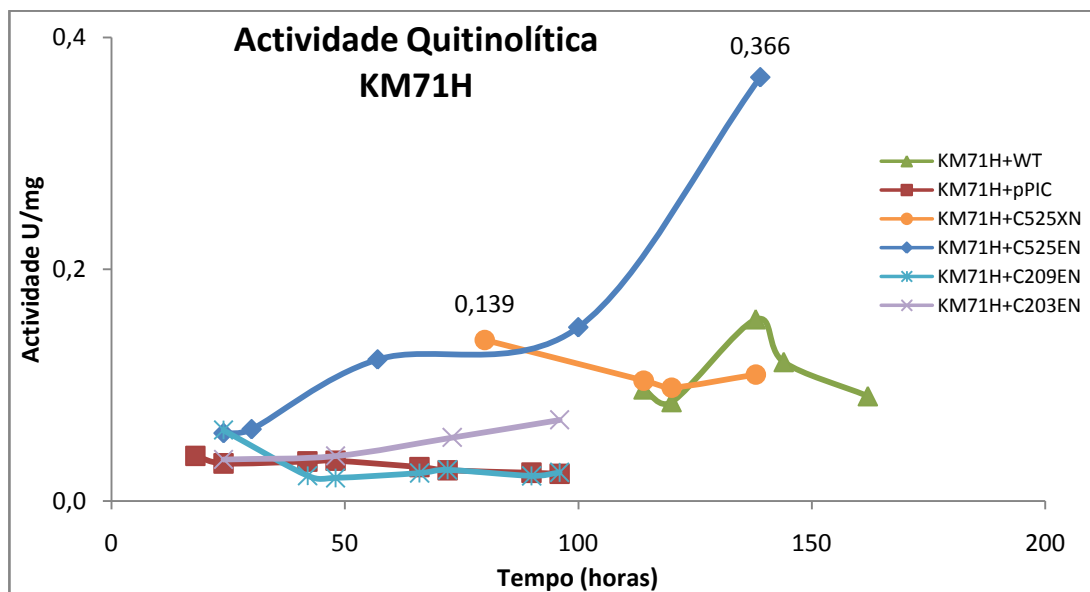


Figura 18 – Variação da actividade quitinolítica ao longo do tempo para as amostras seleccionadas das quatro construções na estirpe KM71H e para a estirpe WT e transfectada com vector nulo.

Pela análise da Figura 18, a recombinante C525EcoNot é a que mais actividade quitinolítica apresenta, quando comparada com a levedura WT, seguida da recombinante C525XhoNot. A recombinante C209EcoNot apresenta uma actividade reduzida e comparável com a da transfectação com o vector nulo, assim como a recombinante C203EcoNot que após as 50 horas aumenta um pouco.

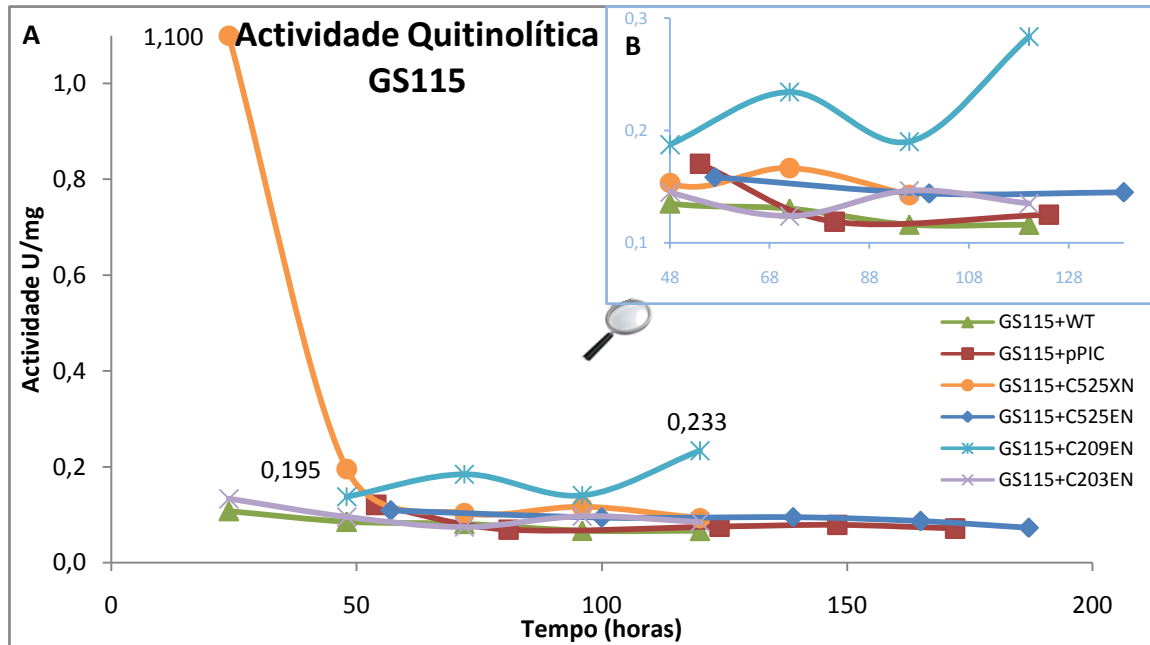


Figura 19 - A) Variação da actividade quitinolítica ao longo do tempo para as amostras seleccionadas das quatro construções no estirpe GS115 e para a estirpe WT e transfectada com vector nulo. B) Pormenor do gráfico de A) entre as 48 e as 139 horas.

Analisando a Figura 19, a recombinante que maior actividade quitinolítica apresenta é a C525XhoNot, quando comparado com a actividade enzimática apresentada pelas estirpes WT e transfectada com o vector nulo. A recombinante que mais actividade enzimática apresenta a seguir à C525XhoNot é a C209EcoNot, sendo que as recombinantes C525EcoNot e C203EcoNot têm uma actividade enzimática baixa, quando comparada com a estirpe WT.

O Quadro 14 resume os resultados obtidos no Quadro 13 e nas Figura 18 e 19, para as recombinantes com maior actividade enzimática no ponto em que essa actividade é máxima. A actividade quitinolítica que as estirpes controlo apresentam deve-se exclusivamente ao(s) quitinase(s) endógeno(s) que a levedura *Pichia pastoris* tem, devendo-se as actividades medidas nos extractos das restantes construções não apenas aos quitinases recombinantes mas também à actividade do(s) quitinase(s) endógeno(s).

Quadro 14 – Valores de actividade quitinolítica em U/mg de proteína total para as amostras que apresentam maior actividade e controlos respectivos.

Estirpe	Construção	Horas _{pico}	U/mg
GS115	WT	24	0,107
	pPICZaA	54	0,120
	C525XhoNot	24	1,100
	C209EcoNot	120	0,233
	C203EcoNot	24	0,133
KM71H	WT	138	0,157
	pPICZaA	48	0,035
	C525EcoNot	139	0,366

É de notar (Quadro 14) que a actividade quitinolítica máxima do sobrenadante de cada recombinante é superior à das estirpes controlo (GS115 e KM71H, WT e transfectadas com pPICZaA nulo), excepto para a recombinante C203EcoNot, que apresenta uma actividade enzimática semelhante à do controlo transformado com o

vector nulo. Este resultado surge de forma algo esperada pois a construção C203 não compreende a região codificante completa de um gene de quitinase, apesar de compreender a região do centro activo do enzima, faltando-lhe cerca de 200pb na extremidade 5'.

Outro factor interessante é a alteração na expressão de quitinases endógenos quando a levedura é transfectada com o vector nulo. Para a estirpe GS115, esta alteração é quase não significativa mas, para a estirpe KM71H, nota-se uma redução drástica na expressão de quitinases endógenos, devido talvez ao facto de a inserção ter ocorrido num local do DNA da levedura essencial à expressão deste enzima. Seleccionaram-se os melhores momentos para cada isoforma para repetir os ensaios de indução para posterior purificação da proteína.

Noutros trabalhos já realizados, a expressão de quitinases de *Beauveria bassiana*, na estirpe GS115, mostrou uma actividade máxima às 48 horas de 1U/mL, aproximadamente (Fan 2007). Tentando converter os resultados obtidos em unidades comparáveis (Quadro 15) verifica-se que a actividade enzimática das proteínas recombinantes expressas neste trabalho é baixa.

Quadro 15 – Valores de U/mL de sobrenadante e mg de GlnNAc/mL de sobrenadante para as amostras que apresentam maior actividade e controlos respectivos.

Estirpe	Construção	Horas _{pico}	U/mL	mg(GlnNAc)/mL
GS115	WT	24	0,298	0,164
	pPICZαA	54	0,360	0,198
	C525XhoNot	24	0,322	0,177
	C209EcoNot	120	0,429	0,236
	C203EcoNot	24	0,306	0,168
KM71H	WT	138	0,342	0,188
	pPICZαA	48	0,337	0,185
	C525EcoNot	139	0,448	0,246

A baixa expressão destas proteínas recombinantes pode dever-se a vários factores, entre os quais a inserção de apenas 1 cópia de DNA recombinante em cada estirpe (Figura 15). Se a inserção de mais cópias tivesse ocorrido, seria de esperar um aumento da expressão dos quitinases. O local em que a cópia é inserida também pode interferir com a melhor ou pior expressão da proteína. Apesar de a integração ser específica, se a molécula recombinante for inserida, por algum erro, num local vital ao desenvolvimento da levedura, pode não se dar a expressão conveniente.

Por outro lado, e apesar de se ter realizado o estudo em estirpes com génotipos diferentes, as estirpes escolhidas podem não ser as mais indicadas para expressar as proteínas em causa, podendo haver outras que melhor expressem estas proteínas. Das duas estirpes utilizadas, a que parece expressar melhor a proteína é a GS115 já que três das quatro proteínas recombinantes expressas apresentam maior actividade enzimática quando clonadas nesta estirpe e apenas uma quando clonada na estirpe KM71H.

As condições de crescimento podem, também, ter afectado, de algum modo, a expressão das proteínas, apesar de terem resultado como as mais indicadas para a expressão, neste caso.

