



UNIVERSIDADE DE SÃO PAULO

FACULDADE DE CIÊNCIAS FARMACÊUTICAS DE RIBEIRÃO PRETO

**Obtenção de peptidases heterólogas do fungo filamentoso
Rhizomucor miehei a partir de análise do transcriptoma**

Júlia Raspante Martins

**Ribeirão Preto
2021**

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE CIÊNCIAS FARMACÊUTICAS DE RIBEIRÃO PRETO

Obtenção de peptidases heterólogas do fungo filamentoso
***Rhizomucor miehei* a partir de análise do transcriptoma**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Biociências e Biotecnologia da Faculdade de Ciências Farmacêuticas de Ribeirão Preto para obtenção do Título de Mestre em Ciências

Área de Concentração: Bioagentes e Biotecnologia Aplicados à Farmácia

Orientado: Júlia Raspante Martins

Orientadora: Prof. Dr. Hamilton Cabral

Ribeirão Preto

2021

LOMBADA

MARTINS,
J.R.

**Obtenção de peptidases heterólogas do fungo filamentoso
Rhizomucor miehei a partir de análise do transcriptoma**

Espaço de 2,5 cm
reservado para
etiqueta de
localização da
biblioteca

MESTRADO

FCFRP-USP

2021

FICHA CATALOGRÁFICA

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Martins, Júlia Raspante

Obtenção de peptidases heterólogas do fungo filamentoso *Rhizomucor miehei* a partir de análise do transcriptoma. Ribeirão Preto, 2021.

59p.: 22il.; 30cm.

Dissertação de Mestrado, apresentada à Faculdade de Ciências Farmacêuticas de Ribeirão Preto/USP – Área de concentração: Bioagentes e Biotecnologia Aplicados à Farmácia

Orientador: Cabral, Hamilton.

1. Peptidases fúngicas. 2. Hidrolases. 3. Enzimas recombinantes.

FOLHA DE APROVAÇÃO

Júlia Raspante Martins

Obtenção de peptidases heterólogas do fungo filamentosso *Rhizomucor miehei* a partir de análise do transcriptoma

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Biociências e Biotecnologia da Faculdade de Ciências Farmacêuticas de Ribeirão Preto para obtenção do Título de Mestre em Ciências

Área de Concentração: Bioagentes e Biotecnologia Aplicados à Farmácia.

Orientadora: Prof. Dr. Hamilton Cabral

Aprovado em:

Banca Examinadora

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

A Neville, por quase metade de meus genes
e toda a liberdade.

AGRADECIMENTOS

Agradeço a todos do Laboratório de Tecnologia Enzimática pelo aprendizado, ajuda e companheirismo;

A minha família de esquisitos pelo suporte;

Aos melhores amigos do mundo por colorir minha vida;

Ao meu escritor favorito, Isaac Asimov, e às minhas bandas favoritas, MGMT e Modest Mouse, por suas obras; e

Ao meu psicólogo e ao meu psiquiatra por ensinarem a me compreender.

Todos vocês foram extremamente importantes para me manter o mais próxima possível da sanidade, fazendo pesquisa durante a pandemia de COVID-19 sob um (des)governo corrupto que nega ciência.

Agradeço também a comunidade de código e ciência aberta.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Existe uma teoria que diz que, se um dia alguém descobrir exatamente para que serve o Universo e por que ele está aqui, ele desaparecerá instantaneamente e será substituído por algo ainda mais estranho e inexplicável.

Existe uma segunda teoria que diz que isso já aconteceu.

Douglas Adams, *O Restaurante no Fim do Universo*

RESUMO

MARTINS, J. R. **Obtenção de peptidases heterólogas do fungo filamentoso *Rhizomucor miehei* a partir de análise do transcriptoma**. 2021. 59f. Dissertação (Mestrado). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2021.

Enzimas são proteínas catalizadoras amplamente utilizadas em diferentes tipos de indústrias, sendo as de origem microbiana as mais comuns. Os microrganismos possuem grande diversidade genética o que os tornam fontes promissoras para pesquisa e identificação de enzimas com o potencial biotecnológico. O fungo termofílico *Rhizomucor miehei* é conhecido por produzir hidrolases com potencial de aplicação industrial, sobretudo as subclasses lipases e peptidases. A aplicabilidade das novas peptidases é influenciada pela caracterização funcional e por fatores que contribuem e modulam sua estabilidade térmica. Este trabalho apresenta ensaios de atividade enzimática do filtrado de cultura de *R. miehei*, construção e análise do seu transcriptoma, análises *in silico* de duas de suas peptidases e expressão das últimas em *Pichia pastoris*. O cultivo submerso do fungo foi realizado em um meio suplementado com quatro fontes orgânicas complexas (celulose microcristalina, azeite, caseína e farinha de pena) que atuam como indutores da produção de enzimas despolimerizantes. O filtrado de cultura, obtido a partir dos cultivos, foi utilizado para determinação da atividade enzimática de peptidases, lipases, amilases, endoglucanases, xilanases e lacases. A massa micelial, também proveniente dos cultivos, foi destinada à extração de RNA para sequenciamento e construção do transcriptoma. Foi detectada atividade hidrolítica de peptidases, amilases e lipases, com perfis diferentes em relação à suplementação do meio de cultivo, evidenciando o potencial das enzimas secretadas pelo fungo. No sequenciamento, foram obtidas 69080 sequências de transcrição e, dessas sequências, quase 70% tiveram correspondência com proteínas revisadas do UniProtKB. Entre as sequências anotadas, aproximadamente um terço são hidrolases, sendo 404 peptidases diferentes. Foram selecionadas duas sequências do transcriptoma que codificam uma aspártico peptidase e uma carboxipeptidase, com base na anotação funcional, para análise do potencial biotecnológico *in silico* e expressão heteróloga. A clonagem, expressão heteróloga e purificação da aspártico peptidase mostrou-se factível enquanto para a carboxipeptidase, apesar de ter sido devidamente clonada, não foi possível detectar expressão no meio extracelular.

Palavras chave: Peptidases fúngicas. Hidrolases. Enzimas recombinantes.

ABSTRACT

MARTINS, J. R. **Obtaining heterologous peptidases from the filamentous fungus *Rhizomucor miehei* from transcriptome analysis**. 2021. 59f. Dissertation (Master). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2021.]

Enzymes are catalyst proteins widely used in different types of industries, with those of microbial origin being the most common. Microorganisms have great genetic diversity, which makes them promising sources for research and identification of enzymes with biotechnological potential. The thermophilic fungus *Rhizomucor miehei* is known for producing hydrolases with potential for industrial application, especially the subclasses lipases and peptidases. The applicability of new peptidases is influenced by their functional characterization and by factors that contribute to and modulate their thermal stability. This work presents assays of enzymatic activity of *R. miehei* culture filtrate, construction and analysis of its transcriptome, *in silico* analysis of two of its peptidases and expression of the latter in *Pichia pastoris*. The submerged cultivation of the fungus was carried out in a medium supplemented with four complex organic sources (microcrystalline cellulose, olive oil, casein and feather meal) that act as inducers of the production of depolymerizing enzymes. The culture filtrate, obtained from the cultures, was used to determine the enzymatic activity of peptidases, lipases, amylases, endoglucanases, xylanases and laccases. The mycelial mass, also from the cultures, was used to extract RNA for sequencing and construction of the transcriptome. Hydrolytic activity of peptidases, amylases and lipases was detected, with different profiles in relation to the supplementation of the culture medium, showing the potential of the enzymes secreted by the fungus. In the sequencing, 69080 transcription sequences were obtained and, of these sequences, almost 70% were matched with reviewed proteins from UniProtKB. Among the annotated sequences, approximately one third are hydrolases, with 404 different peptidases. Two transcriptome sequences encoding an aspartic peptidase and a carboxypeptidase were selected, based on functional annotation, for analysis of *in silico* biotechnological potential and heterologous expression. Cloning, heterologous expression and purification of the aspartic peptidase proved to be feasible while for the carboxypeptidase, despite having been properly cloned, it was not possible to detect expression in the extracellular medium.

Keywords: Fungal peptidases. Hydrolases. Recombinant enzymes.

LISTA DE FIGURAS

Figura 1. Fungo <i>R. miehei</i> cultivado em meio ágar de batata dextrose.....	5
Figura 2. Esquema da interação entre sítio ativo de uma peptidase e o substrato	7
Figura 3. Fluxograma do desenvolvimento do projeto.....	9
Figura 4. Atividade enzimática do filtrado de cultura do fungo <i>R. miehei</i>	20
Figura 5. Número de sequências totais no transcriptoma de <i>R. miehei</i>	22
Figura 6. Distribuição da porcentagem de identidade das sequências do transcriptoma	22
Figura 7. Expressão relativa das de enzimas do transcriptoma	23
Figura 8. Expressão relativa das subclasses de hidrolases do transcriptoma.....	23
Figura 9. Expressão relativa das famílias de peptidases do transcriptoma.....	24
Figura 10. Alinhamento entre as sequências peptídicas	25
Figura 11. Predição de presença de peptídeo-sinal.....	26
Figura 12. Predição de N-glicosilação.	26
Figura 13. Representação da estrutura tridimensional da RmAspPep.	27
Figura 14. Eletroforese em gel de agarose do produto de PCR de colônia	30
Figura 15. Atividade proteolítica do meio de cultivo das colônias de <i>P. pastoris</i> KM71H/rRmAspPep.....	31
Figura 16. SDS-PAGE das colônias de <i>P. pastoris</i> KM71H/rRmAspPep.....	31
Figura 17. Purificação de rRmAspPep.....	33
Figura 18. Predição de presença de peptídeo-sinal.....	35
Figura 19. Predição de N-glicosilação	35
Figura 20. (A) Representação da estrutura tridimensional da RmCarPep	36
Figura 21. Eletroforese em gel de agarose de PCR de colônia	38
Figura 22. SDS-PAGE das colônias de <i>P. pastoris</i> KM71H/rRmCarPep	39

LISTA DE TABELAS

Tabela 1. Identificação das sequências da aspártico peptidase.....	25
Tabela 2. Identificação das sequências de serino carboxipeptidases	34

LISTA DE ABREVIATURAS E SIGLAS

ABTS	Ácido 2,2'-Azino-bis(3-etilbenzotiazolina-6-sulfônico)
Abz	O-aminobenzoil
AOX1	Gene da Álcool oxidase 1
AOX2	Gene da Álcool oxidase 2
ATC	Ácido tricloroacético
BLAST	<i>Basic Local Alignment Search Tool</i>
BMGY	<i>Buffered glycerol-complex medium</i>
BMMY	<i>Buffered methanol-complex medium</i>
CMC	Carboximetilcelulose
DNA	Ácido desoxirribonucleico
DNS	Ácido dinitrosalicílico
D.O. ₆₀₀	Densidade ótica com leitura a 600 nm
DTT	Ditiotreitol
EC	<i>Enzyme Commission</i>
EDDnp	Etileno diamino 2,4-dinitrofenil
EDTA	Ácido etilenodiamino tetra-acético
FASTA	Formato baseado em texto para sequências de nucleotídeos e aminoácidos representados por uma letra
FRET	<i>Fluorescence Resonance Energy Transfer</i>
HEPES	4-(2-hidroxietil)-1-piperazinetanolsulfônico
ID	Identidade
IMAC	Cromatografia de afinidade por metal imobilizado
Kex2	Kexin (serino endopeptidase)
LB	Luria-Bertani
LNBR	Laboratório Nacional de Biorrenováveis
pb	Pares de base
<i>Pmel</i>	Enzima de restrição de <i>Pseudomonas mendocina</i>
PMSF	Fenilmetilsulfonil fluoreto
PCR	Reação em Cadeia da Polimerase
RmAspPep	Aspártico peptidase de <i>R. miehei</i>
RmCarPep	Carboxipeptidase de <i>R. miehei</i>

RNA	Ácido ribonucleico
RNAm	RNA mensageiro
rRmAspPep	Aspártico peptidase recombinante de <i>R. miehei</i>
rRmCarPep	Carboxipeptidase recombinante de <i>R. miehei</i>
SDS-PAGE	<i>Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis</i>
TPM	Transcritos Por Milhão
UA	Unidade de Absorbância
<i>Xba</i> I	Enzima de restrição de <i>Xanthomonas badrii</i>
YEPD	<i>Yeast Extract Peptone Dextrose</i>
YNB	<i>Yeast Nitrogen Base</i>
YPDS	<i>Yeast Extract Peptone Dextrose Sorbitol</i>

Sumário

RESUMO	i
ABSTRACT.....	ii
LISTA DE FIGURAS.....	iii
LISTA DE TABELAS	iv
LISTA DE ABREVIATURAS E SIGLAS.....	v
1. Introdução.....	3
1.1. Enzimas e microrganismos com potencial biotecnológico	3
1.2. O fungo <i>Rhizomucor miehei</i>	4
1.3. Classificação das peptidases	6
1.4. Sistema de expressão de proteínas em <i>Pichia pastoris</i>	7
2. Objetivos.....	8
2.1. Objetivo geral	8
2.2. Objetivos específicos	8
3. Materiais e métodos	8
3.1. Origem e manutenção do fungo <i>R. miehei</i>	9
3.2. Cultivo submerso do fungo <i>R. miehei</i>	9
3.3. Atividade enzimática do filtrado de cultura de <i>R. miehei</i> e quantificação de proteínas totais.....	10
3.3.1. Atividade de peptidases	10
3.3.2. Atividade de glicosidases.....	11
3.3.3. Atividade de lipases.....	11
3.3.4. Atividade de lacases	12
3.3.5. Quantificação de proteínas totais	12
3.4. Transcriptoma	12
3.4.1. Extração e sequenciamento do RNA.....	12
3.4.2. Análises bioinformáticas.....	14
3.5. Análise e seleção das sequências de peptidases	14
3.6. Clonagem e expressão heteróloga das peptidases.....	15
3.6.1. Construção dos vetores de expressão	15
3.6.2. Transformação em <i>Escherichia coli</i> DH5 α	16
3.6.2. Transformação em <i>P. pastoris</i> KM71H.....	16
3.6.3. Triagem das colônias de <i>P. pastoris</i> KM71H	17
4. Resultados de discussão	19
4.1. Atividade enzimática do filtrado de cultivo	19

4.2. Mapeamento do transcriptoma.....	21
4.3. Aspártico peptidase	24
4.3.1. Seleção e análise da sequência proteica	24
4.3.2. Clonagem em <i>P. pastoris</i> KM71H.....	29
4.3.3. Análise da expressão da rRmAspPep	30
4.4. Carboxipeptidase.....	33
4.4.1. Seleção e análise da sequência proteica	33
4.4.2. Clonagem em <i>P. pastoris</i> KM71H.....	38
5. Conclusões e perspectivas	40
REFERÊNCIAS BIBLIOGRÁFICAS	41
ANEXOS	44

1. Introdução

1.1. Enzimas e microrganismos com potencial biotecnológico

Enzimas são macromoléculas cruciais para a manutenção da vida. Elas catalisam reações que não ocorreriam em velocidade suficiente para sobrevivência de um organismo vivo mesmo sendo termodinamicamente favoráveis. São proteínas especializadas com aminoácidos reativos, formando o sítio catalítico, promove a catálise sendo mais eficientes que catalisadores sintéticos ou inorgânicos (WOLFENDEN; SNIDER, 2001). O uso de enzimas como biocatalisadores em diversos setores industriais é desejável pela ação rápida e eficiente mesmo em concentrações baixas da enzima e do substrato, com alto grau de especificidade e baixa toxicidade. A reação ocorre em soluções aquosas sob condições brandas de temperatura e pH e pode ser interrompida facilmente (ALI et al., 2018; BANERJEE; RAY, 2017; SANCHEZ; DEMAIN, 2011). O impacto ambiental do uso de enzimas é baixo considerando, mesmo considerando o tipo de resíduo gerado e o modo de obtenção (BANERJEE; RAY, 2017) que podem ser obtidos de fontes animais, vegetais e microbianas (GURUMALLESH et al., 2019).

Os microrganismos são fontes boas para a produção de enzimas devido a diversidade bioquímica, podendo obter enzimas em larga escala apresentando rendimento maior do que as outras fontes, pois não dependem de condições climáticas como as plantas e não enfrentam as questões éticas do uso como em animais (GURUMALLESH et al., 2019). Microrganismos naturalmente produzem enzimas extracelulares para manter o próprio metabolismo ou para defesa, têm crescimento rápido e entende-se bem seu funcionamento orgânico e genético, características que facilitam a produção de enzimas naturais e recombinantes (BANERJEE; RAY, 2017), cuja construção da sequência do gene pode ser otimizada para superexpressar a proteína (GURUMALLESH et al., 2019).

Dentre os microrganismos, os fungos são fontes interessantes de enzimas, direcionadas para setores industriais e biotecnológicos, devido à variabilidade genética destes organismos que permitiu adaptação em condições ambientais diversas (SILVA et al., 2005) e então a capacidade de metabolizar múltiplos polímeros complexos numa faixa ampla de estabilidade e especificidade (BANERJEE; RAY, 2017).

Enzimas de microrganismos são produzidas em larga escala por processo de fermentação em biorreatores com meio de cultura que permita o desenvolvimento celular e estimule a expressão de genes que codificam enzimas de interesse (GURUMALLESH et al., 2019). Esses microrganismos podem ser isolados de amostras de vários locais naturais, como solo e lagos, também de resíduos de efluentes industriais. A escolha do microrganismo apropriado é fundamental na viabilidade das enzimas produzidas por eles (GURUMALLESH et al., 2019).

Fungos filamentosos são capazes de secretar proteínas funcionalmente ativas que apresentam grande vantagem quando comparadas a enzimas bacterianas pois é possível obter extrato enzimático por etapas simples de purificação que separam a solução de enzimas do micélio e dos outros elementos do meio de cultivo (WU et al., 2006). Enzimas para uso industrial são produzidas em larga escala e quando requerem pouca purificação (BANERJEE; RAY, 2017) têm custo de produção baixo por não utilizam técnicas e aparelhos refinados para obtenção do extrato enzimático (ALVES et al., 2007). Por outro lado, enzimas utilizadas nos setores farmacêuticos e clínicos devem ser altamente purificadas (BANERJEE; RAY, 2017), elevando assim o custo do produto final.

1.2. O fungo *Rhizomucor miehei*

O fungo filamentoso *R. miehei* foi descrito inicialmente por Schipper em 1978, que o diferenciou do gênero *Mucor* pelo desenvolvimento de estolões e rizóides, e pela temperatura requerida para crescimento que o caracteriza como termofílico. A classificação dessa espécie segue a ordem: Fungi, Mucoromyceta, Mucoromycota, Mucoromycotina, Mucoromycetes, Mucorales, Lichtheimiaceae, *Rhizomucor*. Suas colônias se apresentam em coloração cinza-esverdeada (Figura 1) e temperatura de crescimento entre 25-57 °C (MYCOBANK, 2018).

Figura 1. Fungo *R. miehei* cultivado em meio ágar de batata dextrose



Fonte: Autor.

O genoma de *R. miehei* CAU432 foi sequenciado e reportado em Zhou et al. (2014). Um total de 155 peptidases foram encontradas, dessas, 43 contém peptídeo-sinal de secreção. O número de peptidases do genoma do *R. miehei* é maior que o reportado anteriormente para *Aspergillus fumigatus* (136 peptidases) e *A. oryzae* (134 peptidases), mas menor que *A. nidulans* (210 peptidases) e *A. niger* (169 peptidases). *R. miehei*, há grande número de serino endopeptidases (EC 3.4.21.-), aspártico endopeptidases (EC 3.4.23.-), metalo endopeptidases (EC 3.4.24.-), treonino endopeptidases (EC 3.4.25.1) e outras duas peptidases ubiquitina-específicas foram encontradas (EC 3.4.19.12) (SUN et al., 2018).

