

**UNIVERSIDADE DE SÃO PAULO**  
**INSTITUTO DE QUÍMICA**  
Programa de Pós-Graduação em Ciências Biológicas (Bioquímica)

**ANDRÉ COPPE PIMENTEL**

**Estudos Funcionais e Moleculares Relativos às**  
**Membranas Apicais Intestinais de *Dysdercus***  
***peruvianus***

***Prof. Dr. Walter Ribeiro Terra***

**Orientador**

**Versão original da Tese defendida**

São Paulo

Data do Depósito na SPG:

**02/06/2016**

ANDRÉ COPPE PIMENTEL

**Estudos Funcionais e Moleculares Relativos às  
Membranas Apicais Intestinais de *Dysdercus  
peruvianus***

*Tese apresentada ao Instituto de  
Química da Universidade de São Paulo para  
obtenção do Título de Doutor em  
Ciências (Bioquímica)*

*Orientador: Prof. Dr. Walter Ribeiro Terra*

São Paulo

2016



***Aos meus pais, irmão e família.***

## AGRADECIMENTOS

Ao Dr. Walter Ribeiro Terra pela orientação, pelo constante suporte e pela oportunidade de realizar esta pesquisa. Sou grato a ele e a Dra. Clélia Ferreira por me aceitarem no Laboratório de Bioquímica de Insetos no qual realizei grande parte da minha formação acadêmica.

A Dra. Adriana Rios Lopes e Dr. Alberto de Freitas Ribeiro pelas discussões e colaborações. Ao Dr. Fernando Genta pelo apoio em terras cariocas. Ao Dr. Ednildo Machado que tornou nosso trabalho mais alegre e unido.

Aos técnicos que apoiaram este trabalho e sem os quais ele não poderia ser realizado Maria Ivanilde Marcelino, Christiane Cardoso, Gilliard de Farias, Waldir Caldeira e Luci Deise Navarro. Aos que me ensinaram o árduo trabalho de bancada ainda na iniciação científica Dra. Thaís Bifano, Dr. Fabio Kendi Tamaki e o amigo Dr. Marcelo Henrique Padilha.

Aos amigos de laboratório Dra. Daniela Beton, Dra. Nathalia Ramalho, Dr. Felipe Jun Fuzita, Dra. Renata Dias, Alexandre Utiyama, Camila Leal, Daniela da Cunha Bataglioli, Ignacio, Felipe Pacheco de Castro e Jessica Maria. Em especial Dra. Walciane da Silva, Ticiane Damasceno e Gabriel Braga Oliveira por dividirem as alegrias e angústias da vida acadêmica.

A Fapesp pela concessão da bolsa de doutorado. O projeto teve o apoio da Fapesp, CAPES, CNPq e INCT-EM.

Agradeço aos meus pais Carmem e Geraldo, que sempre estiveram do meu lado, apoiando nos momentos mais difíceis e ao meu irmão Renê, pela inspiração. Essas três pessoas formam a base de quem sou e por isso serei eternamente grato a elas. Agradeço aos meus avós, minha cunhada, meus tios, primos e familiares que sempre acompanharam minha carreira com palavras de incentivo e carinho.

A querida Sheina Koffler pelas constantes revisões e sugestões que me fizeram um melhor cientista. Obrigado pelo carinho e imensa paciência. Independente de situação, pude sempre contar com seu apoio.

Aos amigos de sempre Ana Carolina, Juliana Vendrami, Diego Vasques, Ivan Dias, Adriana Sandre, Leticia Yoshida e tantos outros pela grande amizade.

***“Sofreu durante muito tempo. Não sabia como abraçar a  
simplicidade do não saber.  
Aos poucos, as coisas começaram a mudar.  
A direção que antes levava a nada virou uma nova estrada.  
Era bom não ser parte de um plano para o mundo. Pela  
primeira vez, o homem velho sentiu-se livre  
Podia sempre continuar a perguntar e a aprender sobre o  
mundo e sobre si próprio.  
Podia sempre amar e esperar que fosse amado.”  
Marcelo Gleiser***

## RESUMO

Pimentel, A.C. **Estudos Funcionais e Moleculares Relativos às Membranas Apicais Intestinais de *Dysdercus peruvianus***. 2016 (137p). Tese de Doutorado - Programa de Pós-Graduação em Ciências Biológicas (Bioquímica). Instituto de Química, Universidade de São Paulo, São Paulo.