Os diferentes sinais de secreção inseridos nas proteínas podem ajudar a perceber as diferenças na expressão das mesmas. Apesar de o factor- α de *S. cerevisiae* já ter demonstrado inúmeros sucessos, não parece ter proporcionado uma melhor expressão neste caso. Ou seja, a isoforma C525 quando expressa com o sinal nativo de secreção é a que mostra maior actividade, quando clonada em GS115.

Infelizmente não é ainda possível inferir sobre se a isoforma C525 expressa uma proteína com maior actividade quitinolítica do que a isoforma C209, pois não só não foi possível obter colónias recombinantes da construção da isoforma C209 com o sinal nativo de secreção, o que permitiria perceber se a maior expressão se deve exclusivamente ao sinal de secreção ou também à isoforma, como a análise da actividade do sobrenadante está directamente dependente do transformante. Será necessário purificar a proteína (trabalho já em curso) para, então, poder avaliar o poder quitinolítico de cada isoforma. Apesar disso, comparando as proteínas recombinantes com o factor- α , nota-se que a que parece ter maior U/mg de proteína total é a isoforma C525, sugerindo que é esta a isoforma com maior potencial quitinolítico.

Por último, um factor que poderia influenciar a baixa expressão da proteína é a diferença de glicosilação entre a levedura e o hospedeiro natural. Os padrões de glicosilação das quatro recombinantes foram analisados quanto aos domínios Asn-X-Ser/Thr (*NetNGlyc 1.0 Server*) e não foram detectados quaisquer aminoácidos passíveis de sofrer N-glicosilação. Relativamente à O-glicosilação, e como mostra o Anexo 7, quando analisado o conteúdo em Ser e Thr (*NetOGlyc 3.1 Server*), encontraram-se resíduos passíveis de serem O-glicosilados. Uma falha neste processo pode impedir o correcto *folding* das proteínas em questão e originar uma actividade enzimática baixa. É ainda de referir que a glicosilação na levedura não tem que ocorrer nos mesmos locais do hospedeiro natural, o que pode levar a diferenças de expressão, e que sequências não glicosiladas no hospedeiro original podem ser glicosiladas na levedura.

No entanto, outros trabalhos apontam também para baixas actividades de genes de quitinases em expressão heteróloga. A expressão de um gene de quitinase de *Oryza sativa* L., na estirpe GS115, mostrou uma concentração de enzima máxima às 36 horas de 4 μ g/mL (Truong 2003), valores ainda mais baixos do que os obtidos neste trabalho (Quadro 15).

A análise por SDS-PAGE realizou-se em condições desnaturantes, e com géis de 5%T (concentração) e 12%T (resolvente), corados com nitrato de prata. As amostras utilizadas foram os sobrenadantes dos pontos escolhidos para análise. Os resultados obtidos não foram satisfatórios já que não ocorreu a identificação de nenhuma banda passível de corresponder à proteína em estudo, com cerca de 31,6kDa. Assim, tentaram analisar-se os precipitados, também referentes aos pontos escolhidos, mas mais uma vez não se obtiveram resultados satisfatórios. Pensou-se que este facto poderia estar relacionado com baixa concentração da proteína no meio e procedeu-se a uma liofilização das amostras e a novo SDS-PAGE mas não se obtiveram resultados, mais uma vez. Provavelmente, a quantidade de amostra liofilizada (500 μ L) não foi suficiente. Optou-se por crescer a cultura em larga escala e purificar primeiro a proteína, com recurso ao kit *ProBond™ Purification System* da Invitrogen (trabalho em curso), para depois realizar nova análise SDS-PAGE.

Assim, em conclusão, os transformantes que melhor expressaram a proteína introduzida são o que transporta a construção C525EcoNot na estirpe KM71H, o que transporta a C525XhoNot e o que transporta a C209EcoNot ambas na estirpe GS115, e destas o que se revela melhor é a recombinante C525XhoNot (Figura 20).

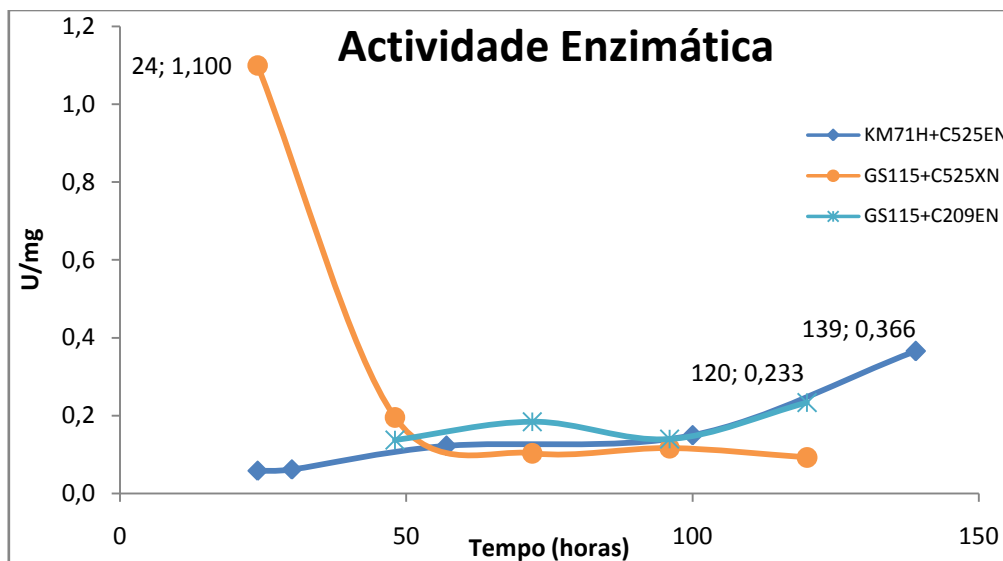


Figura 20 – Variação da actividade enzimática ao longo do tempo para as construções seleccionadas das duas estirpes GS115 e KM71H.

É necessário ter em conta que estes resultados são apenas preliminares e que é necessário realizar replicados destas determinações experimentais.

As construções escolhidas serão usadas para a produção em larga escala das isoformas de quitinase correspondentes, para depois proceder à purificação das proteínas e avaliação do potencial antifúngico que possam ter, em relação a fungos fitopatogénicos de macieira e pereira.

CONCLUSÃO E PERSPECTIVAS

Este trabalho visou a clonagem de sequências que codificam quitinases ácidos de classe III de macieira 'Royal Gala', e a expressão das proteínas correspondentes, em *Pichia pastoris*.

Para tal, clonaram-se três genes que codificam os quitinases (C525, C209 e C203), com sequências secretoras diferentes (num total de cinco construções distintas), em duas estirpes de *Pichia pastoris* (GS115 e KM71H), e avaliou-se a actividade quitinolítica das culturas contendo estas isoformas. Clonaram-se com sucesso quatro das cinco construções preparadas, a C525, C209 e C203 com o sinal de secreção *α -factor* de *S. cerevisiae* e a isoforma C525 com o sinal nativo de secreção, nas duas estirpes. Não foi possível clonar a isoforma C209 com o sinal nativo de secreção, apesar de a transfecção de ambas as estirpes de levedura ter dado origem a colónias. Uma análise de *Southern* sugere a presença de apenas uma cópia de cada construção recombinante correspondente.

Fazendo uso das diferentes sensibilidades ao metanol que as duas estirpes apresentam, induziu-se a expressão das proteínas clonadas e realizaram-se testes enzimáticos com o intuito de perceber quais as construções que permitiram registar maior actividade quitinolítica nas respectivas culturas.

Verificou-se que culturas com os genes C525 ou C209 introduzidos apresentaram maior actividade enzimática, quando comparadas com as estirpes correspondentes WT e transfectada com o vector nulo, evidenciando a expressão das proteínas heterólogas. Escolheram-se as três construções cujos transformantes apresentaram maior actividade quitinolítica, em relação à proteína total expressa. São elas as isoformas C525 com o sinal de secreção nativo e C209 com o sinal de secreção *α -factor*, clonadas na estirpe GS115 (1,100U/mg de proteína às 24h e 0,233U/mg de proteína às 120h, respectivamente), e a isoforma C525 com o *α -factor*, clonada na estirpe KM71H (0,366U/mg de proteína às 139h). Estes valores são intermédios em relação a valores encontrados na literatura para expressão heteróloga de quitinases em *Pichia pastoris*. Das três construções, a primeira foi a que expressou melhor e mais rapidamente a proteína.

Estes resultados sugerem, assim, que a clonagem heteróloga em *Pichia pastoris* é um método adequado para expressar estas proteínas. Não obstante, a análise de SDS-PAGE dos precipitados e dos secretomas respeitantes aos diferentes momentos não se revelou conclusiva, provavelmente devido à baixa concentração proteica das amostras.

O projecto em que se insere este trabalho visa a avaliação do potencial antifúngico dos quitinases expressos. Os transformantes das três construções seleccionadas, e as condições de crescimento optimizadas, serão utilizados em

culturas de larga escala para produzir e purificar grandes quantidades das proteínas correspondentes. Estas proteínas puras serão caracterizadas (por SDS-PAGE e teste enzimático) e seguidamente testar-se-á o seu poder antifúngico em macieira e pereira, ou seja, determinar-se-á qual apresenta maior actividade quitinolítica em relação à quitina de fungos fitopatogénicos destas duas fruteiras.