R. miehei é conhecido por sua elevada capacidade de produção de enzimas hidrolíticas, principalmente peptidases e lipases (SUN et al., 2018). A lipase comercial Lipozyme[®] RM IM (Novozymes A/S) de *R. miehei* foi empregada no estudo de Cozertino et al. (2020) para reação de acidólise de óleo de semente de uva. Essa enzima tem capacidade de síntese de triacilglicerol a partir de ácidos graxos do óleo e, uma vez modificado, pode ter efeitos positivos para a saúde quando consumido. Outro potencial para lipases de *R. miehei* é produção de biodiesel. A mistura de lipases imobilizadas de *Candida rugosa* e *R. miehei* aplicada em resíduo de óleo de cozinha produziu biodiesel com rendimento maior que 96% (BINHAYEEDING et al., 2020).

Novas enzimas têm sido exploradas para potencial aplicação em indústria alimentícia, como a aspártico peptidase com potencial para tenderização de carne (SUN et al., 2018) e a α -amilase capaz de produzir alto nível de maltose a partir de amido liquefeito (WANG et al., 2020). Nos dois exemplos anteriores, os genes referentes às enzimas foram obtidos de *R. miehei* CAU432, clonados e expressos em *Pichia pastoris*.

Considerando o potencial de aplicação, é importante conhecer novas enzimas e suas características. A função das enzimas pode ser otimizada em ambientes não-naturais utilizando técnicas de engenharia de proteínas, como *design* racional, no qual são feitas modificações na estrutura primária dessas novas enzimas ou, baseando-se nelas, modificações em enzimas já conhecidas (PLANAS-IGLESIAS et al., 2021).

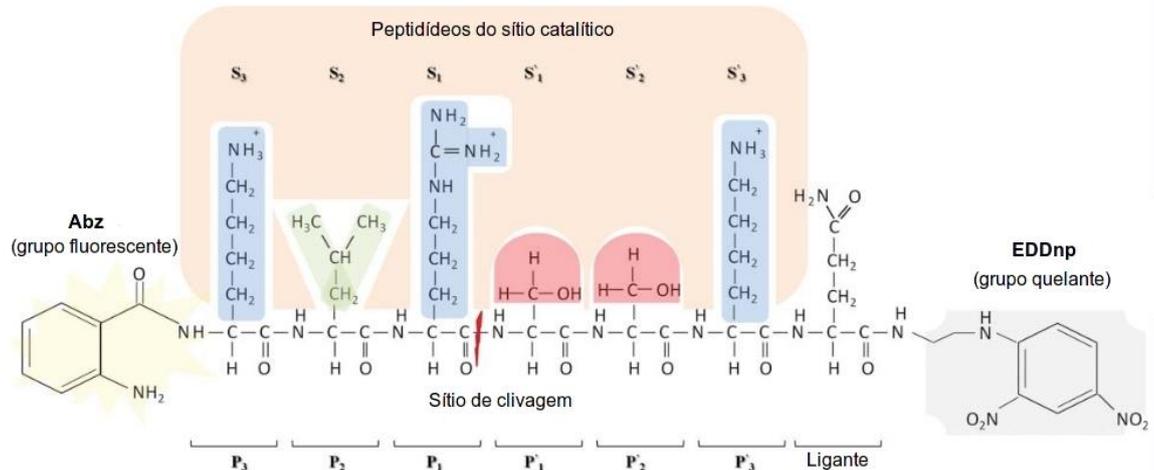
O genoma de *R. miehei* evidencia o potencial de produção de enzimas pelo fungo. Entretanto, estudos transcriptômicos podem revelar variedade maior de possíveis enzimas devido a modificações pós-transcricionais, como *splicing* alternativo e embaralhamento de éxons (GRABHERR et al, 2013).

1.3. Classificação das peptidases

Entre as várias enzimas as hidrolases são as mais utilizadas nos setores industriais e desse grupo a subclasse das peptidases apresenta potencial de aplicação alimentício, farmacêutico, tratamento de efluentes e pesquisa básica e aplicada (BANERJEE; RAY, 2017; GURUMALLESH et al., 2019).

As peptidases são enzimas pertencentes a classe das hidrolases, que clivam ligações peptídicas, recebendo numeração EC 3.4 (Comissão de Enzimas, União Internacional de Bioquímica e Biologia Molecular). Existem dois grandes mecanismos químicos de ação das peptidases separando-as em exopeptidases, que clivam ligações peptídicas nas posições N-C terminais da cadeia polipeptídica e endopeptidases, que atuam geralmente após o quarto resíduo de aminoácido (internamente na cadeia polipeptídica) (RAO et al., 1998). O tamanho e a composição do sítio catalítico através dos resíduos de aminoácidos que o compõe, têm importância na atividade e especificidade que podem ser medidas pela cinética da reação com substratos e inibidores diferentes determinando a eficiência da catálise (SCHECHTER; BERGER, 1967), assim cada resíduo de aminoácido do sítio catalítico da enzima é considerado subsítio denominado como $S_n \dots S_1$ a partir do aminoterminal até o ponto de clivagem e $S'_1 \dots S'_n$ deste ponto até carboxiterminal, e, da mesma forma, os subsítios do substrato são denominados $P_n \dots P_1$ e $P'_1 \dots P'_n$ (SCHECHTER; BERGER, 1967) como representado pela interação de uma peptidase hipotética com substrato Abz-KLRSSKQ-EDDnp na Figura 2. A faixa de pH ótimo de atividade, também subclassifica as peptidases em ácidas (pH 2-6), neutras (pH 6–8) e alcalinas (pH 8–13) (RAO et al., 1998).

Figura 2. Esquema da interação entre sítio ativo de uma peptidase e o substrato Abz-KLRSSKQ-EDDnp. Grupo fluorescente Abz: O-aminobenzoil. Grupo quelante EDDnp: etileno diamino 2,4-dinitrofenil.



Fonte: Adaptado de ALI et al. (2018).

Peptidases são agrupadas em famílias de acordo com similaridades na estrutura primária entre elas e em clãs com estruturas secundárias e/ou terciárias, como motivos e domínios, similares dentro de cada família. Tais elementos de similaridade são importantes para inferência de relações evolutivas e construção de árvores filogenéticas. As famílias têm o nome conforme o aminoácido principal do sítio catalítico, sendo aspártico, cisteíno, glutâmico, metálico, asparagino, serino e treonino. Há também peptidases com sítio catalítico misto ou desconhecido (RAWLINGS et al., 2018).

A aplicação de peptidases pode ser baseada em sua especificidade, inibidores, ativadores e condições de atividade, características definidas pelo arranjo de aminoácidos que determinam o perfil de interação enzima-substrato (ALI et al., 2018).

1.4. Sistema de expressão de proteínas em *Pichia pastoris*

A elaboração de um sistema de expressão heteróloga pode ser mais eficiente do que a fermentação do organismo que produz a enzima endógena (PEREIRA et al., 2004), além de conhecer exatamente a sequência de nucleotídeos que será expressa. O sistema de expressão em *P. pastoris* é adequado para produção de proteínas recombinantes e apresenta alguns benefícios como alta taxa de produção da proteína de interesse e baixa das proteínas nativas (SHU et al., 2016).

Por ser um microrganismo unicelular, o cultivo e manipulação de leveduras é mais fácil comparando a outros organismos eucariotos. *P. pastoris* tem capacidade de crescimento até alta concentração de biomassa, apresenta promotores induzíveis e alta eficiência de secreção de proteínas. Além disso, desempenha modificações pós-traducionais complexas como processamento proteolítico de sequências de sinal, dobramento correto de proteínas, formação de ponte dissulfeto e N- e O-glicosilação (CREGG et al., 2000; YANG et al., 2018).

2. Objetivos

2.1. Objetivo geral

Análise do transcriptoma do fungo *R. miehei* submetido a diferentes cultivos, seleção de duas peptidases, para clonagem e expressão recombinante em *P. pastoris*.

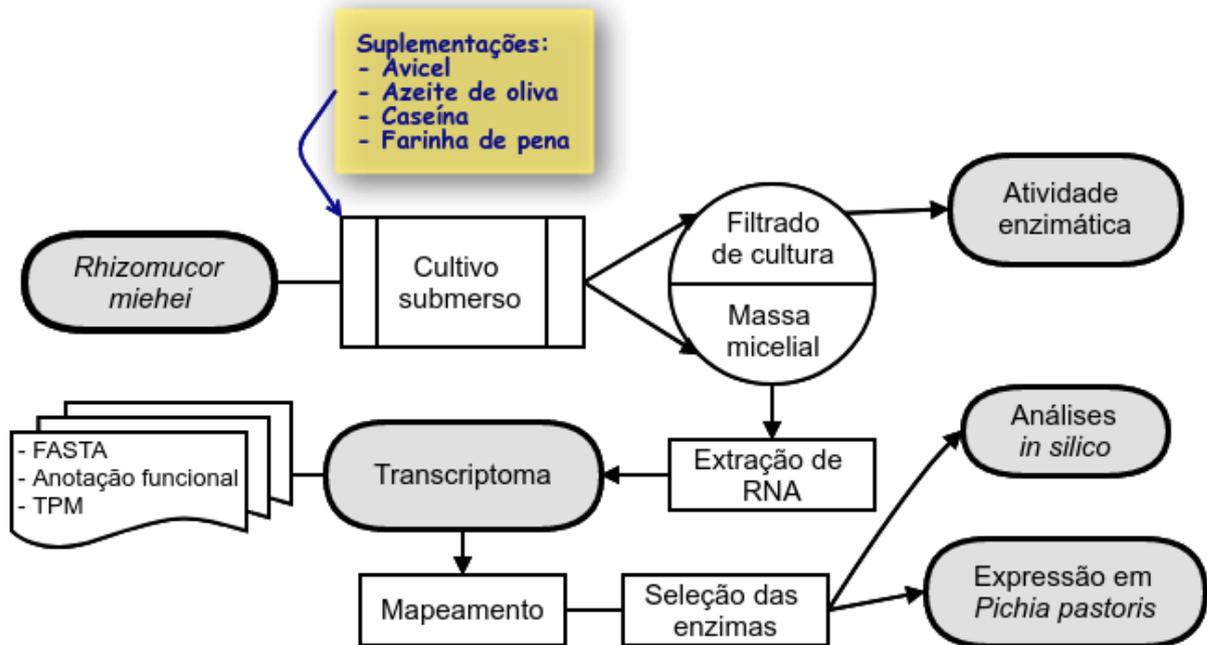
2.2. Objetivos específicos

- Obter panorama de atividade enzimática do filtrado dos cultivos de *R. miehei*;
- Mapeamento do transcriptoma do fungo *R. miehei*;
- Análise *in silico* das peptidases selecionadas; e
- Clonagem e expressão de peptidases selecionadas.

3. Materiais e métodos

As etapas seguidas para alcançar cada objetivo estão resumidas no fluxograma da Figura 3. Nos tópicos seguintes a metodologia é apresentada detalhadamente.

Figura 3. Fluxograma das etapas realizadas no desenvolvimento do projeto, onde as formas elípticas em destaque representam o início (*Rhizomucor miehei*) e resultados finais. TPM: transcritos por milhão.



Fonte: Autor.

3.1. Origem e manutenção do fungo *R. miehei*

O fungo termofílico *R. miehei* foi isolado a partir de matéria orgânica em decomposição no solo e está disponível na micoteca de fungos filamentosos do Laboratório de Tecnologia Enzimática sob a responsabilidade do Prof. Dr. Hamilton Cabral.

Para utilização neste projeto o fungo foi repicado em placas de Petri contendo meio de cultivo ágar de batata dextrose 39 g.L^{-1} mantidos em estufa por 72 h a $45 \text{ }^\circ\text{C}$.

3.2. Cultivo submerso do fungo *R. miehei*

O cultivo do fungo foi conduzido conforme estudos prévios realizados no nosso grupo de pesquisa (DA SILVA et al., 2016) de modo a induzir a expressão e secreção de hidrolases e oxirredutases e obter o filtrado de cultura para realização de ensaios de atividade enzimática e massa micelial para extração de RNA usado para obtenção do transcriptoma. A suplementação dos meios e tempos de cultivo diferentes foram usados para maximizar a cobertura de transcritos, como recomendado pelo método

do programa Trinity (GRABHERR et al., 2013) usado para sequenciamento de RNA.

O cultivo submerso foi realizado em frascos Erlenmeyer de 250 mL, onde foram adicionados 50 mL de solução contendo KH_2PO_4 0,7%, K_2HPO_4 0,2%, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 0,01%, citrato $2\text{H}_2\text{O}$ 0,05%, extrato de levedura 0,1% e $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ 0,01% (TRAN; NAGANO, 2002) ajustado para pH 5. Foram estabelecidas quatro condições de cultivo diferenciadas pela suplementação do meio por adição de celulose microcristalina (Avicel – Sigma), caseína, azeite de oliva ou farinha de pena (penas de frango trituradas) na concentração 1%. O inóculo foi realizado com a adição de 10 discos de 7 mm de diâmetro do pré-inóculo do fungo em cada frasco. O cultivo foi realizado durante 144 h, a 45 °C e agitação 120 rpm. A cada 24 h foi coletada a triplicata de cada cultivo. As culturas foram centrifugadas a 10000 x g por 30 min a 4 °C para separar o filtrado de cultura (sobrenadante) e a massa micelial (precipitado).

3.3. Atividade enzimática do filtrado de cultura de *R. miehei* e quantificação de proteínas totais

Foram realizados ensaios de atividade para hidrolases (peptidases, amilases, endoglucanases, xilanases e lipases) e oxirredutases (lacases) e quantificação de proteínas totais segundo o método de Bradford (1976) do filtrado de cultura do *R. miehei* a fim obter perfil de enzimas secretadas de acordo com o tempo e a suplementação do cultivo. Todas as reações foram feitas em triplicata.

3.3.1. Atividade de peptidases

A atividade de peptidases do filtrado de cultura foi realizado em triplicata conforme protocolo descrito por Sarath et al. (1996) com modificações. Foi utilizada solução caseína 1% preparada em solução de NaH_2PO_4 a 50 mM, pH 6,5. Para a reação foram incubados 100 μL de cada amostra com 500 μL de substrato em banho-maria a 45 °C por 2 h. Em seguida, foi adicionado 300 μL de ácido tricloroacético (ATC) 10% para interromper a reação. Para os controles, foi adicionado ATC 10% antes do início da reação. Em seguida os tubos de reação foram centrifugados a 10000 x g por 10 min a temperatura ambiente (25 °C). A atividade proteolítica foi mensurada a partir do valor de absorbância a 280 nm do sobrenadante dos tubos de reação contra seus respectivos controles. O valor de unidades de enzima ($\text{U} \cdot \text{mL}^{-1}$) foi definida como a quantidade de enzima requerida para promover a liberação de 1 μmol

tirosina por minuto sob condições definidas de ensaio (MEYERS; AHEARN, 1977).

3.3.2. Atividade de glicosidases

Os ensaios de atividade para amilases, CMCases e xilanas foram realizados de acordo com o método de quantificação de açúcares redutores de Miller (1959) com substratos amido solúvel, carboximetilcelulose (CMC) e xilana de madeira de faia, respectivamente, na concentração 0,5% em tampão acetato de sódio a 50 mM e pH 5. Para as reações foram incubados 25 μ L de cada amostra com 25 μ L de substrato a 45 °C por 30 min. A reação foi interrompida com a adição de 50 μ L de solução ácido 3,5-dinitrosalicílico (DNS) em banho fervente por 5 min e, após, foram adicionados 400 μ L de água ultrapura. Os controles foram feitos com amostra previamente fervida. A absorbância das amostras foi lida a 540 nm em leitor de placas SpectraMax (Molecular Devices). A unidade enzimática ($\text{U}\cdot\text{mL}^{-1}$) foi definida como a quantidade de enzima necessária para liberar 1 μ g de açúcar redutor nas condições de ensaio.

3.3.3. Atividade de lipases

O substrato para os ensaios de atividade para lipase pela mistura de solução de palmitato de p-nitrofenila a 8 mM em 2 mL isopropanol com solução de 10 mg de goma arábica e 50 μ L de Triton X-100 em tampão bicina a 50 mM e pH 8,5. Para a reação foram utilizados 50 μ L da amostra e 450 μ L do substrato, incubados em banho-maria a 45 °C. Após 30 min a reação foi interrompida por fervura durante 1 min seguida de banho de gelo e então foi adicionado 500 μ L de solução saturada de tetraborato de sódio. Os controles foram feitos pela fervura das amostras antes da incubação com substrato. A leitura da absorbância a 410 nm foi feita em leitor de placas SpectraMax (Molecular Devices). Para determinar a concentração de p-nitrofenol liberado durante a reação, foi utilizada a equação da regressão linear entre os valores de concentração ($\mu\text{mol}\cdot\text{mL}^{-1}$) e absorbância a 410 nm. A unidade de atividade enzimática ($\text{U}\cdot\text{mL}^{-1}$) foi definida como a quantidade de enzima necessária para liberar 1 μ M de p-nitrofenol por min sob as condições do ensaio (PEREIRA et al., 2014).

3.3.4. Atividade de lacases

Os ensaios de atividade para lacases foram realizados incubando 50 μL de cada amostra com 100 μL de solução ácido 2,2'-azino-bis(3-etilbenzotiazolina-6-sulfônico) (ABTS) a 1 mM. A reação foi interrompida após 10 min pela adição de 150 μL de ATC 10% e o controle de cada amostra foi feito pela adição de ATC 10% antes da adição do substrato. A leitura da absorbância a 405 nm foi feita em leitor de placas SpectraMax (Molecular Devices) e a unidade de atividade enzimática ($\text{U}\cdot\text{mL}^{-1}$) foi definida como a quantidade de enzima necessária para aumentar a absorbância em 1 unidade sob as condições do ensaio (NING et al., 2016).

3.3.5. Quantificação de proteínas totais

A concentração ($\text{mg}\cdot\text{mL}^{-1}$) de proteínas totais foi determinada de acordo com o método de Bradford (1976) usando albumina de soro bovino como padrão. Para a reação foram utilizados 20 μL de amostra e 1 mL de reagente de Bradford incubados por 10 min a temperatura ambiente (25 °C) com leitura de absorbância a 595 nm. Um controle foi feito com 20 μL de água substituindo a amostra.

3.4. Transcriptoma

A fonte de RNA para a construção do transcriptoma foi a mistura da massa micelial proveniente do cultivo do fungo das quatro suplementações (item 3.2.) nos tempos de 24, 48 e 72 h com intuito de obter a maior variedade possível de transcritos produzidos pelo *R. miehei* nessas condições de acordo com o sugerido pela documentação do programa de sequenciamento de RNA Trinity (GRABHERR et al., 2013). A massa micelial foi separada do meio de cultivo por filtração em papel qualitativo, lavada com solução NaCl 0,9% e congelada com N_2 líquido. Dessa mistura foram usados 100 mg para maceração na presença de N_2 líquido.

3.4.1. Extração e sequenciamento do RNA

A extração do RNA da massa micelial macerada procedeu-se com a adição de 500 μL TRI Reagent (Sigma Aldrich), em seguida foram adicionados 100 μL de

clorofórmio, agitada em vórtex e incubada por 12 min a 24 °C. O material foi centrifugado a 12000 x *g* por 5 min a 4 °C, após a centrifugação, retirou-se o sobrenadante e foram adicionados mais 100 µL de clorofórmio e novamente centrifugado. Em seguida, o RNA foi precipitado com 0,7% de isopropanol gelado e incubado por 10 min. Após, a amostra foi centrifugada a 12000 x *g* por 10 min a 4 °C, o precipitado foi lavado com etanol 75% gelado, seco em temperatura ambiente (25 °C) e foram ressuspensos em água livre de RNase. A amostra foi tratada com Dnase I Rnase Free (Invitrogen), em uma reação contendo 1 unidade de enzima para cada 3 µg de RNA total, incubando por 30 min a 37 °C seguido por inativação com 1 µL de EDTA a 50 mM incubada por 5 min a 65 °C.