Os insetos da ordem Hemiptera apresentam membranas lipoproteicas que revestem as microvilosidades das células intestinais, como se fossem dedos de luva, e formam expansões para o lúmen do intestino que parecem terminar em fundo cego. A presença das duas membranas no ápice dos enterócitos gera questões intrigantes de como se dá a formação da membrana perimicrovilar, como ocorre a absorção de nutrientes e como se dá o encaminhamento de enzimas digestivas para o lúmen intestinal. A digestão de proteínas baseada em enzimas originalmente lisossômicas é uma característica marcante nos Hemiptera, em especial os Heteroptera que evolutivamente voltaram a uma alimentação de polímeros. O presente estudo indica que os genes das proteinases tipicamente lisossômicas sofrem uma série de duplicações, havendo a manutenção de um gene para função puramente lisossômica e a divergência funcional dos demais genes para a função de digestão extracelular. O gene que mantém a função lisossômica não é modulado pela alimentação, além de ser expresso nos mais diversos tecidos. Já os genes que se especializaram na digestão extracelular têm a expressão aumentada com a ingestão de alimento, indicando sua função. Parece não haver diferença nas características relacionadas ao encaminhamento celular das proteínas produzidas por esses genes, indicando que o direcionamento para a rota secretória é devido à superexpressão dos genes relacionados à digestão. Enzimas como alfa-glicosidases, alfa-manosidases e aminopeptidases que participam da digestão extracelular seguem a rota secretória que envolve a formação de vesículas de dupla membrana. Foi possível ampliar o modelo de digestão incluindo a participação das catepsinas D, de uma alfa-glicosidase solúvel e a possível participação de uma tiolredutase além de definir o local de atuação das lipases. Temos agora uma visão global da participação das enzimas digestivas que atuam na digestão em *D. peruvianus*.

**Palavras-chave:** Manchador do algodoeiro, Membrana perimicrovilar, proteinase ácida.



## ABSTRACT

Pimentel, A.C. **Functional and Molecular Studies Related to Apical Midgut Membranes of *Dysdercus peruvianus***. 2016 (137p). PhD Thesis - Graduate Program in Biochemistry. Instituto de Química, Universidade de São Paulo, São Paulo

The insects of the order Hemiptera have lipoprotein membranes lining the microvilli of midgut cells, like glove fingers, and form expansions into the lumen of the intestine. The presence of two membranes on the apex of enterocytes thus generates intriguing questions about the formation of perimicrovillar membrane, the absorption of nutrients, and the targeting of digestive enzymes into the intestinal lumen. The digestion of proteins based on originally lysosomal enzymes is an important feature in Hemiptera, especially in Heteroptera that evolutionary returned to feed on polymers. The genes of typical lysosomal proteinases undergo a series of duplications followed by the maintenance of a gene for purely lysosomal function and functional divergence of other genes for extracellular digestion function. The gene that maintains the lysosomal function is not modulated by feeding, in addition to being expressed in diverse tissues. By the other hand, genes specialized in extracellular digestion are up regulated by food intake, indicating its function. No difference was found in the targeting in the proteins produced by these genes, which indicates that targeting to the secretory route is due to overexpression of digestion-related genes. Enzymes involved in extracellular digestion as alpha-glucosidase, alpha-mannosidase and aminopeptidase follow the secretory route that includes the formation of double membrane vesicles. In this study we increased the digestion model adding the participation of cathepsin D, a soluble alpha-glucosidase, and the possible participation of a tiolredutase, and also to defining the place of lipases operation. We now have a global view of the participation of digestive enzymes involved in digestion *D. peruvianus*.