BIBLIOGRAFIA

- Ano, A. (2002). "Characterization of a Class III Chitinase from *Vitis vinifera* cv. Koshu " Journal of Bioscience and Bioengineering **95**(6): 645-647.
- Arruda, A. (2008). Utilização do promotor do gene PGK1 de *Pichia pastoris* para a expressão heteróloga, Dissertação de Mestrado, Universidade de Brasília, Brasília.
- Balbás, P. (2001). "Understanding the Art of Producing Protein and Nonprotein Molecules in *Escherichia coli*." Molecular Biotechnology **19**: 251-267.
- Cereghino, J. L. (2000). "Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*." FEMS Microbiology Reviews **24**: 45-66.
- Collinge, D. (1993). "Plant chitinases." The Plant Journal **3**(1): 31-40.
- Datta, S. K. a. M., S. (1999). Pathogenesis-related proteins in plants, CRC Press LLC.
- Dessai, P. N. (2010). "Productin of heterologous proteins in plants: Strategies for optimal expression." Biotechnology Advances **28**: 427-435.
- Domínguez, A. (1998). "Non-conventional yeasts as hosts for heterologous protein production." Internatl Microbiol **1**: 131-142.
- Fan, Y. (2007). "Expression of a *Beauveria bassiana* chitinase (Bbchit1) in *Escherichia coli* and *Pichia pastoris*." Protein Expression and Purification **56**: 93-99.
- Fukamizo, T. (2003). "Plant Chitinases: Structure-Function Relationships and Their Physiology." Foods Food Ingredients **208**(8): 631-638.
- Gomord, V. (2010). "Plant-specific glycosylation patterns in the context of therapeutic protein production." Plant Biotechnology Journal **8**: 564-587.
- Higgins, D. R. C., J. M. (1998). Methods in Molecular Biology Pichia Protocols, Humana Press.
- Huynh, Q. K. (1992). "Antifungal Proteins from Plants." The Journal of Biological Chemistry **267**(10): 6635-6640.
- Invitrogen *Pichia pastoris* Expression System.
- Invitrogen (2005). EasySelect™ *Pichia* Expression Kit - User Manual, Invitrogen. **Version H**: 1-74.
- IUBMB. from <http://www.chem.qmul.ac.uk/iubmb/>.
- Kasprzewska, A. (2003). "Plant Chitinases - Regulation and function." Cellular & Molecular Biology Letters **8**: 809-824.
- Koga, D. (1999). "Biochemistry of chitinases." Chitin and Chitinases, **EXS 87**: 111-123.
- Kramer, K. J. (1997). "Insect Chitinases: Molecular Biology and Potencial Use as Biopesticides." Insect Biochem. Molec. Biol **27**(11): 887-900.
- Nagasaki, H. (1997). "Rice Class III Chitinase Homologues Isolated by Random Cloning of Rice cDNAs." DNA Research **4**: 379-385.
- Nicholl, D. S. T. (2002). An Introduction to Genetic Engineering, Cambridge University Press.
- Rai, M. (2001). "Expression systems for production of heterologous proteins." Current Science **80**(9): 1121-1128.
- Singh, A. (2007). "Heterologous expression of new antifungal chitinase from wheat." Protein Expression and Purification **56**: 100-109.
- Truong, N.-H. (2003). "Structure, Heterologou Expression, and Properties of Rice (*Oryza sativa* L.) Family 19 Chitinases." Biosci. Biotechnol. Biochem. **67**(5): 1063-1070.
- Uchibori, T. (2000). "Purification and characterization of elicitor-induced chitinase from grape berries." **47**(2): 163-168.
- Yanai, K. (1992). "Purification of Two Chitinases from *Rhizopus oligosporus* and Isolation and Sequencing of the Encoding Genes." **174**(22): 7398-7406.

ANEXOS

ANEXO 1: Composição dos Meios de Cultura

Todos os reagentes que compõem os meios foram dissolvidos em água destilada e autoclavados a 120°C, 1,4 bar, por 20 minutos, salvo quando indicação do contrário. Os antibióticos foram adicionados após arrefecimento até 50°C, aproximadamente. Os meios de cultura utilizados foram:

- LB (*Luria-Bertani*) com canamicina (50 µg/mL):

<u>Para 1000mL - líquido</u>	
Tryptona – 1%	10g
Extracto de levedura – 0,5%	5g
Cloreto de sódio – 1%	10g
Água destilada	950mL
Acertar o pH a 7,0	
Para LB sólido	15g de agar
Canamicina 50mg/mL	1000µL

- LB_{lowsalt} (*Luria-Bertani Low Salt*) com zeocina (25 µg/mL):

<u>Para 1000mL - líquido</u>	
Tryptona – 1%	10g
Extracto de levedura – 0,5%	5g
Cloreto de sódio – 0,5%	5g
Água destilada	950mL
Acertar o pH a 7,5	
Para LB_{is} sólido	15g de agar
Zeocina 100mg/mL	250µL

- LB_{comp} (adaptado de *Luria-Bertani*):

<u>Para 100mL - líquido</u>	
PEG 3350 – 20%	20g
DMSO– 10%	10mL
Cloreto de magnésio – 50mM	1,01g
LB Caldo – 0,5x	50mL
Água destilada	40mL

- SOC:

<u>Para 1000mL - líquido</u>	
Tryptona – 2%	20g
Extracto de levedura – 0,5%	5g
Cloreto de sódio	0,58g
Cloreto de potássio	0,19g
Cloreto de magnésio	2,03g
Sulfato de magnésio	3,96g
Água destilada	2,46g
Acertar o pH a 7,0	
Glucose 1M	20mL
Para SOB	Não adicionar glucose

- YPD com zeocina (100 µg/mL):

<u>Para 1000mL - líquido</u>	
Tryptona – 2%	20g
Extracto de levedura – 1%	10g
Água destilada	900mL
Dextrose 10x – 2%	100mL, após autoclavar
Para YPD sólido	20g de agar
Zeocina 100mg/mL	1mL

- YPDS com zeocina (100 µg/mL):
Para 1000mL - líquido
 Triptona – 2% 20g
 Extracto de levedura – 1% 10g
 Sorbitol - 1M 182,2g
 Água destilada 900mL
 Dextrose 10x – 2% 100mL, após autoclavar
Para YPD sólido 20g de agar
Zeocina 100mg/mL 1mL

- MD e MDH (Meio Mínimo com Dextrose + Histidina):
Para 1000mL – líquido: Autoclavar a água destilada e adicionar os reagentes aos 60°C
 YNB – 1,34% 100mL YNB 10x
 Biotina – 4x10⁻⁵% 2mL Biotina 500x
 Dextrose – 2% 100mL Dextrose 10x
 Água destilada 800mL
Para MDH 10mL Histidina 100x
Para meio sólido 15g de agar à água a autoclavar

- MM e MMH (Meio Mínimo com Metanol + Histidina):
Para 1000mL – líquido: Autoclavar a água destilada e adicionar os reagentes aos 60°C
 YNB – 1,34% 100mL YNB 10x
 Biotina – 4x10⁻⁵% 2mL Biotina 500x
 Metanol – 0,5% 100mL Metanol 10x
 Água destilada 800mL
Para MMH 10mL Histidina 100x
Para meio sólido 15g de agar à água a autoclavar

- BMGY e BMMY (Meio Tamponado Complexo com Glicerol/Metanol):
Para 1000mL – líquido: Autoclavar o extracto de levedura e a peptona dissolvidos em água e adicionar os reagentes aos 60°C
 YNB – 1,34% 2mL Biotina 500x
 Biotina – 4x10⁻⁵% 10g
 Extracto de levedura – 1% 20g
 Peptona – 2% 100mL de um tampão 1M
 Fosfato potássio pH 6 – 100mM 700mL
 Água destilada 100mL Glicerol 10x - 1% glicerol
Para BMGY 100mL Metanol10x - 0,5% metanol
Para BMMY 100mL YNB 10x

ANEXO 2: Composição de Soluções e Tampões

Todas as soluções e tampões foram preparados segundo *Sambrooks et al.* (1989), excepto quando indicado em contrário.

Quadro 16 – Composição e concentrações das soluções utilizadas ao longo do trabalho.

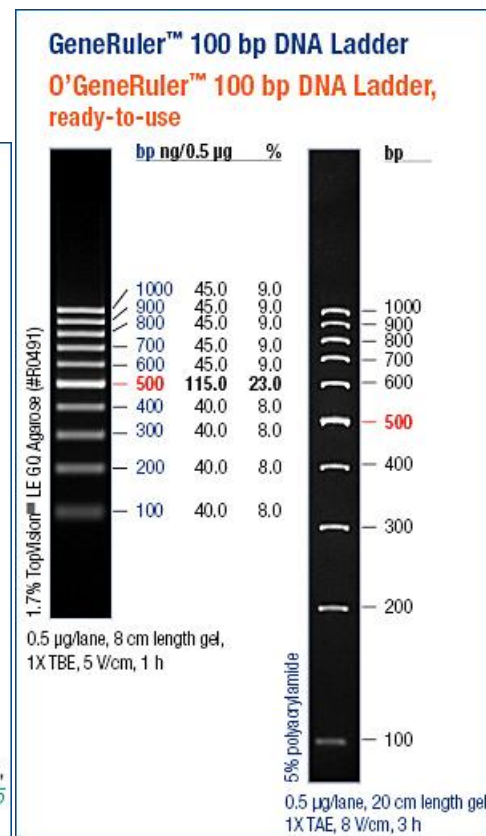
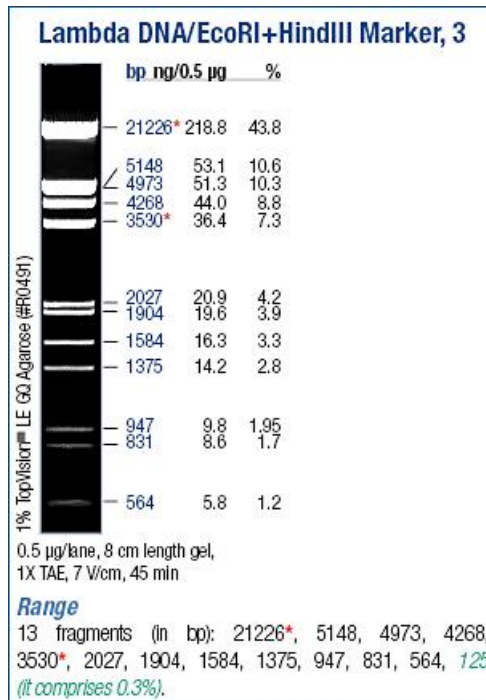
Solução	Composição	Concentração final	Utilidade
Tampão TAE 1x (tampão Tris-acetato)	Tris-Acetato	40mM	
	EDTA	1mM	
Tampão de aplicação	Tris pH 7,5	0,1M	Análise por electroforese em gel de agarose
	Glicerol	2%	
	EDTA	0,1M	
	Azul bromofenol	20%	
Etanol 70%	Etanol	70%	
	Água destilada	30%	
Acetato de Sódio 3M, pH5,2	Acetato Sódio hidratado	4,97M	Precipitação e purificação de DNA
	Água destilada	Acertar volume	
	Ácido acético glacial	Acertar pH	
Acetato de Amónio 7,5M	Acetato de amónio	7,5M	
	Água destilada	Acertar volume	
Fenol: Clorofórmio:	Fenol	25%	
Álcool isoamílico	Clorofórmio	24%	
Conservada a 4°C, tóxico	Álcool isoamílico	1%	
dNTPs, pH 7	Adenina		Reacção de PCR
	Guanina		
	Citosina	100mM	
	Timina		
Conservada a -20°C			
Agarose 2%	Agarose (Invitrogen)	2%	Géis para electroforese
	TAE 50x	2%	
	Greensafe	0,025µL/mL	
Géis conservados a 4°C	Água destilada	Acertar volume	
Agarose 1,2%	Agarose (Invitrogen)	1,2%	
	TAE 50x	2%	
	Greensafe	0,025µL/mL	
Géis conservados a 4°C	Água destilada	Acertar volume	
Canamicina*	Canamicina	50mg/mL	Seleção de microrganismos
Conservada a -20°C	Água destilada	Acertar volume	
Zeocina*	Zeocina	100mg/ml	
Conservada a -20°C	Água destilada	Acertar volume	
X-gal*	X-gal	50 mg/mL	*soluções filtradas para esterilização
	N,N - Dimetilformamida	2 mL	
Conservada a -20°C			