As etapas de análise da qualidade de RNA e sequenciamento foram realizadas no Laboratório Nacional de Biorrenováveis (LNBR). A análise da qualidade da amostra de RNA foi feita em equipamento NanoVue (GE healthcare), considerando amostra de boa qualidade com valor da relação de absorvância 260/280 nm maior que 1,7, indicando alto grau de pureza de RNA, e por eletroforese em gel de agarose 1% para verificação das bandas de RNA ribossomal e grau de degradação. Outra etapa de análise de qualidade foi realizada com *kit* Total RNA Nano no equipamento 2100 Bioanalyzer (Agilent Technologies) considerando amostra de boa qualidade com valor de integridade de RNA (que varia entre 1-10, sendo 1 degradado e 10 intacto) maior que 7. Antes do preparo da biblioteca, o RNA da amostra foi quantificado em fluorímetro QU Bit 3.0 (Life Technologies) e a concentração foi ajustada para 6 ng.µL⁻¹ em um volume final de 50 µL.

O RNA mensageiro (RNAm) foi separado da amostra total utilizando *beads* carregadas com cauda poli-timina (Oligo-dT beads) para hibridizar com a cauda poli-adenina das moléculas de RNAm e então foi fragmentado e purificado. Foi feita PCR por transcrição reversa e, após, PCR convencional. O material resultante foi tratado com exonucleases para garantir extremidades cegas e com polimerase para adição de adenina nas extremidades 3' das sequências. Foi feita ligação de adaptadores nas sequências e mais uma PCR para enriquecimento delas utilizando *primers* que anelam com os adaptadores.

A validação da biblioteca de sequências geradas foi feita por quantificação por fluorímetro QU Bit 3.0 (Life Technologies) e análise do tamanho das sequências utilizando *kit* DNA 1000 em equipamento Bioanalyzer (Agilent Technologies). Uma segunda quantificação foi feita com as amostras diluídas 1000x por PCR quantitativo

seguindo instruções do Sequencing Library qPCR Quantification Guide. Finalmente, o material foi sequenciado no equipamento Next-seq 500 (Illumina), gerando *reads* de 150 pares de base (pb) em construção *paired-ends*.

3.4.2. Análises bioinformáticas

A construção do transcriptoma também foi feita em parceria com LNBR, utilizando o programa Trinity (GRABHERR et al., 2013) na construção dos transcritos e estimativa de abundância de cada transcrito em transcritos por milhão (TPM). As extensões do Trinity, Trasndecoder e Trinotate, foram usadas para predição de regiões codificadoras e para anotação funcional das sequências, respectivamente. A documentação completa do programa pode ser encontrada em <https://github.com/trinityrnaseq/trinityrnaseq/wiki>.

A partir dos dados da anotação funcional foi feito mapeamento do transcriptoma. Foi feita a separação das sequências de transcritos pelo número EC em conjuntos de enzimas, classes e subclasses enzimáticas e famílias de peptidases. A expressão relativa de cada conjunto foi obtida pela soma dos valores de TPM.

3.5. Análise e seleção das sequências de peptidases

Foram selecionadas sequências de duas peptidases do transcriptoma do *R. miehei* para expressão recombinante e posterior caracterização bioquímica, uma aspártico peptidase (RmAspPep) e uma carboxipeptidase (RmCarPep).

A primeira enzima escolhida foi a RmAspPep previamente estudada pelo grupo de pesquisa, com objetivo de continuar o estudo da caracterização e comparar resultados entre a enzima nativa e a recombinante. A identificação da sequência se deu por busca textual dos quatro fragmentos peptídeos (TWSISYGDGSSASGILAK, ASNGGGGEYIFGGYDSTK, GSLTTVPIDNSR e GWWGITVDRA) gerados por análise de espectrometria de massas da peptidase purificada do filtrado de cultura do *R. miehei* por Da Silva et al., (2016). A segunda enzima escolhida, RmCarPep, baseado na anotação funcional de carboxipeptidase Y homologa A (EC 3.4.16.5), a fim de explorar uma exopeptidase. A escolha dentre as sequências das isoformas da RmAspPep e entre as carboxipeptidases foi feita considerando o valor de abundância relativa dos transcritos, uma vez que valores baixos ou artificialmente nulos podem

representar montagem incompleta e serem biologicamente questionáveis (GRABHERR et al., 2013).

As sequências selecionadas foram analisadas com os programas SignalP-5.0 (NIELSEN et al., 2019) para predição do local de clivagem do peptídeo-sinal, NetNGlyc1.0 (GUPTA; BRUNAK, 2002) para predição de N-glicosilação e NetOGlyc4.0 (STEENTOFT et al., 2013) para predição de O-glicosilação. A ferramenta Swiss-Model (WATERHOUSE et al., 2018) foi usada para criação de modelo de estrutura das sequências e o programa PyMol-2.4.1 (Schrödinger, LLC) utilizado para visualização, alinhamento e construção das figuras de estrutura tridimensional das enzimas. A ferramenta Compute pI/Mw (BJELLQVIST et al., 1994) usada para estimar ponto isoelétrico (pI) e massa molecular. Os bancos de dados Merops (RAWLINGS et al., 2018) e UniProtKB (BATEMAN et al., 2021) foram usados para buscar por proteínas similares utilizando BLAST (*basic local alignment search tool*), e a partir do resultado fazer comparações entre sequências homólogas, como identificação de padrão de estrutura secundária e terciária, prováveis aminoácidos catalíticos e função celular.

3.6. Clonagem e expressão heteróloga das peptidases

Expressão heteróloga foi a estratégia utilizada para produção das enzimas. Os métodos seguiram as orientações do manual do usuário do EasySelect Pichia Expression Kit (INVITROGEN, 2010).

3.6.1. Construção dos vetores de expressão

Para construção do vetor pPICZ α -A (Invitrogen), foi removida a sequência de nucleotídeos correspondentes ao peptídeo sinal das sequências da RmAspPep e da RmCarPep e também a pro-região da sequência da RmCarPep por ser relacionada a endereçamento intracelular, considerando a similaridade entre sequências de outras enzimas da mesma família. Em ambas as construções foram adicionados às sequências, o código para cauda poli-histidina e códon de terminação da tradução. As sequências das enzimas de interesse foram inseridas entre o sinal de clivagem por Kex2 e sítio de restrição por *Xba*I do vetor. A sequência completa dos vetores encontra-se no Anexo 1, com destaque para os nucleotídeos das peptidases

recombinantes. Os vetores de expressão foram sintetizados pela empresa Biomatik, entregues na forma liofilizada, foram ressuspensos em água livre de nuclease para concentração final de $500 \text{ ng} \cdot \mu\text{L}^{-1}$ e estocados a $-20 \text{ }^\circ\text{C}$.

3.6.2. Transformação em *Escherichia coli* DH5 α

A etapa de transformação dos vetores em *E. coli* foi realizada para aumentar o número de cópias dos mesmos. Para a transformação, $200 \mu\text{L}$ de cultura de células *E. coli* da linhagem DH5 α competentes foram incubados com 100 ng do vetor durante 10 min em gelo, em seguida por 90 s em banho-maria a $42 \text{ }^\circ\text{C}$ e novamente em gelo. Foram adicionados $800 \mu\text{L}$ de meio *low salt* LB (Luria-Bertani) e incubado por 1 h em banho-maria a $37 \text{ }^\circ\text{C}$. Após, o cultivo foi centrifugado a $10000 \times g$ por 1 min , foram dispensados $400 \mu\text{L}$ do sobrenadante e as células precipitadas foram ressuspensas no meio restante e plaqueadas em meio seletivo *low salt* LB ágar com $25 \mu\text{g} \cdot \text{mL}^{-1}$ de zeocina e incubadas em estufa por 24 h a $37 \text{ }^\circ\text{C}$.

Para extração e purificação do DNA plasmidial das bactérias transformadas, uma colônia foi selecionada foi cultivada em 5 mL de meio *low salt* LB por 16 h . Foi utilizado o *kit* Wizard Plus SV Minipreps DNA Purification System (PROMEGA, 2010) seguindo as orientações do protocolo do fabricante. A concentração do DNA extraído foi quantificada em equipamento NanoVue (GE healthcare) considerando amostras de boa qualidade quando apresentaram valores de relação de absorbância $260/280 \text{ nm}$ e $260/230 \text{ nm}$ maiores que $1,7$. O DNA purificado foi estocado a $-20 \text{ }^\circ\text{C}$.

3.6.2. Transformação em *P. pastoris* KM71H

A transformação em células de *P. pastoris* consiste na inserção do DNA do vetor ao genoma da levedura para posterior indução de expressão das enzimas em estudo. Antes de iniciar essa etapa, o DNA plasmidial foi linearizado utilizando a enzima de restrição *PmeI*. Foram incubados $6 \mu\text{g}$ de DNA com $1,5 \mu\text{L}$ de *PmeI*, $5 \mu\text{L}$ de CutSmart Buffer $10\times$ e água livre de nuclease para volume final de $50 \mu\text{L}$ por 16 h em banho-maria a $37 \text{ }^\circ\text{C}$ e 20 min a $65 \text{ }^\circ\text{C}$ para inativação da enzima. Foi realizado eletroforese em gel de agarose 1% em tampão trizma a 40 mM , ácido acético a 20 mM e EDTA a 1 mM e pH 8 , corado com SYBR Safe DNA Gel Stain (Invitrogen) $0,01\%$ com amostras do DNA plasmidial e posteriores a restrição para confirmar linearização.

O DNA linearizado foi purificado com o *kit* Wizard SV Gel and PCR Clean-Up System (CORPORATION, 2010) seguindo orientações do protocolo do fabricante. A concentração de DNA foi quantificada como descrito no tópico anterior. O DNA linear purificado foi estocado a -20 °C.

As células de *P. pastoris* KM71H competentes foram preparadas após o cultivo de uma colônia em 5 mL de meio contendo de extrato de levedura 1%, peptona 2% e dextrose 2% (YEPD) em tubo cônico de 50 mL em estufa sob agitação de 200 rpm por 8 h a 30 °C. A concentração de células no cultivo foi determinada por densidade ótica a 600 nm (D.O.₆₀₀). Em seguida, 100 µL ($7,8 \times 10^6$ células) do cultivo foram transferidos para um frasco Erlenmeyer de 250 mL contendo 50 mL de meio YEPD que foi novamente mantido a 200 rpm e 30 °C até atingir valor de D.O.₆₀₀ entre 1 e 1,5 (cerca de 18 h). Após, o frasco foi mantido em gelo por 10 min e o conteúdo foi centrifugado a 1500 x g por 5 min a 4 °C. O sobrenadante foi descartado e o precipitado de células foi ressuspensionado em 10 mL de YEPD mais 2 mL de tampão ácido 4-(2-hidroxietil)-1-piperazinetanosulfônico (HEPES) a 1 M, pH 7, depois foram adicionados 250 µL de ditioneitol (DTT) a 1 M e incubado em estufa por 15 min a 30 °C. Após, foram realizadas três etapas de centrifugação a 1500 x g por 10 min a 4 °C e descarte do sobrenadante: (1) foi adicionado água gelada para completar o volume de 50 mL, (2) em seguida adicionados 25 mL de água gelada, e (3) depois 2 mL de sorbitol a 1 M gelado. Por fim as células precipitadas foram ressuspensas em 100 µL de sorbitol a 1 M gelado.

A transformação foi realizada por eletroporação com pulso de 1500 V com 40 µL da suspensão de células competentes e 3 µg do DNA linearizado em cubetas GenePulser 0,2 cm (BioRad). Após, foram adicionados 500 µL de sorbitol a 1 M e 500 µL de YEPD e as células foram incubadas em estufa por 2 h a 30 °C. O material foi plaqueado em meio seletivo de YEPD com sorbitol a 1 M (YPDS), ágar 2% e zeocina 100 µg.mL⁻¹ e incubado em estufa por 72 h a 30 °C.

As colônias visualmente isoladas foram inoculadas em placas com meio YPDS ágar 2% com zeocina 100 µg.mL⁻¹ em quadrantes delimitados e numerados para cada colônia e foram incubadas em estufa por 72 h a 30 °C e armazenadas a 4 °C.

3.6.3. Triagem das colônias de *P. pastoris* KM71H

O objetivo da triagem é selecionar dentre as colônias de células transformadas

uma que expressa e secreta a proteína de interesse e o tempo necessário para a indução da expressão. A triagem foi realizada com 14 colônias selecionadas arbitrariamente para as duas enzimas recombinantes e uma colônia de *P. pastoris* KM71H sem o vetor de expressão foi usada como controle.

A primeira etapa da triagem foi PCR de colônia utilizando *primers* 5'AOX1 e 3'AOX1 conforme descrito por Linder, Schliwa e Kube-Grandenrath (1996) para confirmar a inserção do vetor ao genoma da levedura. Para extração do DNA a ser submetido a PCR, a lise celular foi feita com colônias suspensas em 20 μL de água, incubadas em banho seco por 10 min a 100 $^{\circ}\text{C}$ e centrifugadas a 10000 $\times g$ por 1 min a temperatura ambiente (25 $^{\circ}\text{C}$). As reações foram preparadas com 2 μL do lisado celular, 2 μL de 10x *buffer* com KCl, 2 μL de MgCl_2 a 25 mM, 0,4 μL de dNTPs a 25 mM, 0,4 μL de cada *primer* (5'AOX1 e 3'AOX1) a 10 μM , 12,4 μL de água livre de nuclease e 0,4 μL de *Taq* polimerase 0,16 $\text{U}\cdot\mu\text{L}^{-1}$. A reação ocorreu em termociclador com incubação inicial por 5 min a 95 $^{\circ}\text{C}$, seguido por 30 ciclos de desnaturação por 1 min a 95 $^{\circ}\text{C}$, anelamento por 1 min a 54 $^{\circ}\text{C}$ e extensão por 1 min a 72 $^{\circ}\text{C}$, e extensão final a 72 $^{\circ}\text{C}$ por 5 min. A análise foi feita por eletroforese em gel de agarose, como descrito no item anterior.

A segunda etapa da triagem foi a análise da expressão das enzimas recombinantes. Foi feito cultivo em pequena escala em minibioreatores Corning de 50 mL incubados em estufa com agitação a 200 rpm a 30 $^{\circ}\text{C}$. Primeiro, as colônias selecionadas foram cultivadas em 10 mL de *buffered glycerol-complex medium* (BMGY) composto por 1% extrato de levedura, 2% peptona, fosfato de potássio a 100 mM, pH 6, 1,34% *yeast nitrogen base* (YNB), $4\cdot 10^{-5}\%$ biotina e 1% glicerol, e cresceram durante 22 h, até atingir valor de D.O.₆₀₀ maior que 2. Depois, para iniciar a indução, os cultivos foram centrifugados a 1500 $\times g$ por 15 min a temperatura ambiente (25 $^{\circ}\text{C}$), o sobrenadante foi descartado e as células ressuspensas em 7 mL de *buffered methanol-complex medium* (BMMY) no qual o glicerol é substituído por 0,5% metanol. A indução durou 144 h, sendo que a cada 24 h era adicionado metanol para concentração final 0,5% e retirada uma alíquota de 200 μL , exceto no último tempo em que todo o material restante foi utilizado. Após a indução, as amostras foram centrifugadas a 1500 $\times g$ por 10 min a temperatura ambiente (25 $^{\circ}\text{C}$) e os sobrenadantes foram separados do precipitado de células e armazenados a -80 $^{\circ}\text{C}$. A detecção das enzimas recombinantes no extrato celular foi feita por eletroforese em gel de poliacrilamida 12% com dodecilsulfato de sódio (SDS-PAGE) usando corante

Coomassie Phast Blue.

Para a RmAspPep recombinante (rRmAspPep) foram realizados ensaios de atividade caseinolítica (descrito no item 3.3.1.) e purificação por cromatografia de afinidade por metal imobilizado (IMAC) seguido por análise por SDS-PAGE da amostra de uma colônia, ambos com amostras obtidas após 144 h de bioprocesso.

A cromatografia foi feita em coluna Tricorn (GE Healthcare) de 1 mL carregada com resina Nickel sepharose (GE Healthcare), a qual tem afinidade pela cauda polihistidina da enzima recombinante. A coluna foi previamente equilibrada com 10 mL de solução KH_2PO_4 a 50 mM e NaCl a 200 mM, pH 8, (solução A). A solução A com adição de imidazol a 250 mM (solução B), que compete com a histidina para interação com a resina, foi usada para eluição. Dois mL da amostra foram injetados no sistema e a eluição foi feita variando a concentração das soluções, de 100% de A até 100% de B em 40 mL com fluxo de 1 mL por min, coletando frações de 1,5 mL.

4. Resultados de discussão

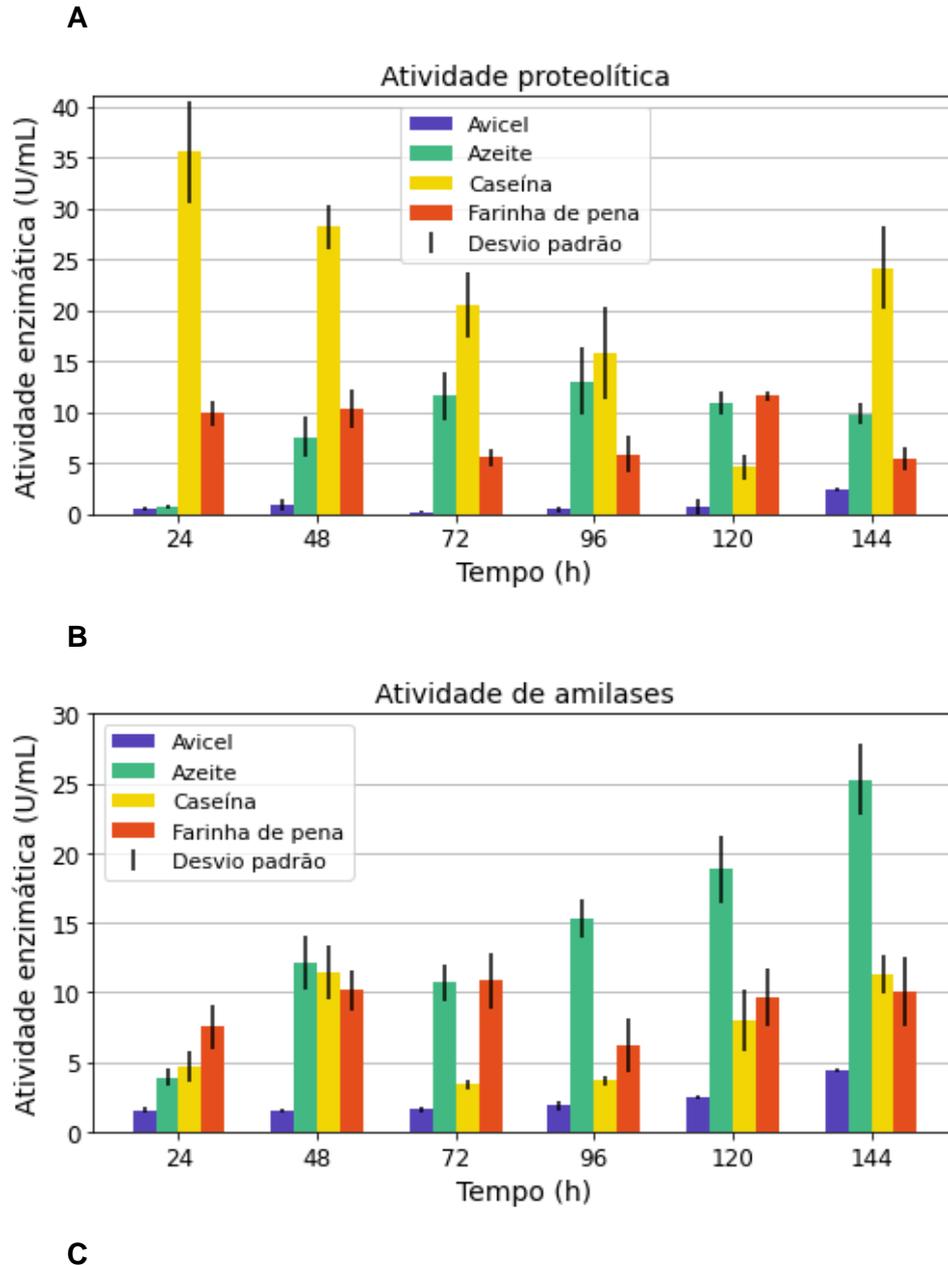
4.1. Atividade enzimática do filtrado de cultivo

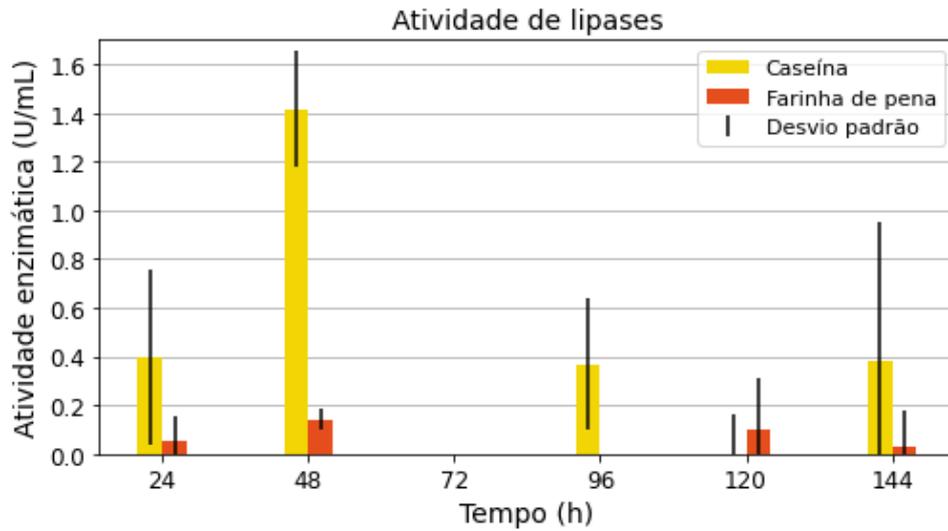
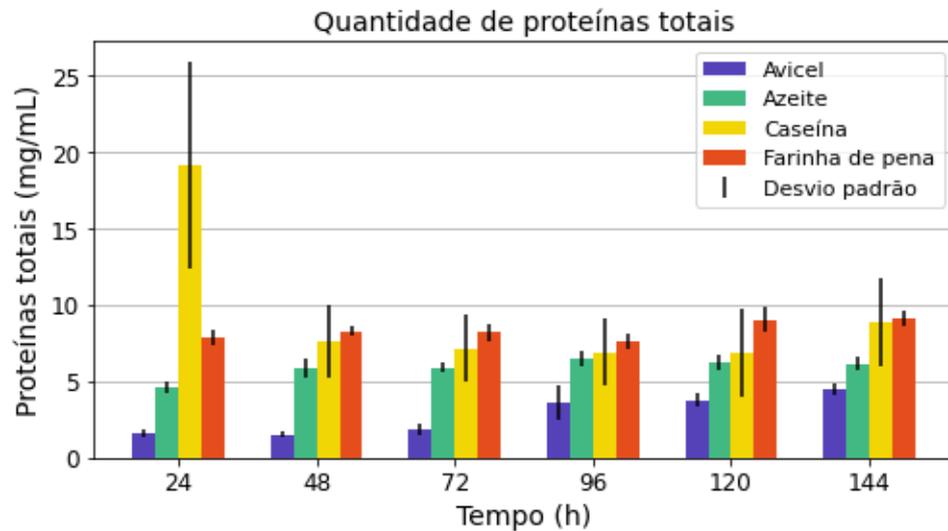
Os perfis de atividade enzimática com o filtrado de cultura do *R. miehei* foram diferentes tanto para o tipo de atividade quanto para a suplementação do meio de cultivo. A maior atividade de peptidases foi $35,51 \text{ U.mL}^{-1}$ para as amostras do cultivo suplementado com caseína no tempo 24 h enquanto as amostras do meio suplementado com Avicel apresentaram atividade baixa para todos os tempos de cultivo (Figura 4 A). A maior atividade de amilases foi determinada no meio suplementado com azeite de oliva, com $25,24 \text{ U.mL}^{-1}$ em 144 h de cultivo. Novamente o meio suplementado com Avicel apresentou o pior desempenho em todos os tempos de cultivo (Figura 4 B) em comparação com as demais condições. Apenas nos meios suplementados com caseína e farinha de pena foi detectado atividade de lipases. As atividades encontradas foram de $1,41 \text{ U.mL}^{-1}$ e $0,14 \text{ U.mL}^{-1}$ para caseína e farinha de pena, respectivamente, em 48 h (Figura 4 C). Não foram detectadas as atividades das enzimas xilanases, endoglucanases e lacases nos filtrados de cultura.

As concentrações de proteínas totais (Figura 4 D) presentes nos filtrados de cultura foram semelhantes entre os meios contendo azeite de oliva, caseína e farinha de pena a partir do tempo 48 h, e crescente para o meio com Avicel em relação ao

tempo do cultivo. Entre os tempos 24 h e 48 h houve redução na quantidade de proteínas totais no meio de cultivo contendo caseína de $19,16 \text{ mg.mL}^{-1}$ para $7,61 \text{ mg.mL}^{-1}$.

Figura 4. Atividade enzimática do filtrado de cultura do fungo *R. miehei* suplementado com Avicel, azeite de oliva, caseína e farinha de pena em relação ao tempo de cultivo para atividade de (A) peptidases, (B) amilases, (C) lipases e (D) quantificação de proteínas



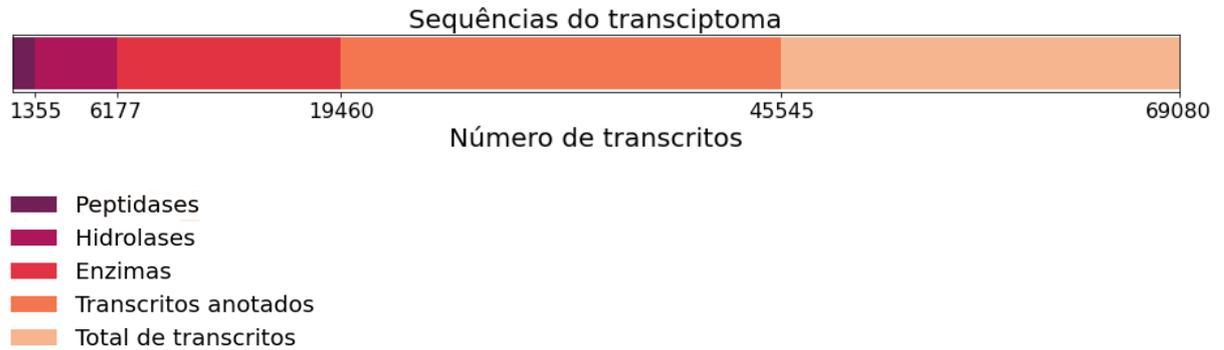
**D**

Fonte: Dados da pesquisa.

4.2. Mapeamento do transcriptoma

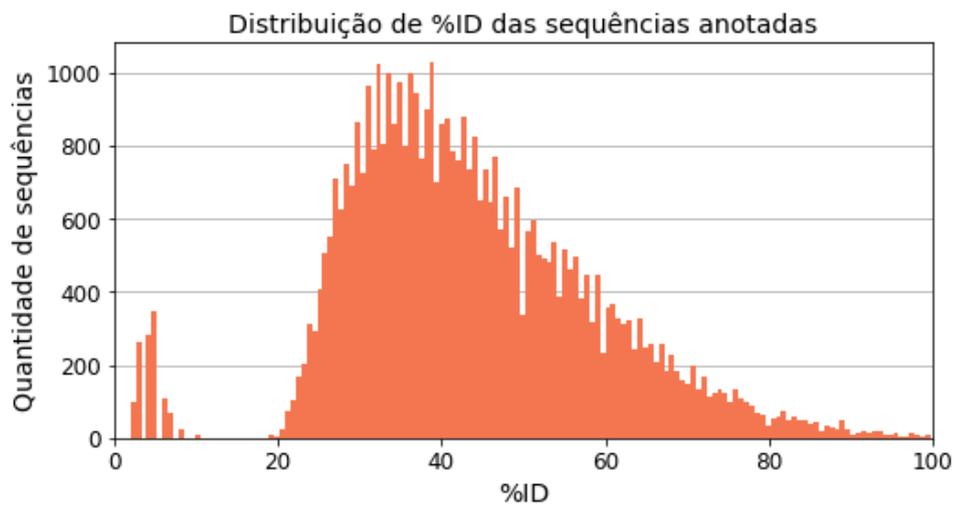
A construção do transcriptoma do fungo *R. miehei* gerou 69080 seqüências das quais 45545 apresentaram anotação funcional (Figura 5), com 11367 anotações diferentes, enquanto as outras 23535 não coincidiram com nenhuma proteína depositada e revisada no UniProtKB. A maioria das seqüências anotadas tiveram identidade com seqüências do banco de dados entre 25-50%, sendo o mínimo 2% e máximo maior que 99% (Figura 6).

Figura 5. Número de seqüências totais no transcriptoma de *R. miehei*, evidenciando valores de quantidade de transcritos anotados como de enzimas, hidrolases e peptidases.



Fonte: Dados da pesquisa.

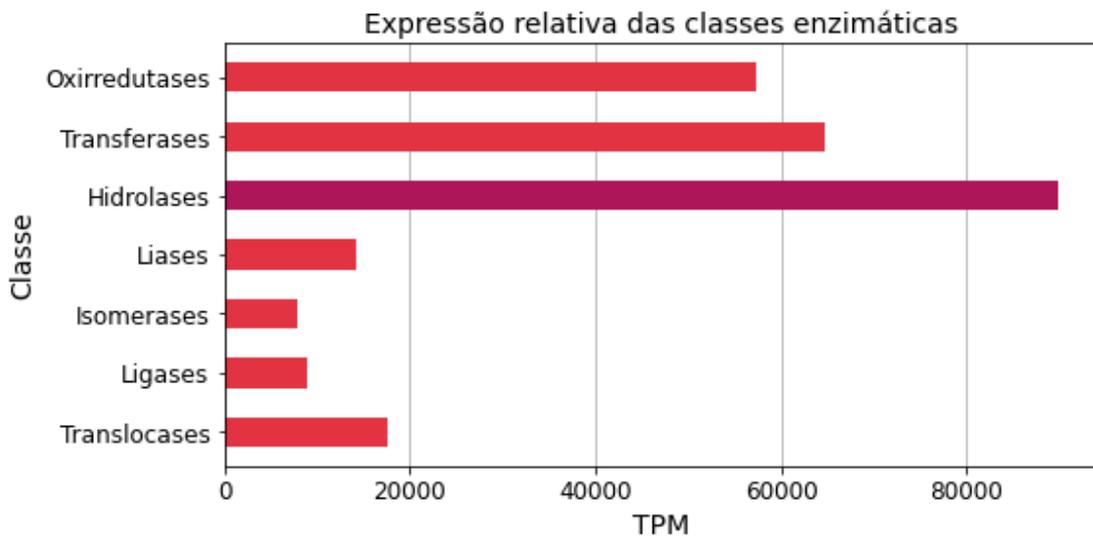
Figura 6. Distribuição da porcentagem de identidade das seqüências do transcriptoma com as seqüências revisadas do UniProtKB. ID: identidade.



Fonte: Dados da pesquisa.

O transcriptoma apresenta as sete classes de enzimas, sendo a maioria das anotações funcionais de transferases, hidrolases e oxirredutases, com 1778, 1650 e 826 anotações, respectivamente. Apesar disso, as hidrolases apresentaram o maior valor de expressão relativa (Figura 7).

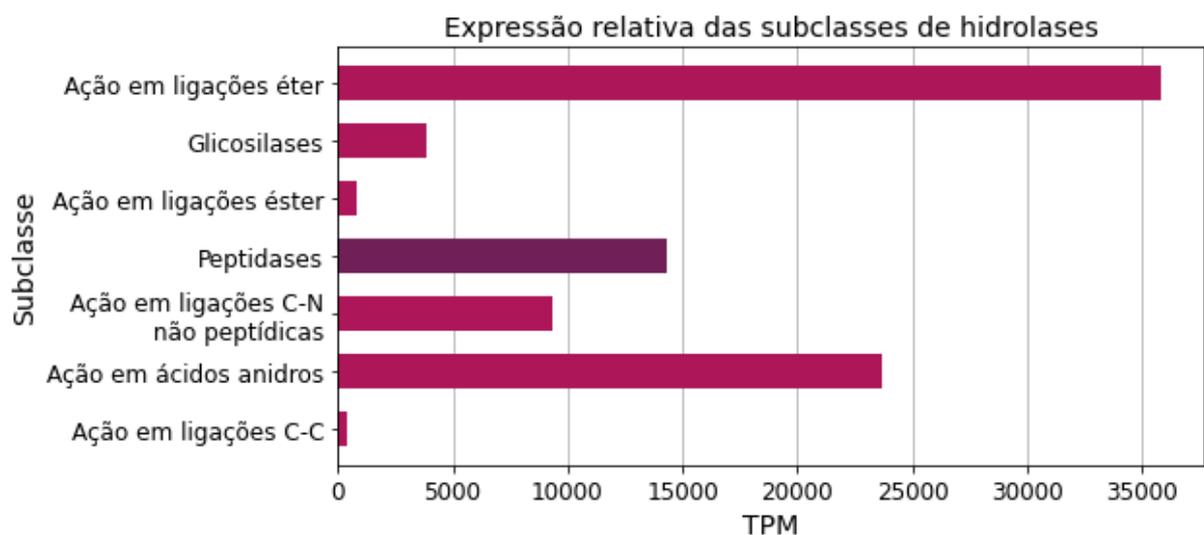
Figura 7. Expressão relativa das sete classes de enzimas do transcriptoma de *R. miehei*. Em destaque, a classe das hidrolases.



Fonte: Dados da pesquisa.

Aprofundando na classe das hidrolases, encontramos oito subclasses, das quais as peptidases foram a terceira maior representante, com 404 anotações, também com terceira maior expressão relativa (Figura 8). Apenas uma anotação para ação em ligações haleto (EC 3.8.1.2) com valor de expressão relativa próximo de 65, não apresentado na Figura 8.

Figura 8. Expressão relativa das subclasses de hidrolases do transcriptoma de *R. miehei*. Em destaque, a subclasse de peptidases. C: carbono. N: nitrogênio.



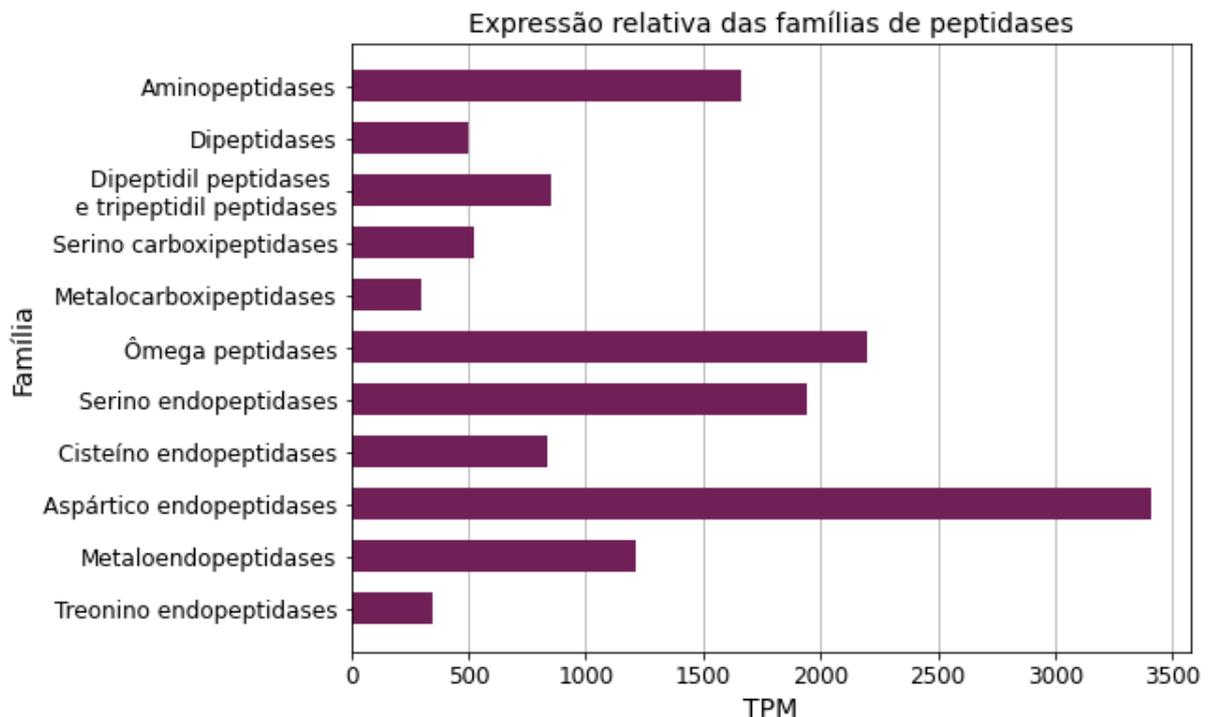
Fonte: Dados da pesquisa.

A baixa expressão de genes com ação em ligações éster e de glicosilases pode

explicar a baixa de atividade lipolítica e ausência de atividade de xilanases dos filtrados de cultivo. Apesar da alta expressão de genes para oxirredutases, não foi detectado atividade de lacases no filtrado de cultura, possivelmente por serem outros tipos de oxirredutases e pela metodologia de atividade.

A variedade de famílias de peptidases e seus respectivos valores de expressão relativa são mostrados na Figura 9. Foram encontradas anotações funcionais de 11 famílias e subfamílias, sendo as aspártico endopeptidases de maior abundância.

Figura 9. Expressão relativa das famílias de peptidases do transcriptoma de *R. miehei*.



Fonte: Dados da pesquisa.

4.3. Aspártico peptidase

4.3.1. Seleção e análise da sequência proteica

Foram identificadas quatro sequências homólogas anotadas como rizopuspepsina de *Rhizopus chinensis* (CARP_RHICH) e rizopuspepsina-I de *Rhizopus niveus* (CARP_RHINI), todas com indicativo de peptídeo-sinal. Essas enzimas são da família A1 tipo A01.012. Na Tabela 1, constam os nomes das sequências no transcriptoma atribuídos na construção da biblioteca pelo Trinity e

respectivos códigos UniProtKB atribuídos pelo Trinotate, valores de identidade (ID) entre a sequência do transcriptoma e a do banco de dados, e os valores de TPM, que indicam abundância relativa dos transcritos.

Tabela 1. Identificação das sequências da aspártico peptidase do transcriptoma do *R. miehei* e os respectivos valores de identidade (ID) e TPM. Os nomes das sequências foram atribuídos pelo programa Trinity na construção da biblioteca e os códigos UniProtKB pelo programa Trinotate. Em destaque, a sequência com maior TPM.

Nome da sequência	Código UniProtKB	ID (%)	TPM
TRINITY_DN16840_c0_g1_i1	CARP_RHICH	94,7	4,6
TRINITY_DN16840_c0_g1_i2	CARP_RHICH	95,9	2,5
TRINITY_DN16840_c0_g1_i3	CARP1_RHINI	81,5	11,7
TRINITY_DN16840_c0_g1_i4	CARP1_RHINI	82,3	5,7

Fonte: Dados da pesquisa.

A sequência escolhida para expressão heteróloga foi TRINITY_DN16840_c0_g1_i3 por apresentar maior valor de TPM. O alinhamento dos quatro peptídeos com a sequência escolhida e a sequência de CARP1_RHINI, mostrado Figura 10, foi feito com a ferramenta de alinhamento do UniProt.