SDS 10%	SDS	10%	Vários fins
	Água destilada	Acertar volume	
Solução I	Glucose	50mM	
pH 8,0	Tis-HCl	25mM	
Conservada a 4°C	EDTA	10mM	
Solução II	NaOH	0,2N	Extração de DNA plasmídico em pequena escala
Preparada na hora	SDS	1%	
	Água destilada	Acertar volume	
Solução III	Acetato de potássio	5M	
Conservada a 4°C	Ácido acético glacial	3M	
	Água destilada	Acertar volume	
YNB 10x*	YNB sem sulfato de amónia nem aminoácidos	13,4%	Preparação de meios para crescimento de leveduras
	Sulfato de amónia	100g-1L	
	Água destilada	Acertar volume	
Biotina 500x*	Biotina	0,02%	
	Água destilada	Acertar volume	
Histidina 100x*	Histidina	0,4%	
	Água destilada	Acertar volume	
Dextrose 10x	Dextrose	20%	
	Água destilada	Acertar volume	
Metanol 10x*	Metanol	5%	*soluções filtradas para esterilização
	Água destilada	Acertar volume	
Glicerol 10x	Glicerol	10%	
	Água destilada	Acertar volume	
Tampão Fosfato de Potássio	K ₂ HPO ₄	1M	
1M, pH6,0	KH ₂ PO ₄	1M	
	Água destilada	Acertar volume	
Solução SCED	Sorbitol	1M	Extração de DNA genómico de levedura em pequena escala
pH 7,5	Citrato	10mM	
	EDTA	10mM	
Preparada no momento Segundo manual do kit	DTT	10mM	
	Água destilada	Acertar volume	
SDS 1%	SDS	1%	
	Água destilada	Acertar volume	
Acetato de potássio	Acetato de potássio	5M	
pH 8,9	Água destilada	Acertar volume	
Tampão TE	Tris-HCl (Acros)	10mM	
pH 7,4 (tampão Tris-EDTA)	EDTA (Sigma)	1mM	
	Água destilada	Acertar volume	

Tampão de breaking pH 7,4	Fosfato de Sódio	50mM	Preparação dos <i>pellets</i> para SDS- PAGE
	PMSF	1mM	
	EDTA	1mM	
	Glicerol	5%	
Segundo manual do kit	Água destilada	Acertar volume	
Tampão SDS-PAGE 4x	Tris HCl pH 6,8	200mM	Preparação das amostras para SDS-PAGE
	DTT	400mM	
	SDS	8%	
	Azul de bromofenol	0,4%	
	Glicerol	40%	
	Água destilada	Acertar volume	
Tampão Tris-glicina 5x	Tris	250mM	Tampão de corrida de SDS- PAGE
	Glicina	250mM	
	SDS	0,1%	
	Água destilada	Acertar volume	
Tampão Tris 1,5M, pH 8,8	Tris-base	1,5M	Preparação dos géis para SDS- PAGE
	Água destilada	Acertar volume	
Tampão Tris 1M, pH 6,8	Tris-HCL	1M	
	Água destilada	Acertar volume	
PSA	Persulfato de amónio	10%	
	Água destilada	Acertar volume	
Nitrato de prata	Nitrato de prata	0,1%	
	Água destilada	Acertar volume	
Etanol:Ác. Acético:H₂O	Etanol	30%	Coloração e revelação dos géis de SDS- PAGE
	Ácido acético	10%	
	Água destilada	60%	
Solução de Revelação	Na ₂ CO ₃	2,5%	
	Formaldeído	0,02%	
	Água destilada	Acertar volume	
Despurinante	HCl	0,25M	
	Água destilada	Acertar volume	
Tampão Desnaturação	NaCl	1,5M	Southern blot
	NaOH	0,5M	
	Água destilada	Acertar volume	
Tampão Neutralização pH 7,5	Tris.HCl	1M	
	NaCl	1,5M	
	Água destilada	Acertar volume	
SSC 20x , pH 7,0	NaCl	0,3mM	
	Citrato sódio	0,03mM	
	NaOH	Acertar pH	
	Água destilada	Acertar volume	

Tampão HB1	DH	10x	<i>Southernblot</i> (cont.)
	SSC	3x	
	SDS	0,1%	
	<i>Salmon sperm</i>	100µg/mL	
Tampão Bloqueio	Reagente de Blocking	10%	
	Tampão ácido maleico	1x	
Tampão ácido maleico 10x, pH 7,5	Ácido maleico	1M	
	NaCl	1,5M	
	Água destilada	Acertar volume	
Tampão Lavagem	Tampão ácido maleico	1x	
	Tween 20	0,3%	
	Água destilada	Acertar volume	
Tampão Detecção pH 9,5	Tris.base	100mM	
	NaCl	100mM	
	Água destilada	Acertar volume	
BSA	Albumina de soro de bovino	1,4 mg/mL	Quantificação proteica
	Água destilada	Acertar volume	
Solução substrato	Glicolquitina	0,1%	
	Acetato de sódio, pH 4	50mM	
Ferricianeto de potássio	Ferricianeto de potássio	0,1%	Actividade enzimática
	Carbonato de sódio	0,5M	
	Água destilada	Acertar volume	
p-Dimetilamino- -benzaldeído	Ácido acético glacial	Acertar volume	
	Ácido clorídrico	0,1%	
	p-DMAB	1%	

ANEXO 3: Padrão de Separação das Bandas após Electroforese dos Marcadores Utilizados



SIZE (bp)	ng/BAND
10000	100
7500	75
6000	60
5000	50
4000	40
3000	30
2500	25
2000	20
1400	14

ANEXO 4: Sequências Isoladas de DNA Genômico

- Sequência de DNA da isoforma C203

CCAGAAATCAAGTCATGCCAAGCCAAGGGGATTAAGTCATACTCTCCAGAGGAGGAGCTTCTGGGAGCT
 ACTCTCTAGCTTCAGCTGATGATGCAAGGCAAGTTGCAACTGACTTGTGGAACAACCTTCTGGGAGGGCA
 TTCTCGTCAAGCCGTTGGGAGCTGCAGTATTGGATGGAATTGACTTTGATATTGAAGGAGGGACTGAC
 CAACATTGGGATGACCTTGCTAAGTACCTATCTGGATATAGCAAGAGAGGCAAGAAAGTCTACTTGACCG
 CTGCCCCACAATGTCCCTTTCTGATGCTTGGATTGGAAATGCACCTTAAGACAGGCCTCTTTGACAATGT
 TTGGGTTTCTAGTTCTACAACAACCCCTCCCTGTCAGTACACTTCCGGCGACGTGGCCAATCTAGAGGACGCT
 TGAAGCAGTGGACTTCAGCCATCCCTGCACATAAGATTTTCTTGGGGTTGCCTGCTGCTCCTCAGGCTG
 CCGGTAGCGGTTTTATTCTGCTTCTGATCTCAACTCACAAGTCCCTTCCGGCTATTAATAAATTCGGCTAA
 GTATGGAGGTGTCATGCTTTGGTCCAAATATTATGATGATCTTGATGGATACAGCTCCTCCATCAAGAGT
 CATGTCTA

- Sequência de DNA da isoforma C209

ATGGCTTCAAAGTCCACAATGTCCCTAGCATTTGCTCTCTTTAGTGACACTAGTTCTGGCTTTGGGGGCTA
 ATGCCGGTGAATAGCTATATATTGGGGTCAAATGGGAATGAAGGCACGTTAGCGGAAACTTGTGCTAC
 AGGGAATTACCAATATGTAAACGTAGCTTTTCTCACCACCTTCGGCAATGGCCAAACCCCGCCATCAAC
 CTAGCCGGTCACTGTGTCCCAACGACCGATGAATGCACCAAGTTGAGCCAGAGATCAGGTCATGCCAAG
 CCAAGGGGATTAAGGTCATACTTTCCATTGGAGGAGCTTCTGGGAGCTACTCTTAACCTCAGCAGCTGA
 TGCAAGGCAAGTTGCAACTTACTTGTGGAACAACCTTCTGGGAGGGCATTGTCGTCGAGGCCATTGGGA
 GCTGCAGTATTGGATGGAATTGACTTTGATATTGAGGGAGGGACTGACCAATATTGGGATGACCTTGCTA
 GGTACCTATCTGGATATAGCAAGAGAGGCAAGAAAGTTTACTTGACTGCTGCCCCACAATGTCCCTTTCC
 TGATGCTTGGATTGGAAATGCACCTTAAGACAGGCCTCTTTGACAATGTTTGGGTTTCTACAACAAC
 CCTCCCTGTCAGTACACTTCCGGCGACGTGGCCAATCTGGAGGACGCTTGAAGCAGTGGACTTCAGCCA
 TCCCTGCACATAAGATTTTCTTGGGGTTGCCTGCTGCTCCTCAGGCTGCCGGTAGCGGTTTTATTCTGCT
 TTCTGATCTCAACTCACAAGTCCCTTTCGGCTATTAAAAGCTCGGCTAAGTATGGAGGTGTCATGCTTTGG
 TCCAAGTATTATGATGATCTTGATGGATACAGCTCCTCCATCAAGAGTTCATGTCTAA

- Sequência de DNA da isoforma C525

ATGGCTTCAAAGTCCACAGCAACGTTCCCTAGCATTTGCTCTCTTTAGTGACACTAGTTCTGGCTCTCGGGG
 CTAATGCCGGTGAATCGCAATATATTGGGGTTCAGAATGGTAATGAAGGCACATTAGCAGAAACATGTGC
 TTCAGGGAATTACCAATTTGTAAACGTAGCTTTTCTCACCACCTTCGGCAACGGCCAAACGCCGTGCAATC
 AACCTAGCCGGTCACTGTGACCCAACGACAGAAGAATGCACCAAATGAGCCAGAAATCAAGTCATGCC
 AAGCCAAGGGGATTAAGTTCATACTCTCCATAGGAGGAGCTTCTGGGAGCTACTCTTAACCTCAGCTGA
 TGATGCAAGGCAAGTTGCAACTTACCTGTGGAATAAATTTCTTGGGAGGGCAATCGTCGTCGAGGCCATTG
 GGAGCTGCCGTTTTGGATGGAATTGATTTTGGACATCGAGGGAGGGACTGACCAACATTGGGATGACCTAG
 CAAGGTACCTCTCTGGATATAGCAAGAGAGGCAAGAAAGTTTACTTGACTGCTGCCCCACAATGTCCCTT
 TCCTGATGCTTATGTTGAAATGCACCTTAAGACGGGCTCTTTGACAATGTTTGGGTTTCTAGTTCTACAAC
 AACCTCCCTGCCAGTACGCTTCTGGGGATGTGACCAACCTTGAAGACGCCGTGAAGCAGTGGACTTCAG
 CCATCCCTGCAGATAAGATTTTCTTGGGATTGCCTGCTGCACCTCAAGCTGCTGGTAGCGGGTTTTATTCC
 TGCTACTGATCTTAGCTCACAAGTCTTCCGGCTATTAATAAAGTTCCGGCTAAGTATGGAGGCGTCATGCTT
 TGGTCCAAGTATTATGATGATCTTGATGGATACAGCTCCTCCATCAAGAATGATGTCTAG

ANEXO 5: Sequência nucleotídica parcial dos vectores utilizados

A) pPICZaA

5' end of AOX1 mRNA
5' AOX1 priming site

811 AACCTTTTTT TTTATCATCA TTATTAGCTT ACTTTCATAA TTGCGACTGG TTCCAATTGA

871 CAAGCTTTTG ATTTTAAACGA CTTTAAACGA CAACTTGAGA AGATCAAAAA ACAACTAATT

931 ATTCGAAACG **ATG** AGA TTT CCT TCA ATT TTT ACT GCT GTT TTA TTC GCA GCA
Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala

983 TCC TCC GCA TTA GCT GCT CCA GTC AAC ACT ACA ACA GAA GAT GAA ACG GCA
Ser Ser Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala

α -factor signal sequence

1034 CAA ATT CCG GCT GAA GCT GTC ATC GGT TAC TCA GAT TTA GAA GGG GAT TTC
Gln Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe

1085 GAT GTT GCT GTT TTG CCA TTT TCC AAC AGC ACA AAT AAC GGG TTA TTG TTT
Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu Phe

Xho I*

1136 ATA AAT ACT ACT ATT GCC AGC ATT GCT GCT AAA GAA GAA GGG GTA TCT CTC
Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val Ser Leu

Kex2 signal cleavage EcoR I Pml I Sfi I BsmB I Asp718 I

1187 GAG AAA AGA GAG GCT GAA GC GAATTCAC GTGGCCCAG CCGGCCGTC TCGGATCGGT
Glu Lys Arg Glu Ala Glu Ala

Ste13 signal cleavage

Kpn I Xho I Sac II Not I Xba I c-myc epitope

1244 ACCTCGAGCC GCGGCGGCC GCCAGCTTTC TA GAA CAA AAA CTC ATC TCA GAA GAG
Glu Gln Lys Leu Ile Ser Glu Glu

polyhistidine tag

1299 GAT CTG AAT AGC GCC GTC GAC CAT CAT CAT CAT CAT CAT TGA GTTTGTAGCC
Asp Leu Asn Ser Ala Val Asp His His His His His His ***

1351 TTAGACATGA CTGTTCTCA GTTCAAGTTG GGCACCTIACG AGAAGACCGG TCTTGCTAGA

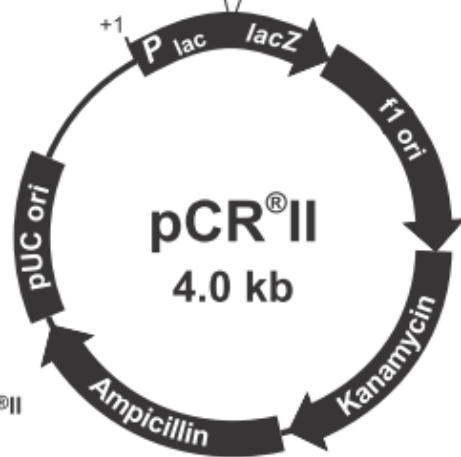
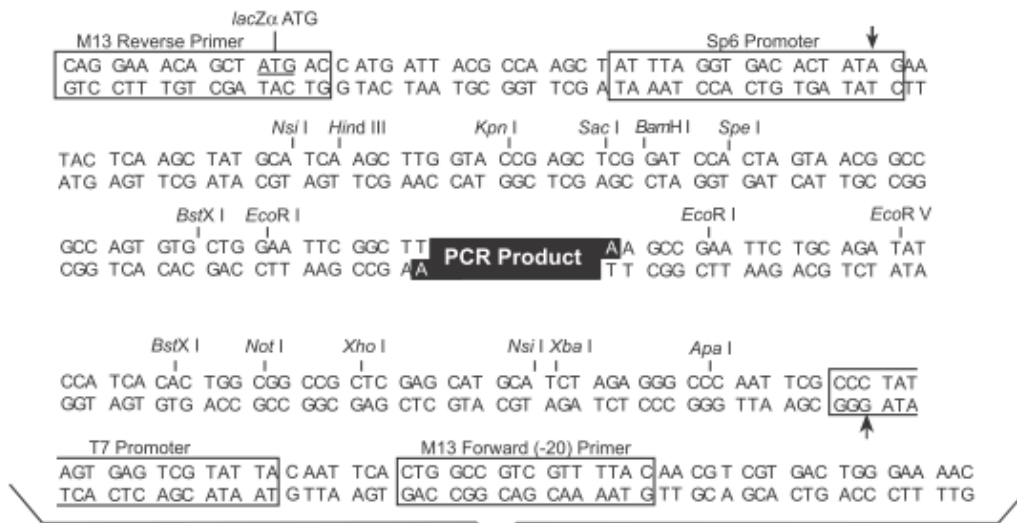
3' AOX1 priming site

1411 TTCTAATCAA GAGGATGTCA GAATGCCATT TGCCCTGAGAG ATGCAGGCTT CATTTTTGAT

3' polyadenylation site

1471 ACTTTTTTAT TTGTAACCTA TATAGTATAG GATTTTTTTT GTCATTTTGT TTCTTCTCGT

B) pCRII



Comments for pCR®II
3971 nucleotides

ANEXO 6: Sequências de DNA dos Fragmentos introduzidos em *Pichia pastoris*

- Sequência de DNA da isoforma C203EcoNot-PIC

```

1      10      20      30      40      50      60      70      80      90
CTAATTATTC GAAACGATGA GATTTCC TTC AATTTTACT GCTGTTTAT TCGCAGCATC CTCCGCATTA GCTGCTCCAG TCAACACTAC

100     110     120     130     140     150     160     170     180
AACAGAAGAT GAAACGGCAC AAATTCCGGC TGAAGCTGTC ATCGGTTACT CAGATTTAGA AGGGGATTTT GATGTTGCTG TTTTGCCATT

190     200     210     220     230     240     250     260     270
TTCCAACAGC ACAAATAACG GGTATTGTTT TATAAATACT ACTATTGCCA GCATTGCTGC TAAAGAGAA GGGGTATCTC TCGAGAAAAG

280     290     300     310     320     330     340     350     360
AGAGGCTGAA GCTGAATTC CAGAAATCAA GTCATGCCAA GCCAAGGGAA TTAAGTCAT ACTCTCCAGA GGAGGAGCTT CTGGGAGCTA
C203Eco-F - Primer

370     380     390     400     410     420     430     440     450
CTCTCTAGCT TCAGCTGATG ATGCAAGGCÁ AGTTGCAACT GACTTGTGGÁ ACAACTTCTT GGGAGGGCAT TCCTCGTCAA GGCCGTTGGG

460     470     480     490     500     510     520     530     540
AGCTGCAGTA TTGGATGGAÁ TTGACTTTGÁ TATTGAAGGA GGGACTGACC AACATTGGGA TGACCTTGTCT AAGTACCTAT CTGGATATAG

550     560     570     580     590     600     610     620     630
CAAGAGAGGC AAGAAAGTCT ACTTGACCCT TGCCCCACAÁ TGTCCTTTC CTGATGCTTG GATTGGAATÁ GCACTTAAGA CAGGCCCTCT

640     650     660     670     680     690     700     710     720
TGACAAATGT TGGGTTTAGT TCTACAACAÁ CCCCCTCCGT CAGTACACTT CCGGCACAGT GGCCAATCTÁ GAGGACGCTT GGAAGCAGTT

730     740     750     760     770     780     790     800     810
GACTTCAGCC ATCCCTGCAC ATAAGATTTT CTTGGGGTTG CTTGCTGCTC CTCAGGCTGC CGGTAGCGGT TTTATTCCCTG CTTCTGATCT

820     830     840     850     860     870     880     890     900
CAACTCACAÁ GTCCCTCCGG CTATTAAAAÁ TTCGGCTAAG TATGGAGGTG TCATGCTTTG GTCCAAAATÁ TATGATGATC TTGATGGATA

910     920     930     940     950     960     970     980     990
CAGCTCCTCC ATCAAGAGTC ATGTCGGGGC CGCCAGCTTT CTAGAACAÁ AACTCATCTC AGAAGAGGAT CTGAATAGCG CCGTCGACCÁ
C209Not-R - Primer

1,000   1,010   1,020   1,030   1,039
TCATCATCAT CATCATGAG TTTGTAGCCT TAGACATGAC TGTTCCTCA