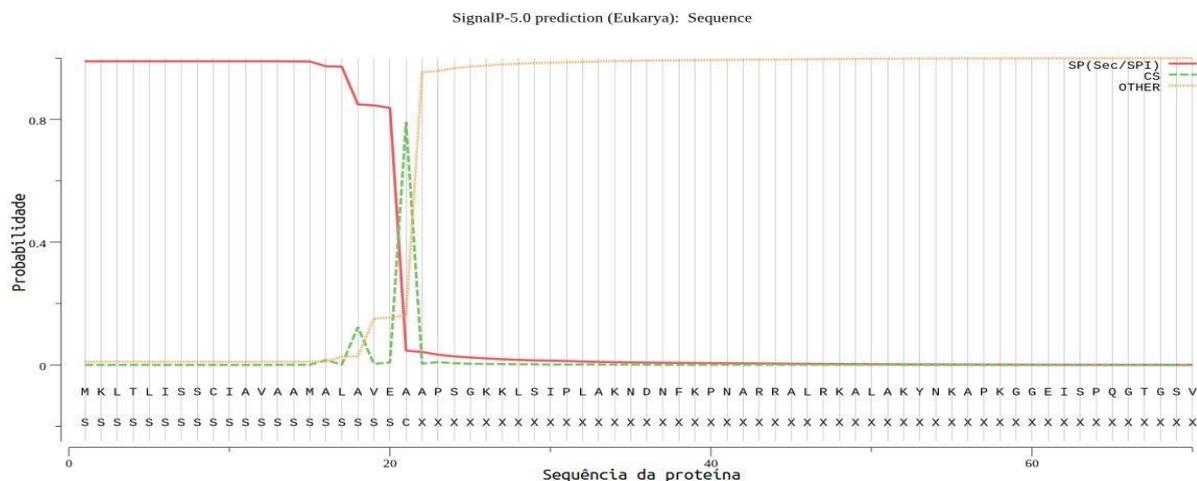
Figura 10. Alinhamento entre as sequências de TRINITY_DN16840_c0_g1_i3 e CARP1_RHINI com os peptídeos de referência. Foi destacada a similaridade entre os aminoácidos.

TRINITY_DN16840	1	MKLTLISSCIAVAAMALAVEAAPSGKKLSIPLAKNDNFKNARRALRKALAKYNKAPKG-	59
CARP1_RHINI	1	MKFTLISSCVALAAMTLAVEAAPNGKKINIPLAKNNSYKPSAKNALNKALAKYNRRKVGSG	60
Peptídeos	1	-----	0
TRINITY_DN16840	60	GEISPQGTGSPVVDVVGSDLEYGIVTVGTPGVKLLKDFDTGSSDLWFASTLCSSCSNSH	119
CARP1_RHINI	61	GGITTEASGSPVMVDYENDVEYYGEVTVGTPGIKLLKDFDTGSSDMWFASTLCSSCSNSH	120
Peptídeos	1	-----	0
TRINITY_DN16840	120	TKYNPNKSSTYKPDGRPWSISYGDGSSASGILAKDTVNLGGLVIKQDTIGLAKRESSAFA	179
CARP1_RHINI	121	TKYDPKKSSTYAADGRTWISISYGDGSSASGILATDNVNLGGLLIKKQTIELAKRESSAFA	180
Peptídeos	1	-----TWSISYGDGSSASGILAK-----	18
TRINITY_DN16840	180	EDVIDGLLGLAFNSITTVRGVKTVPVDNLISQGLISRPIYGVYLGKASNGGGGEYIFGGYD	239
CARP1_RHINI	181	TDVIDGLLGLGFNTITTVRGVKTVPVDNLISQGLISRPIYGVYLGKQSNNGGGGEYIFGGYD	240
Peptídeos	19	-----ASNGGGGEYIFGGYD-----	33
TRINITY_DN16840	240	STKFKGSLTTVPIDNSEGFVGVTVKSTKIGGKSIGGSFSAILDGTGTTTTLLLPDVAASVA	299
CARP1_RHINI	241	SSKFKGSLTTVPIDNSEGFVGVTVKSTKIGGTTVSASFDAILDGTGTTTTLLLPDDVAAKVA	300
Peptídeos	34	STK--GSLTTVPIDNSRGVWGITVDRA-----	58
TRINITY_DN16840	300	RTYGATDNGDGYTITCDASKLQPLVFTLGSSTFQVPADSLIYEKSGSQCIAGFAAGGDL	359
CARP1_RHINI	301	RSYGASDNGDGYTISITCDTSKLLQPLVFTLGSSTFQVPADSLIFEKDGKNCIAGFAAGGDL	360
Peptídeos	59	-----	58
TRINITY_DN16840	360	AILGDVFLKNNYVVFNQVPEVQIAPVAH	388
CARP1_RHINI	361	AILGDVFLKNNYVVFNQEVPEVQIAPVAN	389
Peptídeos	59	-----	58

Fonte: UniProt Align (BATEMAN et al., 2021).

A identificação do peptídeo-sinal foi feita pelo servidor SignalP-5.0 e é ilustrada na Figura 11. A probabilidade dos aminoácidos 1-21 ser peptídeo-sinal é de 98,9% e de clivagem entre 21-22 é de 79,1%. O peptídeo-sinal foi desconsiderado na construção do vetor de expressão da rRmAspPep.

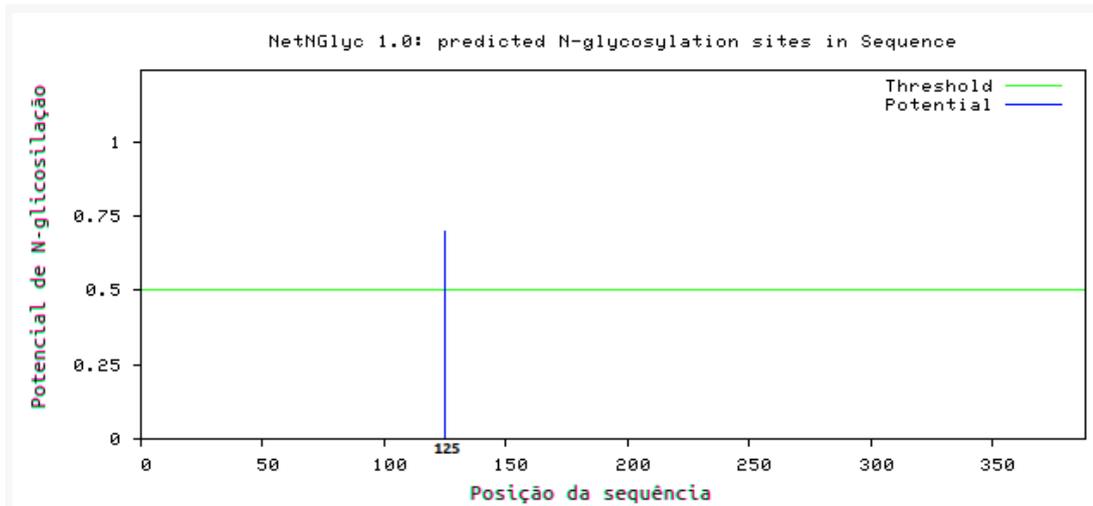
Figura 11. Predição de presença de peptídeo-sinal e localização do sítio de clivagem da sequência TRINITY_DN16840_c0_g1_i3. A relação entre probabilidade e posição e natureza do aminoácido é mostrado pelas linhas no gráfico. Em vermelho, a probabilidade de ser peptídeo-sinal (SP); em verde, a probabilidade de ser sítio de clivagem (CS); e em amarelo, a probabilidade de ser outro tipo de sequência (OTHER).



Fonte: SignalP-5.0 Server (Nielsen et al., 2019).

Na Figura 12 é representada a predição de N-glicosilação da sequência de TRINITY_DN16840_c0_g1_i3, sendo que o resíduo de Asn-125 apresentou potencial de 69,9% para N-glicosilação. Para O-glicosilação foram identificadas oito posições (95, 99, 107, 108, 174, 280 e 284) com potencial maior que 50%. A estabilidade térmica de uma proteína pode ser relacionada com sua glicosilação. Enquanto a maioria dos genes de peptidases apresentam vários sítios de N-glicosilação (ZHOU et al., 2014) a RmAspPep apresentou apenas um local.

Figura 12. Predição de N-glicosilação na sequência TRINITY_DN16840_c0_g1_i3, relacionando potencial de N-glicosilação e posição da sequência. A linha vertical azul indica a probabilidade (*Potential*) do resíduo de asparagina da sequência ser N-glicosilado. A linha horizontal verde indica o *threshold* de 50%.

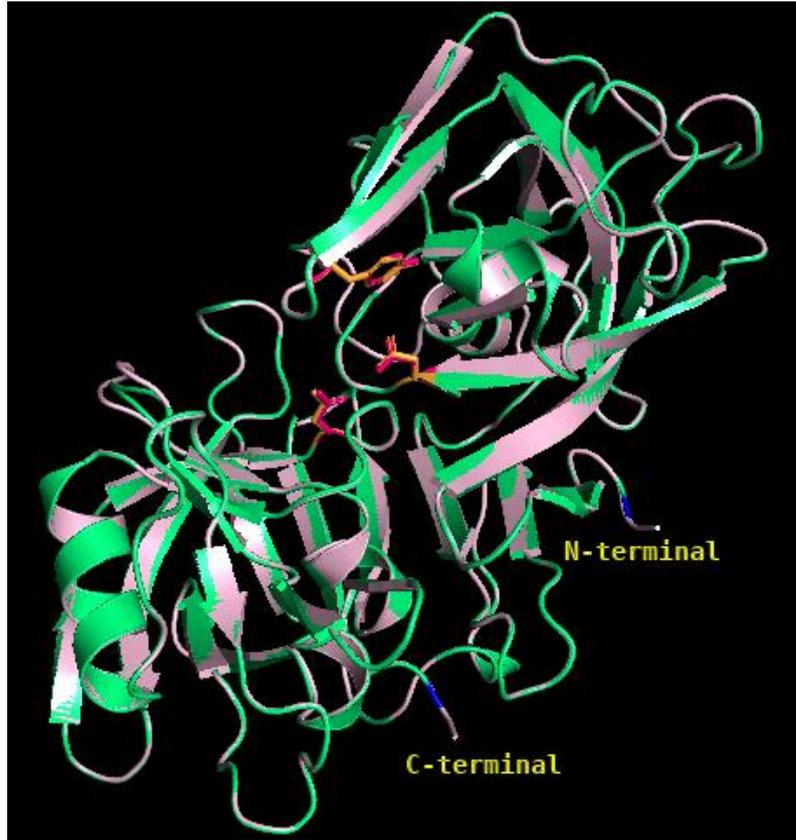


Fonte: NetNGlyc 1.0 Server (GUPTA; BRUNAK, 2002).

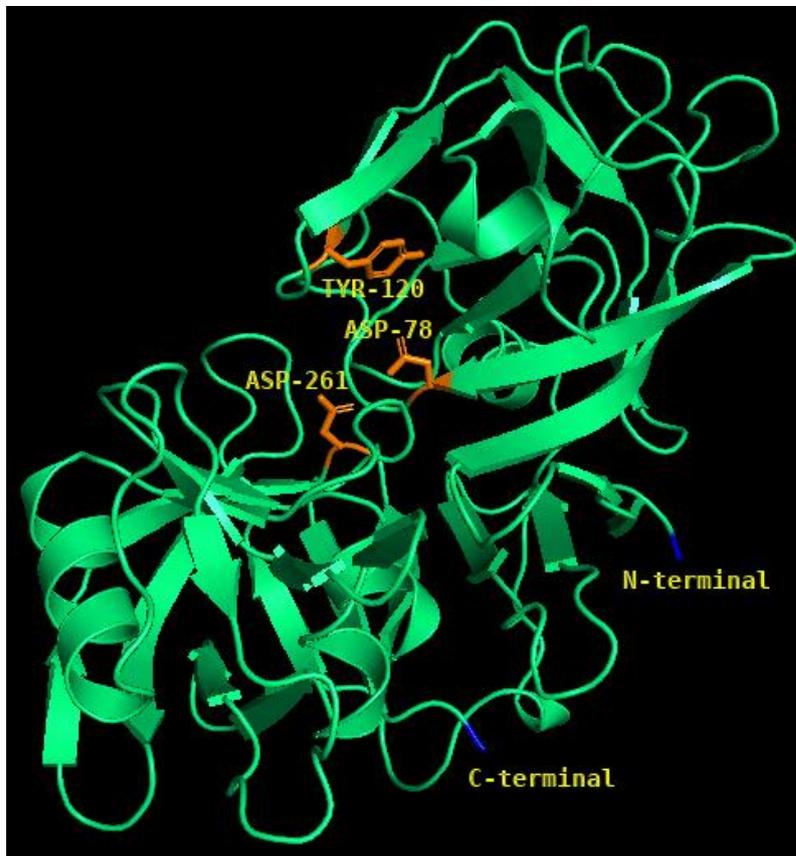
Foi feita modelagem da estrutura da sequência de TRINITY_DN16840_c0_g1_i3 pelo servidor SWISS-MODEL, o qual utilizou a rhizopuspepsina madura (entrada do banco de dados de proteínas, PDB, 1uh9) de *Rhizopus microsporus* var. *chinensis* (PRASAD; SUGUNA, 2003) como modelo (Figura 13). Na Figura 13 A é mostrado o alinhamento entre o modelo criado e a estrutura molde com valor de QMEAN de 0,89, indicando boa qualidade do modelo. Os possíveis aminoácidos catalíticos (Asp-78, Tyr-120 e Asp-261) foram identificados no modelo a partir de comparações com a estrutura molde e informações sobre outras proteínas da mesma família do banco de dados Merops. Foi identificado o propeptídeo (aminoácidos 22-65), que foi mantido na construção da rRmAspPep.

Figura 13. (A) Representação da estrutura tridimensional da RmAspPep, em verde, alinhada a estrutura molde, em rosa claro. As formas espirais representam α -hélices e as setas representam folhas- β . Os grupos laterais dos aminoácidos da tríade catalítica estão destacados em laranja para a RmAspPep e em magenta para o molde. Os resíduos em azul correspondem ao N- e C-terminal da RmAspPep. (B) Semelhante a A, o modelo da RmAspPep. (C) Representação da superfície da RmAspPep, destacando o sítio catalítico.

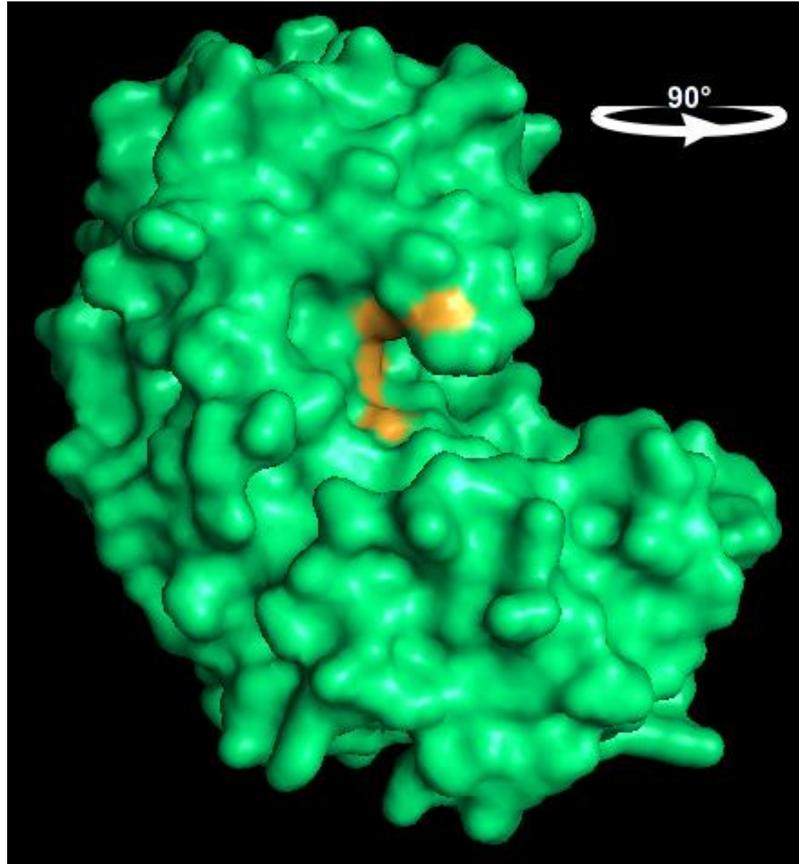
A



B



C



Fonte: Autor.

O vetor de expressão para rRmAspPep foi construído com a sequência dos aminoácidos 22-388 de TRINITY_DN16840_c0_g1_i3 mais cinco resíduos de histidina. A massa molecular estimada para rRmAspPep é de 34.3 kDa e pI 5,6, considerando maturação adequada (perda do propeptídeo).

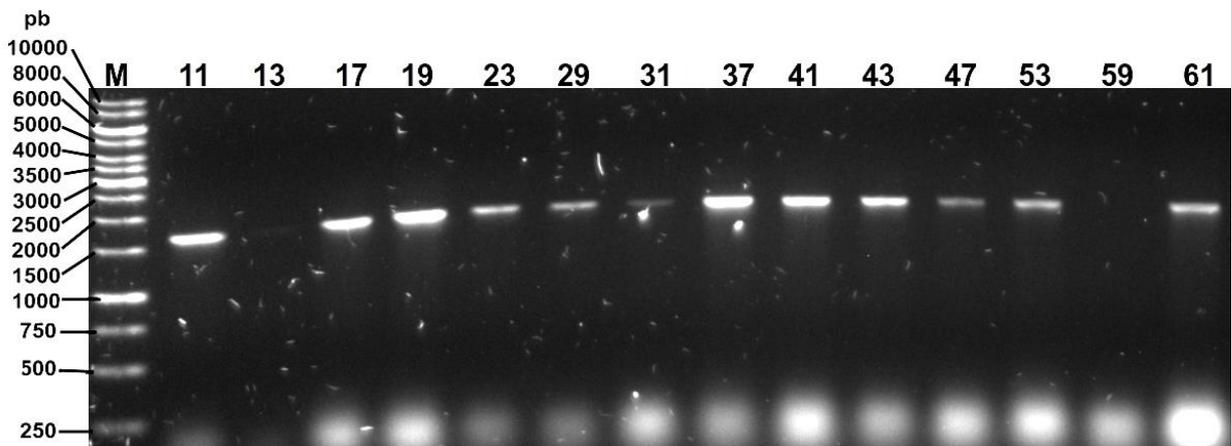
4.3.2. Clonagem em *P. pastoris* KM71H

Após a etapa de transformação, foram isoladas e numeradas 65 colônias de *P. pastoris* KM71H/rRmAspPep que cresceram em meio seletivo YPDS ágar com zeocina a $100 \mu\text{g}\cdot\text{mL}^{-1}$. As colônias escolhidas arbitrariamente para as etapas da triagem foram as de número 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59 e 61 para rRmAspPep.

A Figura 14 é resultado da PCR das colônias de *P. pastoris* KM71H/rRmAspPep recuperadas após a etapa de indução da expressão enzimática. O tamanho entre os sítios de anelamento dos *primers* 5'AOX1 e 3'AOX1 para a rRmAspPep é de 1633 pb. Foi possível ver bandas alinhadas entre 1500-2000 pb

confirmando a presença do inserto, mas há diferença na intensidade das bandas entre as colônias analisadas. As bandas das colônias 11, 17, 19, 37, 41 e 43 são mais notáveis que as bandas das colônias 13, 23, 29, 31, 47, 53 e 61, e não foi possível visualizar banda referente ao inserto para a colônia 59.

Figura 14. Representação fotográfica da eletroforese em gel de agarose 1% do produto de PCR de colônia depois da etapa de indução de expressão enzimática. Os números das raias correspondem ao número das colônias *P. pastoris* KM71H/rRmAspPep. À esquerda, valores de pb para cada banda do marcador (M) GeneRuler 1 kb DNA ladder (Thermo Fisher).



Fonte: Dados da pesquisa.