```

- Sequência de DNA da isoforma C209EcoNot-PIC

```

1      10      20      30      40      50      60      70      80      90      100
CAAAAAACA CTAATTATTC GAAACGATGA GATTTCC TTC AATTTTACT GCTGTTTAT TCGCAGCATC CTCCGCATTA GCTGCTCCAG TCAACACTAC

110     120     130     140     150     160     170     180     190     200
AACAGAAGAT GAAACGGCAC AAATTCCGGC TGAAGCTGTC ATCGGTTACT CAGATTTAGA AGGGGATTTT GATGTTGCTG TTTTGCCATT TTCCAACAGC

210     220     230     240     250     260     270     280     290     300
ACAAATAACG GGTATTGTTT TATAAATACT ACTATTGCCA GCATTGCTGC TAAAGAGAA GGGGTATCTC TCGAGAAAAG AGAGGCTGAA GCTGAATTCG
C209Eco-F - P...

310     320     330     340     350     360     370     380     390     400
GTGGAATAGC TATATATTGG GGTCAAAATG GGAATGAAG CACGTTAGCG GAAACTTGTG CTACAGGGAÁ TTACCAATAT GTAACGATAG CTTTCTCAC
C209Eco-F - Primer

410     420     430     440     450     460     470     480     490     500
CACCTTCGGC AATGGCCAÁ CCCCCCCAAT CAACCTAGCC GGTCACTGTG TCCCAACGAC CGATGAAATG ACCAAGTTGA GCCCAGAGAT CAGGTCATGC

510     520     530     540     550     560     570     580     590     600
CAAGCCAAGG GGATTAAGGT CATACTTCC ATTGGAGGÁ CTTCTGGGÁ CTACTCTCTÁ ACTTCAGCÁ CTGATGCAAG GCAAGTTGÁ ACTTACTTGT

610     620     630     640     650     660     670     680     690     700
GGAAACAATT CTTGGAGGG CATTCCCGCT CGAGGCCAT GGGAGCTGCÁ GTATTGGATG GAATTGACTT TGATATTGAG GGAGGGACTG ACCAATATTG

710     720     730     740     750     760     770     780     790     800
GGATGACCTT GCTAGGTACC TATCTGGATÁ TAGCAAGAGÁ GGCAAGAAÁ TTTACTTGAC TGCTGCCCCÁ CAATGTCCCT TTCTGATGC TTGGATTGGA

810     820     830     840     850     860     870     880     890     900
AATGCACTTÁ AGACAGGCC CTTTGACAAT GTTTGGGTTT AGTTTCAAÁ CAACCCTCCÁ TGTCTAGTACÁ CTTCGGCGÁ CGTGGCCAAT CTGGAGGACG

910     920     930     940     950     960     970     980     990     1,000
CTTGAAGCÁ GTGGACTTCÁ GCCATCCCTG CACATAAGAT TTTCTTGGGG TTGCCTGCTG CTCCTCAGGC TGCCGGTAGC GGTTTATTTC CTGCTTCTGÁ

1,010   1,020   1,030   1,040   1,050   1,060   1,070   1,080   1,090   1,100
TCTCAACTCÁ CAAGTCCTTC CGGCTATTAA AAGCTCGGCT AAGTATGGAG GTGTCATGCT TTGGTCCAAG TATTATGATG ATCTTGATGG ATACAGCTCC
C209Not-R - P...

1,110   1,120   1,130   1,140   1,150   1,160   1,170   1,180   1,190   1,200
TCCATCAAGÁ GTCATGTCGÁ GGCCGCCAGC TTTCTAGAAC AAAAATCAT CTCAGAAGAG GATCTGAATA GCGCCGTCGÁ CCATCATCAT CATCATCATT

1,210   1,220   1,230   1,240   1,246
GAGTTTGTAG CCTTAGACAT GACTGTTCCCT CAGTTCAAGT TGGGCA

```

- Sequência de DNA da isoforma C209XhoNot-PIC

```

1      10     20     30     40     50     60     70     80     90     100    110
ACATTATTCG AAACGATGAG ATTCCTTCÁ ATTTTACTG CTGTTTTATT CGCAGCATCC TCCGCATTA GCTGCTCCAGT CAACACTACÁ ACAGAAGATG AAACGGCACÁ

120     130     140     150     160     170     180     190     200     210     220
AATTCGGGCT GAAGCTGTCÁ TCGGTTACTC AGATTTAGAA GGGGATTCG ATGTTGCTGT TTTGCCATT TCCAACAGCÁ CAAATAACGG GTTATTGTTT ATAATACTÁ

230     240     250     260     270     280     290     300     310     320     330
CTATTGCCAG CATTGCTGCT AAAGAAGAA GGGTATCTCT CGAGCGGCCG CCAGTGTGAT GGATATCTGC AGAATTCGGC TTAAGCTCGÁ GAAAAGAATG GCTTCAAAGT
C209Xho-F - Primer

340     350     360     370     380     390     400     410     420     430     440
CCACAATGTC CCTAGCATTG CTCTCTTTAG TGACACTAGT TCTGGCTTTG GGGGCTAATG CCGGTGGAAT AGCTATATAT TGGGGTCAAÁ ATGGGAATGÁ AGGCACGTTÁ
C209Xho-R - Primer

450     460     470     480     490     500     510     520     530     540     550
GCGGAAACTT GTGCTACAGG GAATTACCAÁ TATGTAACG TAGCTTTTCT CACCACCTC GGCAATGGCC AAACCCCGC CATCAACCTÁ GCCGGTCACT GTGTCCCAAC

560     570     580     590     600     610     620     630     640     650     660
GACCGATGAA TGCACCAAGT TGAGCCCGA GATCAGGTCÁ TGCCAAGCCA AGGGGATTA GGTCACTACT TCCATTGGAG GAGCTTCTG GAGCTACTCT CTAACTTCAG

670     680     690     700     710     720     730     740     750     760     770
CAGCTGATGC AAGGCAAGT GCAACTTACT TGTGGAACA CTTCTTGGG GGGCATTCTG CTCTCAGGCC ATTGGGAGCT GCAGTATTGG ATGGAATTGÁ CTTTGATATT

780     790     800     810     820     830     840     850     860     870     880
GAGGGAGGGA CTGACCAATÁ TTGGGATGAC CTTGCTAGGT ACCTATCTGG ATATAGCAA GAGGGCAAGÁ AAGTTTACTT GACTGCTGCC CCACAATGTC CTTTCTGTA

890     900     910     920     930     940     950     960     970     980     990
TGCTTGGATT GAAATGCAC TTAAGACAGG CCTCTTTGAC AATGTTTGG TTCAGTTCÁ CAACAACCTT CCCTGTCAGT ACATTCCGG CGACGTGGCC AATCTGGAGC

1,000   1,010   1,020   1,030   1,040   1,050   1,060   1,070   1,080   1,090   1,100
ACGCTTGGAA GCAGTGGACT TCAGCCATCC CTGCACATAA GATTTTCTTG GGGTTGCCG CTGCTCCTCÁ GGCTGCCGGT AGCGGTTTTÁ TTCTGCTTC TGATCTCAA

1,110   1,120   1,130   1,140   1,150   1,160   1,170   1,180   1,190   1,200   1,210
TCACAAGTCC TTCGGCTAT TAAAGCTCG GCTAAGTATG GAGGTGTCAT GCTTTGGTCC AAGTATTATG ATGATCTTGA TGGATACAGC TCCTCCATCÁ AGAGTCATGT
C209Not-R - Primer

1,220   1,230   1,240   1,250   1,260   1,270   1,280   1,290   1,300   1,310
CGCGGCCGCC AGTTTCTAG AACAAAACT CATCTCAGAA GAGGATCTGÁ ATAGCGCCGT CGACCATCAT CATCATACGÁ TCATTGAGTT TGTAGCCTTÁ GACATGACTG
C209Not-R - Primer

1,320
TTCTCTC

```

- Sequência de DNA da isoforma C525EcoNot-PIC

```

1      10     20     30     40     50     60     70     80     90     100
GACACTTGAG AGATCAAAAÁ ACAACTAAT TATCGAAAAG ATGAGATTTT CTTCAATTTT TACTGCTGTT TTATTCGCAG CATCCTCCGC ATTAGCTGCT

110     120     130     140     150     160     170     180     190     200
CCAGTCAACA CTACAACAGÁ AGATGAAACG GCACAAATTC CGGCTGAAGC TGTCAATCGT TACTCAGATT TAGAAGGGGÁ TTTTCGATGT GCTGTTTTGC

210     220     230     240     250     260     270     280     290     300
CATTTTCCAA CAGCACAAAT AACGGGTTAT TGTTTATAAA TACTACTATT GCCAGCATTG CTGCTAAAGA AGAAGGGGTA TCTCTCGAGA AAAGAGAGGC

310     320     330     340     350     360     370     380     390     400
TGAAGCTGAA TTCCTCGGGG CTAATGCCGG TGGAAATCGCÁ ATATATTGGG GTCAGAAATG TAATGAAGGC ACATTAGCAG AAACATGTGC TTCAGGGAAT
C525Eco-F - Primer

410     420     430     440     450     460     470     480     490     500
TACCAATTTG TAAACGTAG TTTTCTCAC ACCTTCGGCÁ ACGGCCAAAC GCCTGCAATC AACCTAGCCG GTCACTGTGÁ CCCAACGACA GAAGAATGCA

510     520     530     540     550     560     570     580     590     600
CCAAATTGAG CCCAGAAATC AAGTCATGCC AAGCCAAGGG GATTGAAGTC ATACTCTCCA TAGGAGGAGC TTCTGGGAGC TACTCTCTAA CTTCAGCTGÁ

610     620     630     640     650     660     670     680     690     700
TGATGCAAGG CAAGTGCAG CTTACCTGTG GAATAATTC TTGGGAGGGC AATCGTCGTC GAGGCCATTG GGAGCTGCCG TTTTGGATGG AATTGATTTT

710     720     730     740     750     760     770     780     790     800
GACATCGAGG GAGGGACTGÁ CCAACATTGG GATGACCTAG CAAGGTACCT CTCTGGATAT AGCAAGAGAG GCAAGAAAGT TACTTTGACT GCTGCCCCAC

810     820     830     840     850     860     870     880     890     900
AATGTCCCTT TCCTGATGCT TATGTTGGAÁ ATGCACCTAA GACGGCCCTC TTTGACAATG TTTGGGTTCA GTTCTACAA CACCTTCCCT GCCAGTACGC

910     920     930     940     950     960     970     980     990     1,000
TTCTGGGGAT GTGACCAACC TTGAAGACGC CTGGAAGCAG TGGACTTCAG CCATCCCTGC AGATAAGATT TTCTTGGGAT TGCCTGCTGC ACCTCAAGCT

1,010   1,020   1,030   1,040   1,050   1,060   1,070   1,080   1,090   1,100
GCTGGTAGCG GTTTATTCCT TGCTACTGAT CTTAGCTCAC AAGTTCTTCC GGCTATTAÁA AGTTCCGGTÁ AGTATGGAGG CGTCATGCTT TGGTCCAAGT

1,110   1,120   1,130   1,140   1,150   1,160   1,170   1,180   1,190   1,200
ATTATGATGÁ TCTTGATGGA TACAGCTCCT CCATCAAGAA TGAATGCGG GCGGCCAGCT TTCTAGAACA AAAACTCATC TCAGAAGAGG ATCTGAATAG

1,210   1,220   1,230   1,240   1,250   1,260   1,270   1,280   1,285
CGCCGTCGAC CATCATCATC ATCATCATTG AGTTGTAGC CTTAGACATG ACTGTTCCCT AGTTCAAGTT GGGCACTACG AGAAG

```

- Sequência de DNA da isoforma C525XhoNot-PIC

```

1          10          20          30          40          50          60          70          80          90          100          110
CAAAAAACA CTAATTATTC GAAACGATGA GATTTCTTC AATTTTACT GCTGTTTAT TCGCAGCATC CTCCGCATT GCTGCTCCA TCAACACTAC AACAGAAGAT

120          130          140          150          160          170          180          190          200          210          220
GAAACGGCAC AAATTCGGGC TGAAGCTGTC ATCGGTACT CAGATTAGA AGGGGATTC GATGTTGCTG TTTTGCCATT TTCCAACAGC ACAAATAACG GGTATTGTGT

230          240          250          260          270          280          290          300          310          320          330
TATAAATACT ACTATTGCCA GCATTGCTGC TAAAGAAGAA GGGGTATCTC TCGAGAAAAA AATGGCTTC AAGTCCACAG CAACGTTCTT AGCATGCTC TCTTAGTGA

230          240          250          260          270          280          290          300          310          320          330
C252Xho-F - Primer
CACTAGTTCT GGCTCTCGGG GCTAATGCCG GTGGAATCGC AATATATTGG GGTCAGAATG GTAATGAAAG CACATTAGCA GAAACATGTG CTTCAGGGA TTACCAATT

340          350          360          370          380          390          400          410          420          430          440
GTAACGTAG CTTTCTCAC CACCTTCGGC AACGGCCAAA GCCTGCAAT CAACCTAGCC GGTCAGTGTG ACCCAACGAC AGAAGAAATG ACCAAATTGA GCCAGAAAT

450          460          470          480          490          500          510          520          530          540          550
CAAGTCATGC CAAGCCAAGG GGATTAAGT CATACTCTCC ATAGGAGGAG CTTCTGGGAG CTACTCTCTA ACTTCAGCTG ATGATGCAAG GCAAGTTGCA ACTTACCTGT

560          570          580          590          600          610          620          630          640          650          660
GGAATAATTT CTTGGGAGGG CAATCGTCGT CGAGGCCATT GGGAGCTGCC GTTTGGGATG GAATTGATTI TGACATCGAG GGAGGGACTG ACCAACATTG GGATGACCTA

670          680          690          700          710          720          730          740          750          760          770
GCAAGGTACC TCTCTGGATA TAGCAAGAGA GGCAAGAAA TTTACTTGA TGTGCCCCA CAATGTCCCT TTCTGATGC TTATGTTGA AATGCACCTA AGACGGGCTT

780          790          800          810          820          830          840          850          860          870          880
CTTTGACAA TTTTGGGTT AGTTCTACAA CAACCTCCC TGCCAGTAG CTTCTGGGGA TGTGACCAAC CTTGAAGAC CCTGGAAGCA GTGGACTTCA GCCATCCCTG

890          900          910          920          930          940          950          960          970          980          990
CAGATAAGAT TTTCTGGGA TTGCTGCTG CACCTCAAGC TGCTGGTAG GGGTTTATT CTGCTACTGA TCTTAGCTCA CAAGTTCCTC CGGCTATTAA AAGTTCGGCT

1,000          1,010          1,020          1,030          1,040          1,050          1,060          1,070          1,080          1,090          1,100
AAGTATGGAG GCGTCATGCT TTGGTCCAAG TATTATGATG ATCTTGATGG ATACAGCTCC TCCATCAAGA ATGATGTGCG GCCCGCCAG TTTCTAGAAC AAAAATCAT

1,110          1,120          1,130          1,140          1,150          1,160          1,170          1,180          1,190          1,200          1,210
C525Not-R - Primer
CTCAGAAGAG GATCTGAATA GCGCCGTCGA CCATCATCAT CATCATCATT GAGTTTGTAG CCTTAGACAT GACTGTTCTT CAGTCAAGT TGGGCACAC G

1,220          1,230          1,240          1,250          1,260          1,270          1,280          1,290          1,300          1,311

```

ANEXO 7: Resultados obtidos

I. Quantificação proteica pelo método de Bradford (1976) e crescimento celular

A quantificação proteica obteve-se pela leitura da absorvância a 595nm de amostras com uma concentração de BSA conhecida traçando-se uma curva de calibração. A concentração de proteína nas amostras desconhecidas foi calculada pela manipulação da equação da linha de tendência ($[proteína](mg.mL^{-1}) = Abs_{595nm} - b/m$) e recorrendo aos valores de Abs_{595nm} destas (Quadro 17). O crescimento celular obteve-se por leitura directa da absorvância a 600nm.

Quadro 17 – Valores de absorvância a 595nm e 600nm das amostras em estudo para as quatro construções nas duas estirpes (GS115 e KM71H) e para as estirpes WT e transfectadas com vector nulo. Concentração proteica total para as mesmas amostras, em mg/mL.

Construção	Estirpe	Horas	Abs _{595nm}	[proteína] mg/mL	OD _{600nm}	Construção	Estirpe	Horas	Abs _{595nm}	[proteína] mg/mL	OD _{600nm}
WT	GS115	24	0,643	0,278	1,500	C525EcoNot	GS115	24	0,662	0,154	4,42
		48	0,693	0,359	1,140			30	0,714	0,223	5,60
		72	0,705	0,379	3,000			57	0,790	0,323	7,62
		96	0,755	0,460	0,280			100	0,855	0,408	8,00
		120	0,774	0,491	0			139	0,849	0,401	9,06
	KM71H	0	0,747	0,142	0,476			163	0,903	0,472	10,22
		66	0,772	0,173	0,538			187	1,045	0,659	10,18
		72	0,804	0,213	0,501			0	0,520	0,005	11,04
		80	0,788	0,193	0,558			6	0,573	0,074	12,44
		114	0,873	0,299	0,547			24	0,981	0,606	12,70
		120	0,896	0,328	0,551		30	1,004	0,636	13,38	
		138	0,808	0,218	0,581		57	0,798	0,368	14,88	
		144	0,850	0,271	0,602		100	0,766	0,326	14,14	
		162	0,884	0,313	0,637		139	0,610	0,123	12,30	
	168	0,811	0,222	0,436	163		0,669	0,199	13,34		
	pPICZαA	GS115	24	0,840	0,389		4,140	187	0,589	0,095	12,16
			30	0,623	0,103		6,180	0	0,519	0,004	1,314
			48	0,699	0,203		6,520	24	0,675	0,207	4,70
54			0,772	0,299	7,800	48	0,752	0,308	7,28		
81			0,935	0,514	8,660	72	0,686	0,222	7,58		
124			0,905	0,474	7,220	96	0,761	0,319	7,46		
148			0,880	0,441	7,640	120	0,657	0,184	5,02		
172		0,926	0,502	8,380	0	0,709	0,322	0,307			
KM71H		0	0,819	0,537	0,409	18	0,784	0,469	0,396		
		18	1,005	0,902	0,501	24	0,800	2,053	0,407		
		24	1,095	1,078	0,575	42	1,326	1,751	0,469		
		42	1,063	1,016	0,612	48	1,383	1,555	0,488		
		48	1,040	0,971	0,661	66	1,318	1,112	0,501		
		66	1,158	1,202	0,659	72	1,356	0,980	0,482		
		72	1,201	1,286	0,665	90	1,316	0,750	0,373		
		90	1,283	1,447	0,676	96	1,323	0,588	0,304		
		96	1,325	1,529	0,681	120	1,268	1,418	0,147		
		120	1,271	1,424	0,647	24	0,687	0,230	7,560		
	C525XhoNot	GS115	24	0,490	0,029	2,260	C209EcoNot	GS115	48	0,760	0,322
48			0,584	0,182	4,160	72			0,837	0,420	9,520
72			0,674	0,328	5,140	96			0,781	0,349	9,100
96			0,658	0,302	2,700	120			0,760	0,322	9,820
120			0,698	0,367	1,500	0			1,138	0,802	17,840
KM71H		0	0,808	0,218	0,473	24			1,146	0,812	20,840
		66	0,801	0,209	0,544	48			1,128	0,789	16,040
		72	0,856	0,278	0,541	72			0,951	0,565	16,200
		80	0,825	0,239	0,553	96			0,886	0,482	15,980
		114	0,880	0,308	0,530						
		120	0,897	0,330	0,577						
		138	0,867	0,292	0,589						
		144	0,848	0,268	0,407						
		162	0,866	0,291	0,593						
		168	0,905	0,340	0,573						

II. Determinação da actividade quitinolítica

Os resultados demonstrados abaixo dizem respeito apenas aos pontos seleccionados no Quadro 12 (pico de concentração proteica e pontos que o flanqueiam).