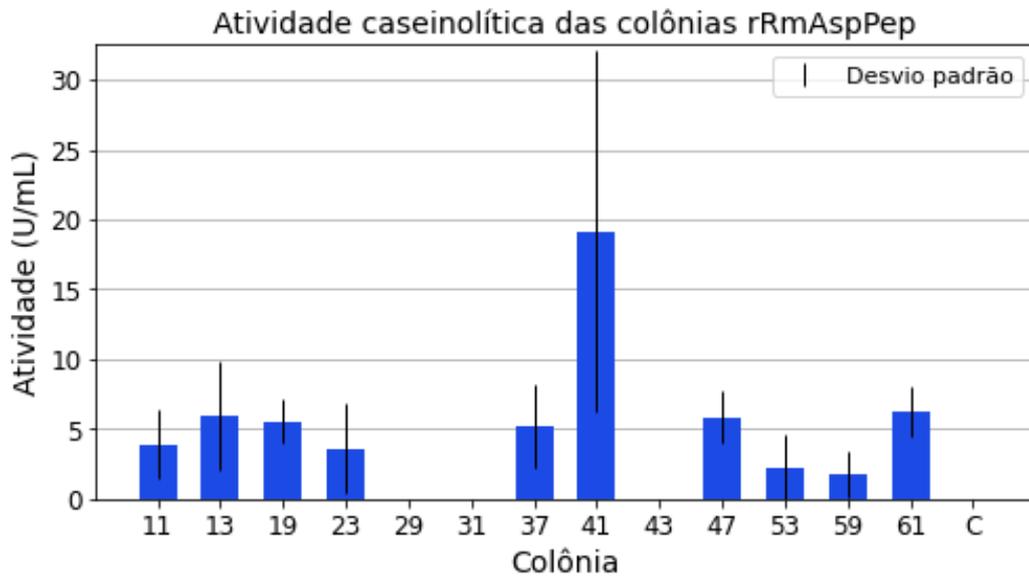
4.3.3. Análise da expressão da rRmAspPep

O ensaio de atividade proteolítica foi feito com as amostras do meio de cultivo das colônias de *P. pastoris* KM71H/rAspPep em 144 h de indução. Como controle, foi utilizado amostra do meio de cultivo de *P. pastoris* KM71H do controle na indução.

Os valores de atividade (U.mL^{-1}) foram parecidos entre as colônias que apresentaram atividade (Figura 15). A maior atividade foi de $11,7 \text{ U.mL}^{-1}$ encontrada para a colônia 41, e a menor atividade foi de $2,6 \text{ U.mL}^{-1}$, para a colônia 59. As colônias

29, 31 e 43 não apresentaram atividade, assim como o controle.

Figura 15. Atividade proteolítica do meio de cultivo das colônias de *P. pastoris* KM71H/rRmAspPep em 144 h de indução. C: controle.



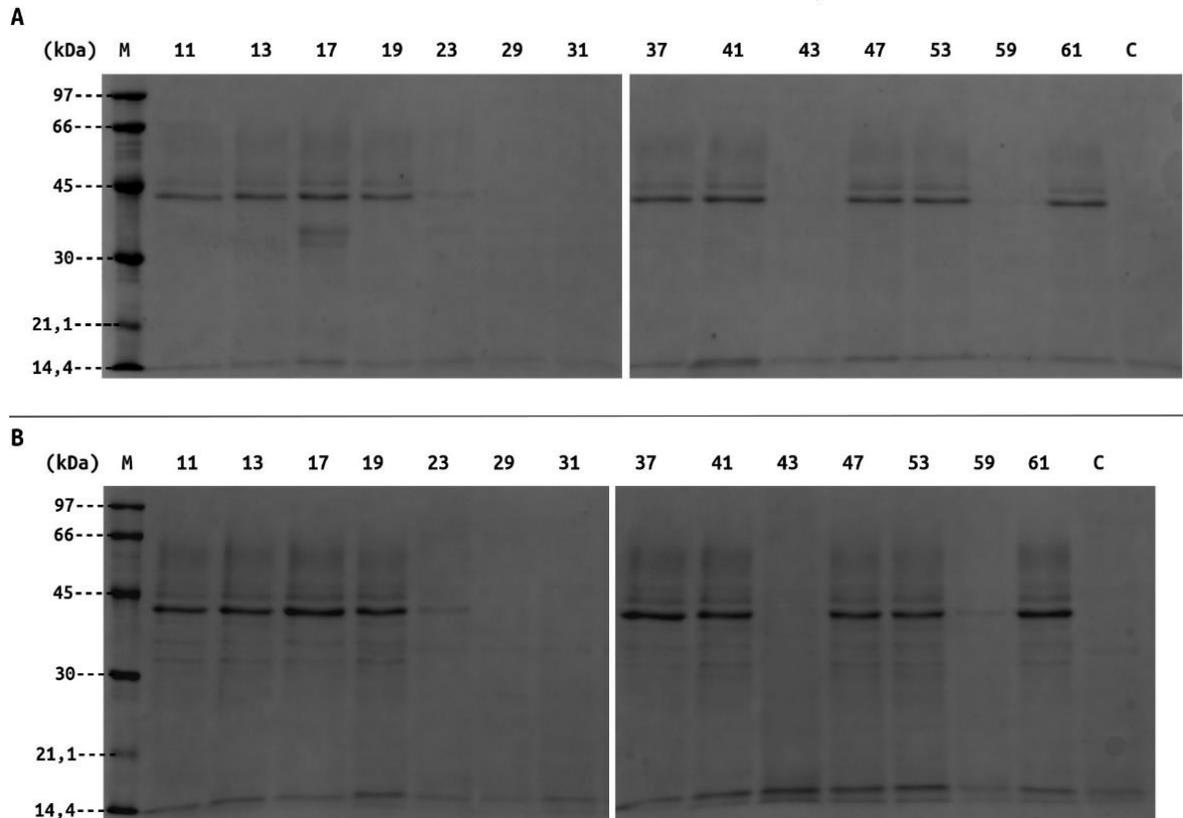
Fonte: Dados da pesquisa.

A análise de SDS-PAGE das colônias de *P. pastoris* KM71H/rRmAspPep apresentou padrão nas bandas entre as colônias e tempo de indução. No primeiro ponto da indução (24 h) foi possível notar duas bandas para as colônias 11, 13, 17, 19, 37, 41, 47, 53 e 61, uma alinhada em 45 kDa e a outra, mais intensa, alinhada entre 30-45 kDa. As mesmas bandas só nas colônias 23 e 59, surgiram a partir de 72 h de indução. O padrão dessas duas bandas foi mantido em todos os pontos da indução. Essas bandas não foram notadas no controle e nas colônias 29, 31 e 43 em nenhum ponto da indução.

Na Figura 16 são mostrados os géis dos pontos de 72 h e 144 h da indução. No ponto de 72 h, pode-se notar outras duas bandas alinhadas entre 30-45 kDa para a colônia 17 que não aparecem para as outras colônias, inclusive no controle. Já nos géis de 144 h de indução, essas duas novas bandas podem ser vistas, novamente para as colônias 11, 13, 17, 19, 37, 41, 47, 53 e 61. Nesse ponto, uma dessas bandas (a maior) apareceu no controle.

Figura 16. SDS-PAGE das colônias de *P. pastoris* KM71H/rRmAspPep e controle nos pontos de (A) 72 h e (B) 144 h de indução. As raias estão identificadas de acordo com o número da

colônia e controle (C). À esquerda, valores da massa molecular (kDa) referentes às bandas do marcador (M) LMW-SDS Marker Kit (GE Healthcare). Corado com Comassie Blue Phast.



Fonte: Dados da pesquisa.

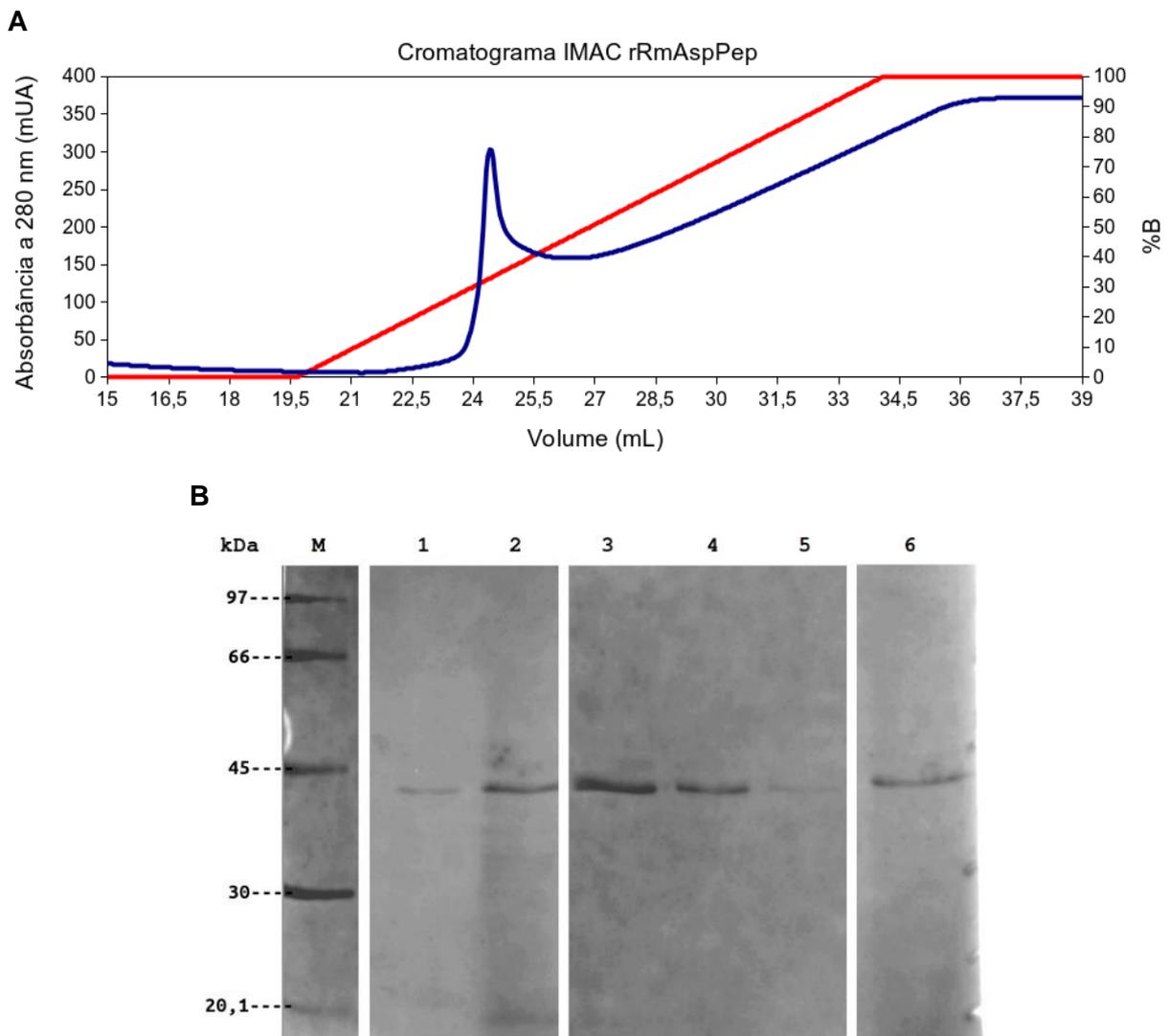
O a ausência de bandas visíveis no SDS-PAGE coincide com a ausência de atividade proteolítica para as colônias 29, 31 e 43, mesmo que confirmada a presença do inserto por PCR de colônia. O oposto aconteceu com a colônia 59, na qual foi identificada atividade e banda no SDS-PAGE, enquanto não foi detectado o inserto na PCR de colônia.

4.3.4. Purificação

A amostra de 144 h de indução da colônia 17 de *P. pastoris* KM71H/rRmAspPep passou por purificação por meio de cromatografia de afinidade. As frações da eluição foram analisadas por SDS-PAGE (Figura 17 B). Na imagem, as raias 1-5 correspondem às frações da eluição no momento do pico da curva de absorvância (280 nm) no cromatograma (Figura 17 A). As bandas fortes das raias 1-5 são alinhadas entre 30 e 45 kDa, assim como a banda da raia 6, proveniente da mesma amostra, mas sem purificação. Há também uma banda menos intensa alinhada em 45 kDa, visível nas raias 2-4. A banda visível no SDS-PAGE da

purificação corresponde a banda mais evidente nos géis da análise de expressão.

Figura 17. (A) Cromatograma da amostra da colônia 17, as faixas verticais indicam as frações usadas no SDS-PAGE. A linha azul corresponde aos valores de absorbância (esquerda) e a linha vermelha à porcentagem da solução de eluição (direita). **(B)** SDS-PAGE do meio de cultivo da colônia 17 de *P. pastoris* KM71H/rRmAspPep purificado (raias 1-5) e não purificado (raia 6) no ponto de 144 h da indução de expressão enzimática. A imagem foi montada apenas com as raias que apresentaram bandas, respeitando a ordem das frações da cromatografia. À esquerda, valores da massa molecular (kDa) referentes às bandas do marcador (M) LMW-SDS Marker Kit (GE Healthcare). Corado com Comassie Blue Phast.



Fonte: Dados da pesquisa.

4.4. Carboxipeptidase

4.4.1. Seleção e análise da sequência proteica

No mapeamento do transcriptoma foram encontradas anotações para 14 serino carboxipeptidases. Foi feita busca textual usando os códigos do UniProtKB para identificar as sequências nos dados de anotação funcional do transcriptoma. Com nome de identificação dessas sequências foi possível verificar o valor de TPM e valor de identidade entre a sequência do *R. miehei* (Tabela 2). A sequência escolhida para expressão heteróloga foi a de maior abundância relativa, a qual tem identidade de 46% com sequência da proteína CBPYA_LEPMJ (carboxipeptidase Y homologa A, *Leptosphaeria maculans*, linhagem JN3 / isolado v23.1.3 / raça Av1-4-5-6-7-8).

Tabela 2. Identificação das sequências de serino carboxipeptidases do transcriptoma do *R. miehei* e os respectivos valores de identidade (ID) e de TPM. Os nomes das sequências foram atribuídos pelo programa Trinity na construção da biblioteca e os códigos UniProtKB pelo programa Tinotate. Em destaque, a sequência com maior TPM.

Nome da sequência	Código UniProtKB	ID (%)	TPM
TRINITY_DN16548_c0_g1_i1	CBPYA_SCLS1	46	18,7
TRINITY_DN16548_c0_g1_i2	CBPYA_ARTOC	45,7	2,9
TRINITY_DN16548_c0_g1_i3	CBPY_CANAX	46	42,1
TRINITY_DN11912_c0_g1_i1	CBPYA_PHANO	42,3	8,3
TRINITY_DN11912_c0_g1_i2	CBPYA_TALSN	46	2,2
TRINITY_DN14171_c0_g1_i1	KEX1_CHAGB	46,9	18,5
TRINITY_DN14171_c0_g1_i2	KEX1_PYRTR	37,8	0,9
TRINITY_DN12590_c0_g1_i1	KEX1_MAGO7	46,9	1,5
TRINITY_DN12590_c0_g1_i2	KEX1_MAGO7	42,8	2,9
TRINITY_DN12590_c0_g2_i1	KEX1_MAGO7	45,5	3,7
TRINITY_DN12590_c0_g2_i2	KEX1_VERA1	41,5	2,1
TRINITY_DN5894_c0_g1_i1	CBPYA_LEPMJ	46	124,8
TRINITY_DN5891_c0_g1_i1	KEX1_UNCRE	46,3	2,1
TRINITY_DN5891_c0_g1_i2	KEX1_UNCRE	46,4	1,7
TRINITY_DN2754_c0_g1_i1	CPS1_PENJA	36	4,1
TRINITY_DN18925_c0_g1_i1	CBPY_CANAX	54,9	78,4
TRINITY_DN18925_c0_g1_i2	CBPY_CANAX	54,7	75,6
TRINITY_DN15941_c0_g1_i1	CBPY_CANAX	56,2	62
TRINITY_DN15941_c0_g1_i2	CBPY_SCHPO	60,9	0
TRINITY_DN15941_c0_g1_i3	CBPY_CANAX	55,9	13,7

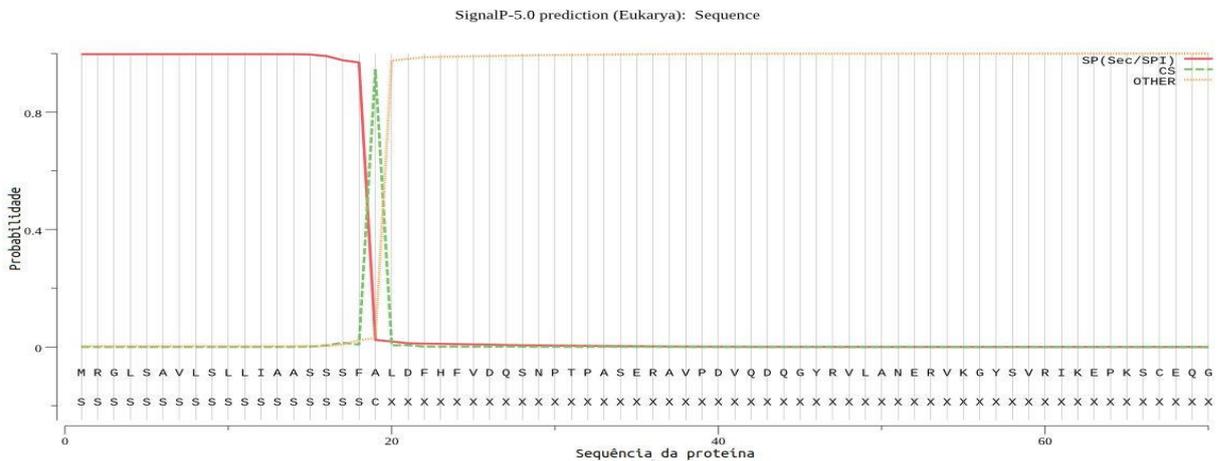
TRINITY_DN15941_c0_g1_i4	CBPY_CANAX	58,8	0
TRINITY_DN15941_c0_g1_i5	CBPY_CANAX	55,4	0,4
TRINITY_DN15941_c0_g1_i6	CBPY_YEAST	53,7	0,1
TRINITY_DN15941_c0_g1_i7	CBPY_SCHPO	60,9	8,9
TRINITY_DN15941_c0_g1_i8	CBPY_CANAX	57,6	48,3

Fonte: Dados da pesquisa.

A enzima carboxipeptidase Y homologa A é da família S10 do tipo S10.001, documentada nos bancos de dados Merops e UniProtKB, é reportada como uma enzima de atuação no vacúolo.

A predição da presença de peptídeo sinal e local do sítio de clivagem na sequência feito no programa SignalP-5.0 é representado na Figura 18. A probabilidade da sequência dos aminoácidos 1-19 ser peptídeo-sinal é de 99,7%, e a probabilidade de clivagem entre os aminoácidos 19-20 é de 94,5%.

Figura 18. Predição de presença de peptídeo-sinal e localização do sítio de clivagem da sequência TRINITY_DN5894_c0_g1_i1. A relação entre probabilidade e posição e natureza do aminoácido é mostrado pelas linhas no gráfico. Em vermelho, a probabilidade de ser peptídeo sinal (SP); em verde, a probabilidade de ser sítio de clivagem (CS); e em amarelo, a probabilidade de ser outro tipo de sequência (OTHER).

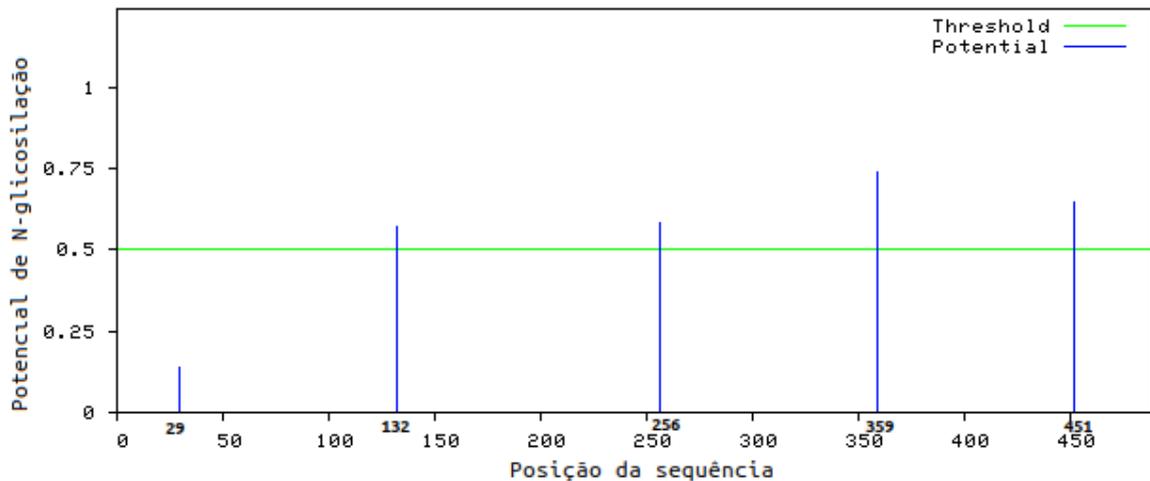


Fonte: SignalP-5.0 Server (NIELSEN et al., 2019).

A predição para N-glicosilação (Figura 19) apresentou quatro posições (132, 256, 359 e 451) com potencial maior que 50% e para O-glicosilação, três posições (9, 12 e 15) com potencial maior que 80%.

Figura 19. Predição de N-glicosilação na sequência TRINITY_DN5894_c0_g1_i1, relacionando potencial de N-glicosilação e posição da sequência. As linhas verticais em azul

indicam a probabilidade (Potential) do resíduo de asparagina da sequência ser N-glicosilado. A linha horizontal verde indica o threshold de 50%.



Fonte: NetNGlyc 1.0 Server (GUPTA; BRUNAK, 2002).

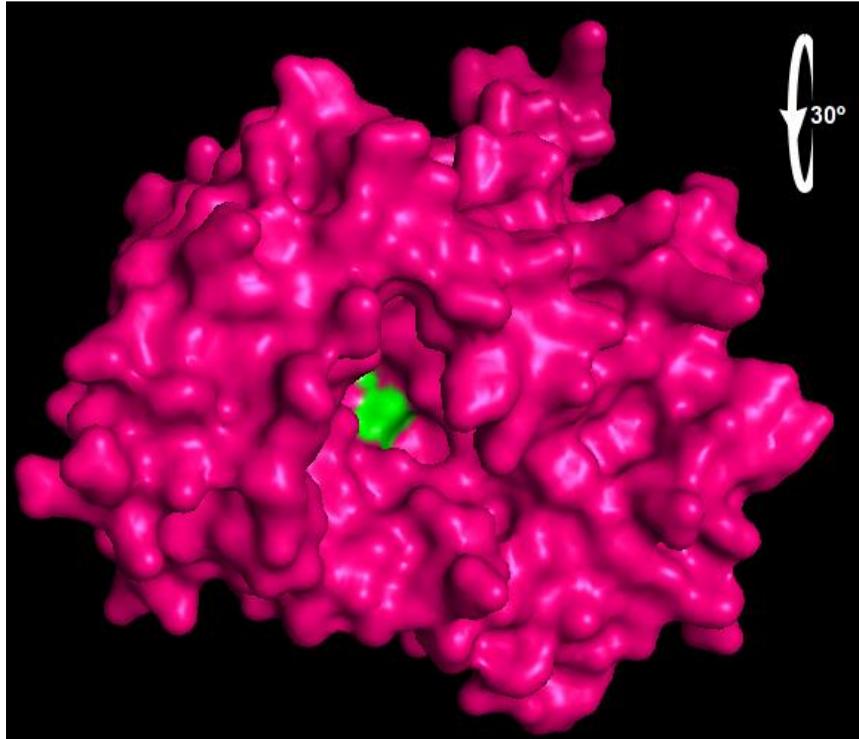
O modelo para RmCarPep foi feito a partir a estrutura da carboxipeptidase Y madura (entrada do PDB 1wpx) de *Saccharomyces cerevisiae* (MIMA et al., 2005). Na Figura 20 é ilustrado o alinhamento entre as duas sequências, que apresentou com valor QMEAM de 0,77, indicando boa qualidade do modelo. Os possíveis aminoácidos catalíticos (Ser-141, Asp-342 e His-403) foram identificados pela comparação entre as estruturas das duas sequências e com informações do Merops e UniProtKB e também os aminoácidos do propeptídeo de endereçamento para o vacúolo, 20-59.

As sequências do peptídeo-sinal e propeptídeo (aminoácidos 1 ao 59) foram desconsideradas para a construção do vetor de expressão. A RmCarPep recombinante (rRmCarPep) é composta pela sequência dos aminoácidos 60-488 mais seis resíduos de histidina, tendo massa molecular estimada de 49,3 kDa e pI 5.

Figura 20. (A) Representação da estrutura tridimensional da RmCarPep, em magenta, alinhada a estrutura molde, em verde claro. As formas espirais representam α -hélices e as setas representam folhas- β . Os grupos laterais dos aminoácidos da tríade catalítica estão destacados em verde para a RmCarPep e em ciano para o molde. Os resíduos em azul correspondem ao N- e C-terminal da RmAspPep. (B) Semelhante a A, o modelo da RmCarPep. (C) Representação da superfície da RmCarPep, destacando o sítio catalítico.

A

**B****C**



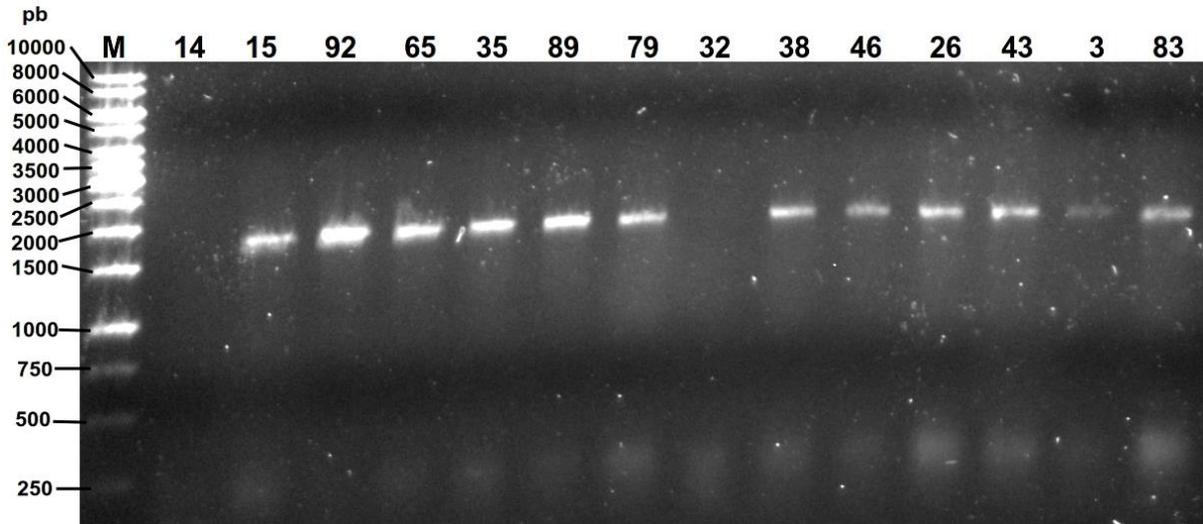
Fonte: Autor.

4.4.2. Clonagem em *P. pastoris* KM71H

Após a etapa de transformação, foram isoladas e numeradas 115 colônias de *P. pastoris* KM71H/rRmCarPep que cresceram em meio seletivo YPDS ágar com zeocina a $100 \mu\text{g}\cdot\text{mL}^{-1}$. As colônias escolhidas arbitrariamente para as etapas da triagem foram as de número 14, 15, 92, 65, 35, 89, 79, 32, 38, 46, 26, 43, 3 e 83 para rRmCarPep.

Após a etapa de indução da expressão, foi feita PCR de colônia para confirmação da transformação. O tamanho entre os sítios de anelamento dos *primers* 5'AOX1 e 3'AOX1 para a rRmCarPep é de 1822 pb. Na Figura 21 é apresentado o resultado para as colônias selecionadas. As colônias 14 e 32 não apresentaram banda referente ao inserto, enquanto para todas as outras colônias as bandas alinham entre 1500-2000 pb. Também houve diferença na intensidade, na qual a colônia 3 tem banda menos visível que as outras.

Figura 21. Eletroforese em gel de agarose 1% do produto de PCR de colônia depois da etapa de indução de expressão enzimática. Os números das raias correspondem ao número das colônias de *P. pastoris* KM71H/rRmCarPep. À esquerda, valores de pb para cada banda do marcador (M) GeneRuler 1 kb DNA *ladder* (Thermo Fisher).



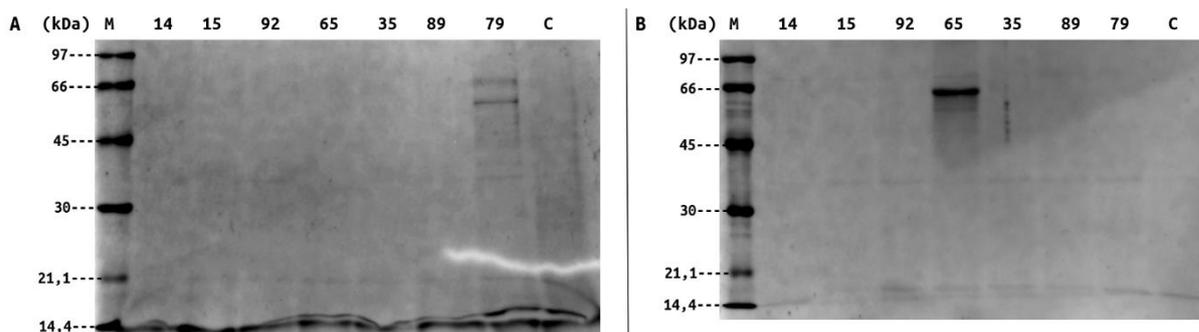
Fonte: Dados da pesquisa.

4.4.3. Análise da expressão da rRmCarPep

A análise de expressão foi feita com as amostras dos seis pontos da indução (24, 48, 72, 96, 120 e 144 h) para todas as colônias selecionadas. As amostras do meio de cultivo de *P. pastoris* KM71H sem inserto dos mesmos pontos da indução foram usadas como controle.

Os resultados de SDS-PAGE das colônias de *P. pastoris* KM71H/rRmCarPep não tiveram um padrão como visto para rRmAspPep. Apenas nos pontos de 48 h e 120 h foi possível visualizar bandas que não aparecem no controle. No gel de 48 h da indução podem ser vistas duas bandas alinhadas entre 45-66 kDa exclusivas para a colônia 79, enquanto no gel de 120 h de indução a colônia 65 apresentou uma banda alinhada entre 45-66 kDa (Figura 22).

Figura 22. SDS-PAGE das colônias de *P. pastoris* KM71H/rRmCarPep e controle nos pontos de **(A)** 48 h e **(B)** 120 h de indução. As raias estão identificadas de acordo com o número da colônia e controle (C). À esquerda, valores da massa molecular (kDa) referentes às bandas do marcador (M) LMW-SDS Marker Kit (GE Healthcare). Corado com Comassie Blue Phast.



Fonte: Dados da pesquisa.

5. Conclusões e perspectivas

Baseado nas informações do transcriptoma foram selecionados dois genes dos quais pela análise da sequência mostraram-se viáveis para a construção do vetor sintético. Os genes foram sintetizados e clonados mostrando a viabilidade da expressão em *P. pastoris*. Foi confirmada a secreção e purificação da rRmAspPep, o que não ocorreu no caso da rRmCarPep. Trabalhos futuros deverão concentrar-se na caracterização da rRmAspPep e na expressão e caracterização da rRmCarPep.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALI, Y. et al. Specificity of peptidases secreted by filamentous fungi. v. 9, n. 1, p. 30–37, 2018.
- ALVES, A. C. V. et al. The correction of reaction rates in continuous fluorometric assays of enzymes. **Journal of Biochemical and Biophysical Methods**, v. 70, n. 3, p. 471–479, abr. 2007.
- BANERJEE, G.; RAY, A. K. Impact of microbial proteases on biotechnological industries. **Biotechnology and Genetic Engineering Reviews**, v. 33, n. 2, p. 119–143, 2017.
- BATEMAN, A. et al. UniProt: The universal protein knowledgebase in 2021. **Nucleic Acids Research**, v. 49, n. D1, p. D480–D489, 2021.
- BINHAYEEDING, N. et al. Improvement of biodiesel production using waste cooking oil and applying single and mixed immobilised lipases on polyhydroxyalkanoate. **Renewable Energy**, v. 162, p. 1819–1827, 2020.
- BJELLQVIST, B. et al. Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions. **Electrophoresis**, v. 15, n. 1, p. 529–539, 1994.
- Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein-dye binding. **Analytical Biochemistry**, v. 72, p. 248–254, 1976.
- CORPORATION, P. Wizard SV Gel and PCR Clean-Up System Technical Bulletin, TB308. **Promega**, p. 1–13, 2010.
- COZENTINO, I. DE S. C. et al. Enzymatic synthesis of structured lipids from grape seed (*Vitis vinifera* L.) oil in associated packed bed reactors. **Biotechnology and Applied Biochemistry**, n. Mlm, p. 1–9, 2020.
- CREGG, J. M. et al. Recombinant protein expression in *Pichia pastoris*. **Applied Biochemistry and Biotechnology - Part B Molecular Biotechnology**, v. 16, n. 1, p. 23–52, 2000.
- DA SILVA, R. R. et al. Evaluation of the catalytic specificity, biochemical properties, and milk clotting abilities of an aspartic peptidase from *Rhizomucor miehei*. **Journal of Industrial Microbiology and Biotechnology**, v. 43, n. 8, p. 1059–1069, 2016.
- GRABHERR, M. G. et al. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. **Nature Biotechnology**, v. 29, n. 7, p. 644–652, 2013.
- GUPTA, R.; BRUNAK, S. Prediction of glycosylation across the human proteome and the correlation to protein function. **Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing**, v. 322, p. 310–322, 2002.
- GURUMALLESH, P. et al. A systematic reconsideration on proteases. **International Journal of Biological Macromolecules**, v. 128, p. 254–267, 2019.
- INVITROGEN, L. T. User manual - EasySelect™ *Pichia* Expression Kit. **Cat. no. K1740-01**, n. 25, p. 95, 2010.
- LINDER, S.; SCHLIWA, M.; KUBE-GRANDERATH, E. Direct PCR screening of *Pichia pastoris* clones. **BioTechniques**, v. 20, n. 6, p. 980–982, 1996.
- MEYERS, S. P.; AHEARN, D. G. Extracellular Proteolysis by *Candida lipolytica*. **Mycologia**, v. 69, n. 3, p. 646, maio 1977.
- MYCOBANK DATABASE. Disponível em: <http://www.mycobank.org/name/Rhizomucor%20miehei> Acesso em 01 de abril de 2018.
- MILLER, G. L. Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. **Analytical Chemistry**, 1959.
- MIMA, J. et al. Structure of the carboxypeptidase Y inhibitor IC in complex with the cognate proteinase reveals a novel mode of the proteinase-protein inhibitor interaction. **Journal of Molecular Biology**, v.

346, n. 5, p. 1323–1334, 2005.

NIELSEN, H. et al. A Brief History of Protein Sorting Prediction. **Protein Journal**, v. 38, n. 3, p. 200–216, 2019.

NING, Y.-J. et al. An extracellular yellow laccase with potent dye decolorizing ability from the fungus *Leucoagaricus naucinus* LAC-04. **International Journal of Biological Macromolecules**, v. 93, p. 837–842, dez. 2016.

PEREIRA, L. A. et al. Proteomic identification, nucleotide sequence, heterologous expression and immunological reactivity of the triosephosphate isomerase of *Paracoccidioides brasiliensis*. **Microbes and Infection**, v. 6, n. 10, p. 892–900, ago. 2004.

PEREIRA, M. G. et al. Screening of filamentous fungi for lipase production: *Hypocrea pseudokoningii* a new producer with a high biotechnological potential. **Biocatalysis and Biotransformation**, v. 32, n. 1, p. 74–83, 7 jan. 2014.

PLANAS-IGLESIAS, J. et al. Computational design of enzymes for biotechnological applications. **Biotechnology Advances**, v. 47, n. January, 2021.

PRASAD, B. V. L. S.; SUGUNA, K. Effect of pH on the structure of rhizopuspepsin. **Acta Crystallographica - Section D Biological Crystallography**, v. 59, n. 10, p. 1755–1761, 2003.

PROMEGA. Wizard® Plus SV Minipreps DNA Purification System Wizard® Plus SV Minipreps DNA Purification System. n. 225, p. 1–12, 2010.

RAO, M. B. et al. Molecular and biotechnological aspects of microbial proteases. **Microbiology and molecular biology reviews : MMBR**, v. 62, n. 3, p. 597–635, set. 1998.

RAWLINGS, N. D. et al. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. **Nucleic Acids Research**, v. 46, n. D1, p. D624–D632, 2018.

SANCHEZ, S.; DEMAIN, A. L. Enzymes and bioconversions of industrial, pharmaceutical, and biotechnological significance. **Organic Process Research and Development**, v. 15, n. 1, p. 224–230, 2011.

SARATH, G.; DE LA MOTTE, R. S.; WAGNER, F. W. Peptidase assay methods. Proteolytic enzymes a practical approach. New York: Oxford University, p. 25-55, 1996.

SCHECHTER, I.; BERGER, A. On the size of the active site in proteases. I. Papain. **Biochemical and Biophysical Research Communications**, v. 27, n. 2, p. 157–162, abr. 1967.

SHU, M. et al. High-level expression and characterization of a novel serine protease in *Pichia pastoris* by multi-copy integration. **Enzyme and Microbial Technology**, v. 92, p. 56–66, 2016.

SILVA, R. DA et al. Production of xylanase and CMCase on solid state fermentation in different residues by *Thermoascus aurantiacus* miehe. **Brazilian Journal of Microbiology**, v. 36, n. 3, p. 235–241, set. 2005.

STEENTOFT, C. et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. **EMBO Journal**, v. 32, n. 10, p. 1478–1488, 2013.

SUN, Q. et al. A novel aspartic protease from *Rhizomucor miehei* expressed in *Pichia pastoris* and its application on meat tenderization and preparation of turtle peptides. **Food Chemistry**, v. 245, n. October 2017, p. 570–577, 2018.

TRAN, L. H.; NAGANO, H. Isolation and characteristics of *Bacillus subtilis* CN2 and its collagenase production. **Journal of Food Science**, v. 67, n. 3, p. 1184–1187, 2002.

WANG, Y. CHUAN et al. A novel high maltose-forming α -amylase from *Rhizomucor miehei* and its application in the food industry. **Food Chemistry**, v. 305, n. May 2019, p. 125447, 2020.

WATERHOUSE, A. et al. SWISS-MODEL: Homology modelling of protein structures and complexes. **Nucleic Acids Research**, v. 46, n. W1, p. W296–W303, 2018.

WOLFENDEN, R.; SNIDER, M. J. The depth of chemical time and the power of enzymes as catalysts. **Accounts of Chemical Research**, v. 34, n. 12, p. 938–945, 2001.

WU, T. Y. et al. Investigations on protease production by a wild-type *Aspergillus terreus* strain using diluted retentate of pre-filtered palm oil mill effluent (POME) as substrate. **Enzyme and Microbial Technology**, v. 39, n. 6, p. 1223–1229, out. 2006.

YANG, J. et al. Controlling AOX1 promoter strength in *Pichia pastoris* by manipulating poly (dA:dT) tracts. **Scientific Reports**, v. 8, n. 1, p. 1–11, 2018.

ZHOU, P. et al. Genome sequence and transcriptome analyses of the thermophilic zygomycete fungus *Rhizomucor miehei*. **BMC Genomics**, v. 15, n. 1, p. 1–13, 2014.

Anexos

Anexo 1. Sequência FASTA dos vetores pPICZαA/rRmAspPep e pPICZαA/rRmCarPep.