A actividade quitinolítica que cada amostra apresenta foi determinada através de um teste enzimático, já anteriormente descrito. As médias das três absorvâncias a 420nm medidas para os pontos escolhidos encontram-se no Quadro 18.

Quadro 18 – Média das três absorvâncias medidas a 420nm das amostras seleccionadas para as quatro construções nas duas estirpes (GS115 e KM71H) e para as estirpes WT e transfectadas com vector nulo.

Construção	Estirpe	Horas seleccionadas	Abs 420nm	Construção	Estirpe	Horas seleccionadas	Abs 420nm
WT	GS115	24	0,188	C525EcoNot	GS115	100	0,298
		48	0,207			139	0,307
		72	0,206			163	0,301
		96	0,205			187	0,278
		120	0,256			24	0,335
	KM71H	114	0,163		KM71H	30	0,436
		120	0,142			57	0,579
		138	0,302			100	0,682
		144	0,257			139	0,577
		162	0,150			48	0,510
pPICZαA	GS115	54	0,349	C209EcoNot	GS115	72	0,474
		81	0,332			96	0,575
		124	0,338			120	0,528
		148	0,317			24	0,208
		172	0,338			42	0,282
	KM71H	24	0,322		KM71H	48	0,265
		42	0,314			66	0,363
		48	0,290			72	0,522
		66	0,329			90	0,265
		72	0,300			96	0,390
C525XhoNot	GS115	90	0,332	C203EcoNot	GS115	24	0,208
		96	0,348			48	0,209
		24	0,250			72	0,219
		48	0,338			96	0,288
		72	0,294			120	0,125
	KM71H	96	0,329		KM71H	24	0,170
		120	0,297			48	0,210
		80	0,277			72	0,218
		114	0,246			96	0,291
		120	0,251				
	138	0,241					

III. O-glicosilação

O padrão de O-glicosilação das proteínas expressas foi determinado com o software *NetOGlyc 3.1 Server* que procura resíduos de serina e treonina passíveis de serem glicosilados. Obtiveram-se os resultados das Figura 21e 22.

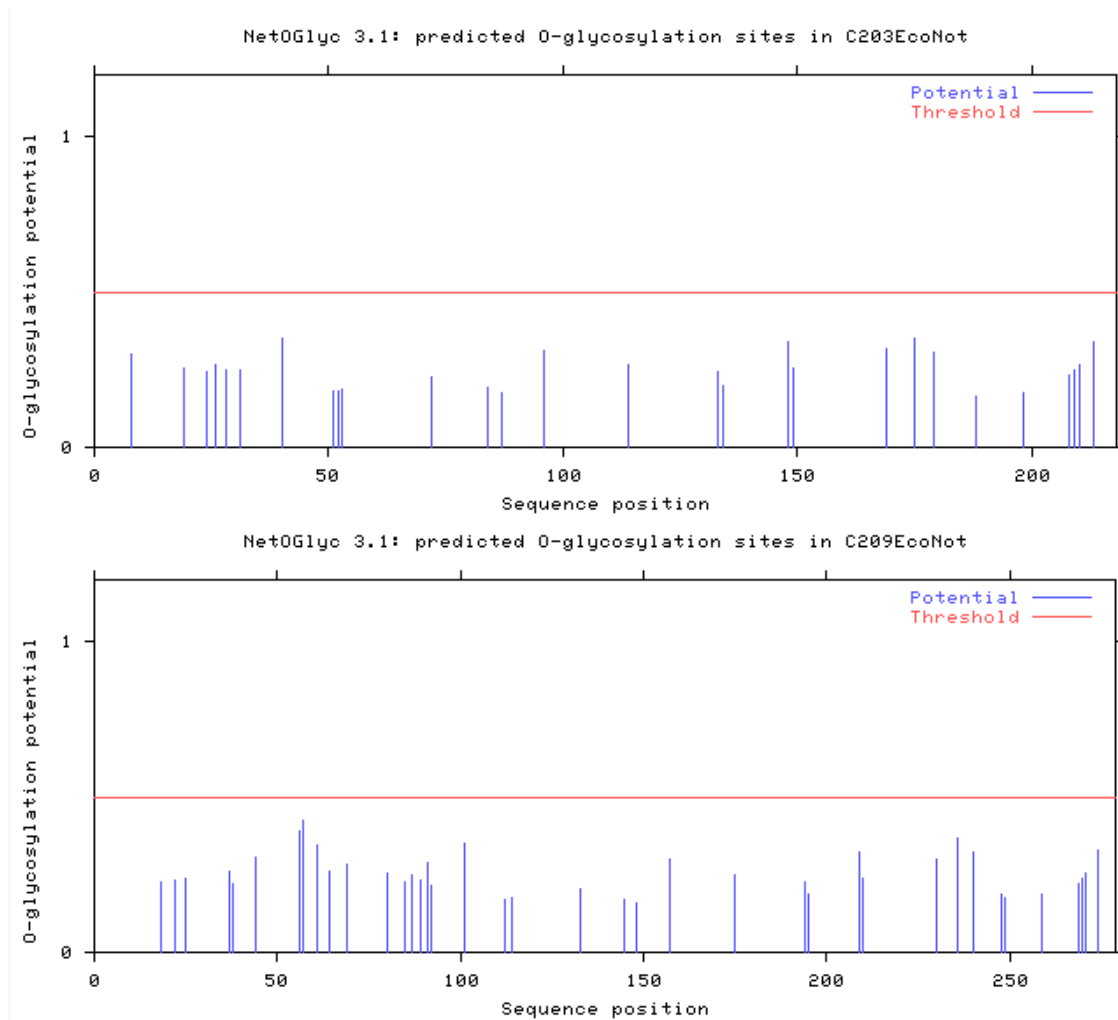


Figura 21 - Output do software *NetOGlyc 3.1 Server* que detecta os resíduos de Ser e Thr passíveis de serem O-glicosilados para as proteínas recombinantes A) C203EcoNot e B) C209EcoNot.

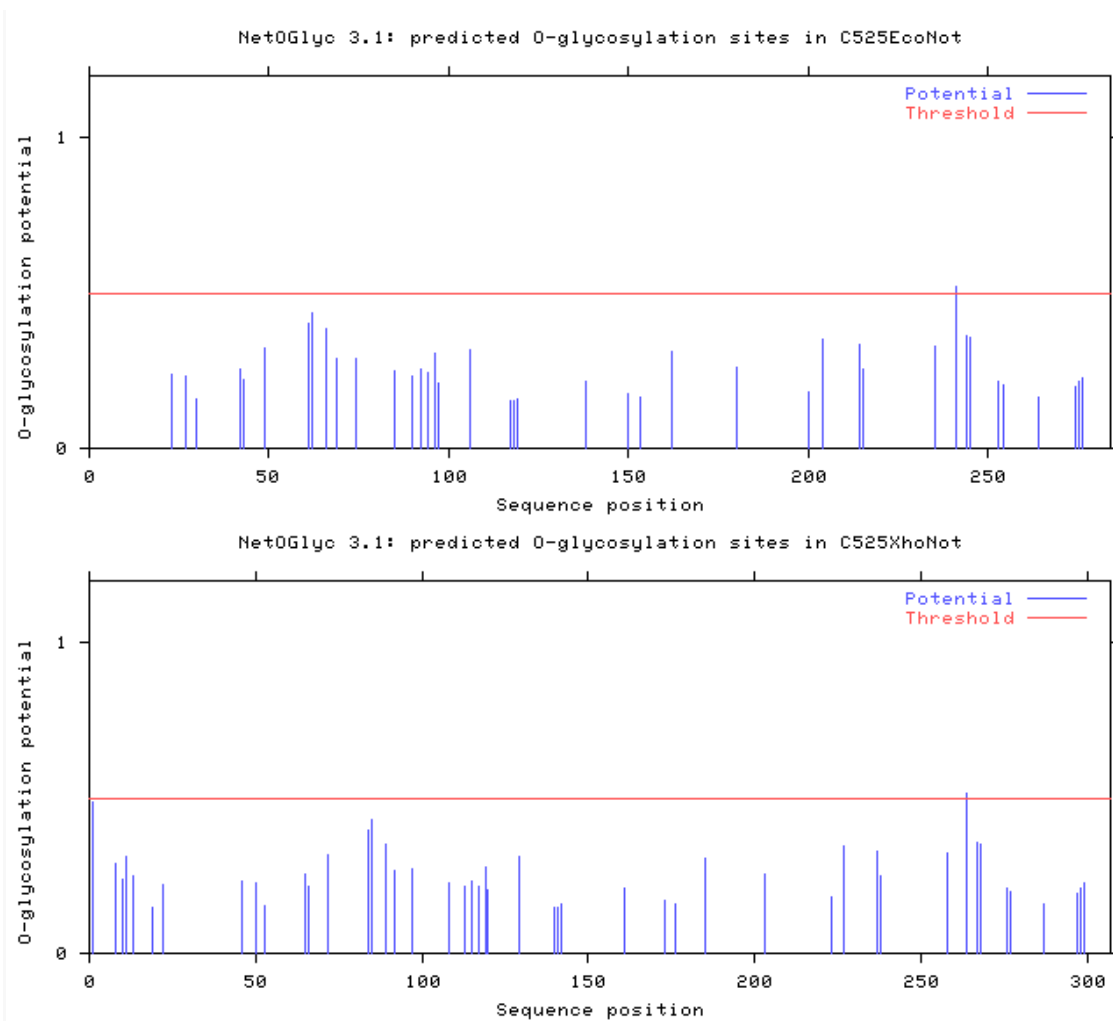


Figura 22 - Output do software *NetOGlyc 3.1 Server* que detecta os resíduos de Ser e Thr passíveis de serem O-glicosilados para as proteínas recombinantes A) C525EcoNot e B) C525XhoNot.