>pPICZαA/ rRmAspPep

```

AGATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTTGCCATCCGACATCCACAGGTCATTCTCACACATAAGT
GCCAAACGCAACAGGAGGGGATACACTAGCAGCAGACCGTTGCAAACGCAGGACCTCCACTCCTTCTCCTCAACACCC
ACTTTTTGCCATCGAAAAACCAGCCAGTTATTGGGCTTGATTGGAGCTCGTCAATCCAATTCTTCTATTAGGCTACTA
ACACCATGACTTTATTAGCCTGTCTATCCTGGCCCCCTGGCGAGGTTTCATGTTTGTATTATTCGAATGCAACAAGCTC
CGCATTACACCCGAACATCACTCCAGATGAGGGCTTTCTGAGTGTGGGGTCAAATAGTTTTCATGTTCCCAAAATGGCCCA
AAACTGACAGTTTAAACGCTGTCTTGGAACTAATATGACAAAAGCGTGATCTCATCCAAGATGAACTAAGTTTGGTTTCG
TTGAAATGCTAACGGCCAGTTGGTCAAAAAGAACTTCCAAAAGTGGGCATACCGTTTGTCTTGGTTTGGTATTGATTGAC
GAAATGGGCAAAATAATCTCATTAAATGCTTAGCGCAGTCTCTATCGCTTCTGAACCCCGTGCACCTGTGCCGAAACG
CAAATGGGCAAAACACCCGCTTTTTGGATGATTATGTCATTGTCTCCACATTGTATGCTTCCAAGATTCTGGTGGGAATACT
GCTGATAGCCTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGACAGCAATATATAAACAGAAGGAAGCTG
CCCTGTCTTAAACCTTTTTTTTTATCATCATTATTAGCTTACTTTTCATAATTGCGACTGGTTCCAATTGACAAGCTTTTG
ATTTTAAAGACTTTTAAACGCAACTTGAGAAGATCAAAAAACAATAATTATTCGAAACGATGAGATTTCTTCAATTTT
TACTGCTGTTTTATTTCGAGCATCCTCCGCATTAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTC
CGGCTGAAGCTGTCTACGGTTACTCAGATTTAGAGGGGATTTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAAT
AACGGTTATTGTTTTATAAATACTACTATTGCCAGCATTGTCTAAAGAAAGAAGGGGTATCTCTCGAGAAAAAGA GCTCC
ATCTGGTAAAGAAGTTGTCCATTCCATTGGCCAAGAACGCAACTTCAAGCCAAACGCTAGAAGGGCTTTGAGAAAGGCTT
TGGCCAAGTACAACAAGGCCCAAAAGGTGGTGAGATTTCCCAACAAGGTTACTGGTCTGTTCAGTTGTTGACGTTGGT
TCCGACTTGGAGTACTACGGTATCGTTACTGTTGGTACTCCAGGTGTCAAGCTGAAGTTGGACTTTGACACTGGTTCCCTC
CGATTTGTGGTTTCGCTCTACTTTGTGTTCCCTCCTGTTCTAAGTCAACCTCCACACTAAGTACAACCCAAACAAGTCCCTCCACCT
ACAAGCCAGATGGTAGACCTTGGTCTATTTCCCTACGGTGACGGTTCTTCTGCTTCCGGTATTTTGGCTAAGGACACCGTT
AACCTTGGTGGTTTGGTTATCAAGGACCAGACTATCGGTCTGGCTAAGAGAGAATCTTCCGCTTTCGCTGAGGACGTTAT
CGATGGTTGTTGGGTCCTTGCCTTCAACTCCATCACTACCGTCAGAGGTGTTAAGACCCCAAGTGCAGACAACCTGATTTCC
AGGGTTTGATCTCCAGACCAATCTACGGTGTCTACCTTGGTAAGGCTTCTAACGGTGGTGGTGGCGAGTACATTTTTGGT
GGTTACGACTCCACCAAGTTCAAGGGTCCATTGACCCTATTTCCAATCGACAACCTCCGAAGGTTTCTGGGGTGTACCGT
TAAGGGTACTAAGATCGGTGGTAAGTCCATCGGTGGTTCCCTCTCTGCTATTTTGGACACCGGTACTACCTTGTGTTGC
TGCCAGATAACGTTGCTGCTTCCGTTGCTAGAACTTACGGTGCTACTGATAACGGCGACGGTACTTACACTATCACTTGT
GACGCTTCCAAGTTGCAGCCATTGGTTTTCACTTTGGGTTCTTCCACCTTCCAGGTTCCAGCTGACTCTTTGATCTACGA
GAAGTCTGGTTCACAGTATCGCTGGTTTTGTGCTGGTGGTGAATGCTTGGGTGATGTTTTCCCTGAAGAACA
ACTACGTTGTTTGAACCCAGGTTCCCTGAGGTTCAAATTTGCTCCAGTTGCTCATCATCACCACTAAATTTAGA
ACAAAACCTCATCTCAGAAGAGGATCTGAATAGCGCCGTCGACCATCATCATCATCATCATTGAGTTTGTAGCCTTAGAC
ATGACTGTTCTCAGTTCAGTTGGGCCTTACGAGAAGACCGGTCTTGTAGATTCTAATCAAGAGGATGTCAGAATGC
CATTTGCCCTGAGAGATGCAGGCTTCATTTTTGATACTTTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTCCATT
TTGTTTCTTCTCGTACGAGCTTGTCTCCTGATCAGCCTATCTCGCAGCTGATGAATATCTTGTGGTAGGGGTTTGGGAAAA
TCATTCAGGTTTGTATGTTTTTCTTGGTATTTCCCACTCCTCTTCCAGGTACAGAAGATTAAGTGAAGCTTCCGTTTGTGC
GGATCCCCACACACCATAGCTTCAAATGTTTCTACTCCTTTTTTACTCTTCCAGATTTTCTCGGACTCCGGCATCGC
CGTACCCTTCAAACACCCAAGCACAGCATACTAAATTTTCCCTCTTTCTTCCCTTAGGGTGTGCTTAATTACCCGTAC
TAAAGGTTTGGAAAAGAAAAAGAGACCGCCTCGTTTCTTTTCTTCTCGTGAAGGCAATAAAAAATTTTTATCACGTT
TCTTTTTCTGAAATTTTTTTTTTTAGTTTTTTTCTCTTTTCAAGTACCTCATTGATATTTAAGTTAATAAACGGTCTTC
AATTTCTCAAGTTTCAAGTTTCAATTTTTCTTGTCTATTACAACCTTTTTTTACTTCTTGTTCATTAGAAAAGAAAGCATAGC
AATCTAATCTAAGGGGCGGTGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTATAATACGACAAGGTGAGGAAC
TAAACCATGGCCAAGTTGACCAGTGGCCTTCCGGTGTCTACCAGCGCGCAGCTCGCCGGAGCGGTGAGTTCTGGACCGA
CCGGCTCGGGTCTCCCGGACTTCTGTTGAGGACGACTTCCCGGTGTGGTCCGGGACGACGTTACCTGTTTCATCAGCG
CGGTCCAGGACAGGTGGTGGCGGACAACACCTGGCCTGGGTGTGGGTGCGCGGCTGGACGAGCTGTACGCCGAGTGG
TCGGAGGTGCTGTCCACGAACCTCCGGGACGCTCCGGGCGGCCATGACCAGATCGGCGAGCAGCCGTGGGGGCGGGA
GTTCCGCCCTGCGCGACCCGCGGCAACTGCGTGCCTTCCGTGGCCGAGGAGCAGGACTGACACGTTCCGACGCGGCCCA
CGGGTCCCAGGCTCGGAGATCCGTCCCCCTTTTCTTGTGATATCATGTAATTAGTTATGTACGCTTACATTCACG
CCCTCCCCCACATCCGCTCAACCGAAAAGGAAGGAGTTAGACAACCTGAAGTCTAGGTTCCCTATTTATTTTTTTATAG
TTATGTTAGTATTAAGAACGTTATTTATATTTCAAATTTTTCTTTTTTTTTTGTACAGACGCGGTGACGCATGTAACATT
ATACTGAAAACCTTGCTTGAAGAAGTTTTTGGGACGCTCGAAGGCTTTAATTTGCAAGCTGGAGACCAACATGTGAGCAAA
AGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCGCGTGTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATC
ACAAAATCGACGCTCAAGTCAAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCTTTCCCCCTGGAAGCTCC
CTCGTGCCTCTCCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCCGGAAGCGTGGCGCTTTC
TCAATGCTCAGCTGTAGGTATCTCAGTTCCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTT
AGCCCCAGCGCTGCGCTTATCCGGTAACTATCGTCTTGAAGTCAACCCGGTAAAGACACGACTTATCGCCACTGGCAGCA
GCCACTGTTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTA
CACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCCGAAAAAGAGTTGGTAGCTCTTGATCCG
GCAAACAACACCGCTGGTAGCGGTGGTTTTTTTTGTTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAA
GATCCTTTGATCTTTTTTACGGGCTGACGCTCAGTGGAAACGAAACCTCACGTTAAGGGATTTTGGTTCATGAGATC

```

>pPICZαA/rRmCarPep

AGATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTTGCCATCCGACATCCACAGGTCCATTCTCACACATAAGT
 GCCAAACGCAACAGGAGGGGATACACTAGCAGCAGACCCTTGCAAACGCGAGACCTCCACTCCTCTTCTCCTCAACACCC
 ACTTTTGCCATCGAAAAACCAGCCCAGTTATTGGGCTTGATTGGAGCTCGCTCATTCCAATTCTCTATTAGGCTACTA
 ACACCATGACTTTATTAGCCTGTCTATCCTGGCCCCCTGGCGAGGTTTCATGTTTGTATTATTCGAAATGCAACAAGCTC
 CGCATTACACCCGAACATCACTCCAGATGAGGGCTTTCTGAGTGTGGGGTCAAATAGTTTCATGTTCCCAAATGGCCCA
 AAAGTACAGTTTAAACGCTGTCTTGGAACTAATATGACAAAAGCGTGATCTCATCCAAGATGAACTAAGTTTGGTTTCG
 TTGAAATGCTAACGGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCGGCATAACCGTTTGTCTGTTTGGTATTGATTGAC
 GAATGCTCAAAAATAATCTCATTAATGCTTAGCGCAGTCTCTCTATCGCTTCTGAACCCCGGTGCACCTGTGCCGAAACG
 CAAATGGGGAAACACCCGCTTTTTGGATGATTATGCATTGTCTCCACATTGTATGCTTCCAAGATTCTGGTGGGAATACT
 GCTGATAGCTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGACAGCAATATATAAACAGAAGGAAGCTG
 CCTGTCTTAAACCTTTTTTTTTATCATCATTTATTAGCTTACTTTTATAAATGCGACTGGTTCCAATTGACAAGCTTTTG
 ATTTTAAACGACTTTTAAACGACAACCTTGAGAAGATCAAAAAACAATAATTATTCGAAACGATGAGATTTCTTCAATTTT
 TACTGCTGTTTTATTTCGCAGCATCCTCCGCATTAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAAAATTC
 CGGCTGAAGCTGTCTATCGTTTACTCAGATTTAGAAGGGGATTTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAAT
 AACGGGTTATTGTTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAAAGAAGGAT
 CAAAGAGCCAAAGCTTGGCAGCAGGGTGTCAATACTCCGGTTACATTGACAACATCGAGACTGACGACCACTTCTTCT
 TCTACTTTTTTCGAATCCAGAAGGGACCCAGCTAACGACCCATTGGTTTTGTGGTTGAATGGTCCAGTTGCTCCTCT
 ATGATGGGTTTTGTGGATGGAATGGGTCCATGCTTGGTTTTCTGACGGTGGTAACCTCCACTGTTAGAAAACCCATACTCCTG
 GAACGAGGTTGCCAACGTTGTTTTCTTGGACCAGCCAGTTAACGTCGGTTACTCTTACGGCAAGTCCAAGATCAGAACCT
 CTAAGGACTCCGCCAGAGATGTTTACGCTTCTTGCAGTTGTTCTGGACGAGTTCAAGCAGTACGCTGATAACCCATTTC
 CACATTTCCGGTGAATCTTACGGTGGTCACTACTTGCAGCTTTGGCTGCTGAAATCTTGCACCAGAACAAGTTGGTTGA
 GAAGTCCGGTCACTTGCCATTGAAGTTCGAGTCCATGTTGATCGGTAACGGTTGGACTGAACCTAGAACTCAGTACAAGG
 GTTACGAAGAGTACTCCTGTGCTAACGACTCTGCCTACAAGCCTTTGTTTCGACGAGCAAGAATGTGCCACATGAGAGAA
 ACTTACCCAAAGATGTAAGGCTTGGCCACCTGTGTGTACAACTACCATCTGCCGTTACTTGTGTTCCAGCTCCTTGTGA
 TTGTTCAGAAAACCCAGCAAGAGAACTTCCGGTAAGACCGGTTTGAACCCATACGACATCAGACGTCATGTGCTGGTGATT
 CCGGTTTTGTGTTACGACGAGATTGACTGGATCGAGACTTACGCTAACAGCCAGAAGTTAGAGCTGCTTTGGGTGTTGAT
 TACGAGGCTGGTAACTACACCGGTTGTGATGACTCTGTGCGTTACAGATTCGCTTTGACTGGTGACGGTAAACAGGACCA
 CTCTCCATTTGTTGCTCAGGCTTTGGAAGAGGGTGTGAGAGTTTTGTTGTACGTCGGTGATATGGACTGGATCTGTAAC
 GGGTGGTAACTTGGCTTGGTCTTGGAAATGAAGTGGTCCGGTCAAGAAGAGTACAGAGCCTCTAAAGAAGAGGAATGG
 CGTAAACAAGGCTACTGGTGTGCTGCTGGTTAGTTAGACGCTGCTGGAACTTGACCTTCGCCAGAGTTTTTGTATGCCGG
 TCACATGGTTCCATACGCAACCAACGAAACCGCTTGGACATGTTTCGCTAGATGGTTGGAGAACGAGCTGTTCCGTTTCTA
 ACCATCACACCACCATCATTAATCTAGAACAACAACTCATCTCAGAAGAGGATCTGAATAGCGCCGTCGACCATCATCA
 TCATCATCATTTGAGTTTTGTAGCCTTAGACATGACTGTTCCCTCAGTTCAAGTTGGGCCTTACGAGAAGACCGGCTTTGCT
 AGATTTCAATCAAGAGGATGTCAGAATGCCATTTGCCTGAGAGATGCAGGCTTCATTTTTTGATACTTTTTTTATTTGTAAC
 CTATATAGTATAGGATTTTTTTTTGTGCTTTTTGTTTCTTCTCGTACGAGCTTGTCTCTGATCAGCCTATCTCGCAGCTGAT
 GAATATCTTGTGGTAGGGTTTTGGGAAAATCATTTCAGTTTTGATGTTTTTCTTGGTATTTCCCACTCCTCTTCAGACTAC
 AGAATAAGTTAGTGAGACTTCGTTTTGTGCGGATCCCCACACACCATAGCTTCAAATGTTTTCTTACTCCTTTTTTTACTCT
 TCCAGATTTTCTCGGACTCCGCGCATCGCCGTACCCTTCAAAAACCCAAAGCAGCATACTAAATTTTTCCCTCTTTCT
 TCCTCTAGGGTGTGTTAATTAACCGTACTAAAGGTTTTGGAAAAGAAAAAAGAGACCGCCTCGTTTTCTTTTTCTTCGTCG
 AAAAAGGCAATAAAAATTTTTATCACGTTTTCTTTTTCTTGAATTTTTTTTTTTAGTTTTTTCTCTTTTTCAGTGACCTCC
 ATTTGATATTTAAGTTAATAAACGGTCTTCAATTTCTCAAGTTTCAGTTTCAATTTTTCTTGTCTATTACAACTTTTTTTA
 CTTCTTGTTCATTAGAAAAGAAAGCATAGCAATCTAATCTAAGGGGCGGTGTTGACAATTAATCATCGGCATAGTATATCG
 GCATAGTATAATACGACAAGGTGAGGAACATAACCATGGCCAAGTTGACCAGTGCCTTCCGGTGTCCACCGCGCGGAC
 GTCGCCCGAGCGGTGAGTTCTGGACCGACCGGCTCGGTTCTCCCGGACTTCGTGGAGGACGACTTCGCCCCTGTGGT
 CCGGGACGACGTGACCCGTTCATCAGCGCGGTCCAGGACCAGGTGGTGGCCGACAACACCCCTGGCCTGGGTGTGGGTGC
 GCGGCCCTGGACGAGCTGTACGCCGAGTGGTCCGAGGTCGTGTCCACGAACCTCCGGGACGCCCTCCGGGCCGGCCATGACC
 GAGATCGGGGAGCAGCCGTGGGGGCGGGAGTTCCGCCCTGCGCGACCCGGCCGCAACTGCGTGCCTTCTGTTGCCGAGGA
 GCAGGACTGACACGTCCGACGGCGGCCACGGGTCCAGGCTCCGAGATCCGTCCTCCCTTTTCCCTTTGTGATATCATG
 TAATTAGTTATGTCACGCTTACATTCACGCCCTCCCCCACAATCCGCTCTAACCGAAAAGGAAGGAGTTAGACAACCTGA
 AGTCTAGGTTCCCTATTTATTTTTTATAGTTATGTTAGTATTAAGAACGTTATTTATATTTCAAATTTTTCTTTTTTTTCT
 TGTACAGACCGGTGTACGCATGTAACATTAATACTGAAAACCTTGCTTGGAGAAGTTTTGGGACGCTCGAAGGCTTTAATT
 TGCAAGCTGGAGACCAACATGTGAGCAAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAAAGGCCGCGTGTGCTGGCGTTTTT
 CCATAGGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAACCCGACAGGACTATAAA
 GATACCAGGCTTTCCCTTGGAAAGCTCCCTCGTGCCTCTCCTGTTCCGACCTGCCCCTTACCAGGATACCTGTCCGCC
 TTTCTCCCTTCGGGAAGCGTGGCGTTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCCGGTGTAGGTCGTTCCGCTCAA
 GCTGGGCTGTGTGCACGAACCCCGCTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGTGAGTCAACCCGG
 TAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAG
 TTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT
 CGGAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAACCCAGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGAG
 TTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACCTCA
 CGTTAAGGGATTTTTGGTTCATGAGATC

Anexo 3. Sequências FASTA de RmAspPep e RmCarPep.

>RmAspPep/TRINITY_DN16840_c0_g1_i3

MKLTLISSCIAVAAMALAVEAAPS GKKLSIPLAKNDNFKNARRALRKALAKYNKAPKGGEISFPQGTGSVPVVDVGSdle
YYGIVTVGTPGVKlKlDFDTGSSDLWFASTLCSSCSNSHTKYNPNKSSTYKPDGRPWSISYGDGSSASGILAKDTVNLGG
LVIKDQTIgLaKRESSAFAEDVIDGllGLAFNSITTVRGVKTpVDNLISQGLISRPIYGVYLGKASNGGGGEYIFGGYDS
TKFKGPLTTIPIDNSEGFwGVTVKGTkIGGKSIgGSFSAILDTGTTllllLPDNVAASVARTYgATDNGDGTyTITCDASK
LQPLVFTLGSSTFQVPADSLIYEKSGSQCIAGFAAGGDLAILGDVFLKNNYVVFNQVPEVQIAPVAH

>RmCarPep/TRINITY_DN5894_c0_g1_i1

MRGLSAVLSLLIAASSSFALDFHFVDQSNPTPAsERAVPDVQDQGYRVLANERVKGYSVRIKEPKSCEQGVQYSGYIDNI
ETDDHFFfyFFESRRDPANDPLVLWLNggPGCSMMGLWMElGPCLVSDGGNSTVRNPYswNEVANVVFLDQpVNVGYSY
GKSKIRTSKDSARDVYAFQLFLDEFKQYADNPFHISGESYGGHYLPALAAEILHQNKlVEKSGHLPLKFESMLIGNGT
EPRTQYKGYEEYSCANDSAYKPLFDEQECAHMRETYPRCKALATLCYNYPSAVTCVPASLYCQKTQQENFGKTGLNPYDI
RRPCAGDSGLCYDEIDWIETyanQPEVRAALGVdYEAGNYTGCDdSVGYRFALTGDGNKDHSPFVAQALEEGVRVLLYVG
DMDWICNWVGNLAWsLEMkwSGQEEYRASKEEWRNKATGAAAGLVRRAGNLTfARVFDAGHMVPYDQPenALDMFARWL
ENELFVSN