**Keywords:** Acid protease, Cotton stainer, Perimicrovillar membrane.

## LISTA DE FIGURAS E TABELAS

<b>Figura 1.</b> Modelo de formação de membrana perimicrovilar .....	17
<b>Figura 2.</b> Esquema do intestino médio de <i>Dysdercus peruvianus</i> .....	18
<b>Figura 3.</b> Fluxograma da preparação das frações subcelulares .....	38
<b>Figura 4.</b> Análise dos genes diferencialmente expressos entre as porções V1 e V2 .....	48
<b>Figura 5.</b> Alinhamento das sequencias de DPCATD .....	51
<b>Figura 6.</b> Comparação da posição da alça de prolina .....	52
<b>Figura 7.</b> Cladograma de sequencias de catepsina D .....	53
<b>Figura 8.</b> Expressão dos genes de catepsina D em diferentes tecidos .....	54
<b>Figura 9.</b> Alinhamento das sequencias de DPCAL .....	56
<b>Figura 10.</b> Expressão do RNAm que codificam as diferentes catepsinas L .....	58
<b>Figura 11.</b> Cladograma de sequencias de catepsina L .....	60
<b>Figura 12.</b> Modelo computacional por homologia da DPCAL1 .....	61
<b>Figura 13.</b> Expressão relativa das catepsinas L de <i>Dysdercus peruvianus</i> por qPCR .....	62
<b>Figura 14.</b> Eletroforeses dos produtos das reações de amplificação (DPCAL1) .....	63
<b>Figura 15.</b> Eletroforese das proteínas expressas em <i>E. coli</i> (DPCAL1 recombinante).....	64
<b>Figura 16.</b> Eletroforese de proteínas do processo de purificação ( DPCAL1 recombinante) .	65
<b>Figura 17.</b> Eletroforese e western blot da DPCAL1 recombinante .....	66
<b>Figura 18.</b> Eletroforese e western blot da DPCAL1 recombinante (teste de diluição) .....	66
<b>Figura 19.</b> Reconhecimento da DPCAL1 de <i>Dysdercus peruvianus</i> por <i>western blotting</i> .....	67
<b>Figura 20.</b> Eletroforese das proteínas expressas em <i>E. coli</i> (DPCAL1 recombinante) .....	68
<b>Figura 21.</b> Atividade sobre Z-FR-MCA do extrato da <i>E. coli</i> (DPCAL1 recombinante) .....	69
<b>Figura 22.</b> Cladograma de diferentes catepsinas L e 26/29-kDa proteinases .....	71
<b>Figura 23.</b> Eletroforese dos produtos de amplificação (clonagem da DPCAL3) .....	72
<b>Figura 24.</b> Eletroforese das proteínas expressas em <i>E. coli</i> (DPCAL3 recombinante) .....	72
<b>Figura 25.</b> Eletroforese do produto de amplificação (clonagem da DPCAL5) .....	73
<b>Figura 26.</b> Expressão do RNAm que codifica DPCAL10 .....	74
<b>Figura 27.</b> Eletroforese dos produtos de amplificação (clonagem da DPCAL10) .....	75
<b>Figura 28.</b> Eletroforeses das proteínas expressas em <i>E. coli</i> (DPCAL10 recombinante) .....	76
<b>Figura 29.</b> Eletroforese (SDS-PAGE) e western blot da DPCAL10 .....	77
<b>Figura 30.</b> Micrografia de enterócito de V1 imunomarcados por anti-DPCAL10 .....	78
<b>Figura 31.</b> Micrografia de enterócito de V1 - organela marcada .....	79
<b>Figura 32.</b> Centrifugação diferencial de homogeneizados V1 em meio isotônico .....	80
<b>Figura 33.</b> Centrifugação diferencial de homogeneizados do tecido de V2 .....	81
<b>Figura 34.</b> Centrifugação diferencial de homogeneizados do tecido de V2 (MTS 1 mM) .....	82
<b>Figura 35.</b> Atividade enzimática de frações subcelulares de V1 .....	83
<b>Figura 36.</b> Alinhamento da porção C-terminal da DPAlphaGlu1 e DPAlphaGlu2 .....	85
<b>Figura 37.</b> Expressão dos RNAm que codificam diferentes alfa-glicosidasas .....	86
<b>Figura 38.</b> Eletroforeses obtidas durante na clonagem e expressão (DPAlpMan) .....	87
<b>Figura 39.</b> Eletroforese das proteínas insolúveis expressas por <i>E. coli</i> (DPAlpMan) .....	88
<b>Figura 40.</b> Eletroforese e western blot da DPAlpMan recombinante .....	88
<b>Figura 41.</b> Micrografia de enterócito de V1 imunomarcado por anti-DPAlpMan .....	90
<b>Figura 42.</b> Eletroforeses obtidas durante na clonagem e expressão (DPAlphaGlu1) .....	92
<b>Figura 43.</b> Micrografia de enterócito de V1 imunomarcado por anti-DPAlphaGlu1 .....	93

<b>Figura 44.</b> Eletroforeses obtidas no processo de clonagem da DPAlphaGlu2 .....	94
<b>Figura 45.</b> Eletroforese das proteínas expressas em <i>E. coli</i> (DPAlphaGlu2 recombinante) ..	95
<b>Figura 46.</b> Eletroforese e western blot da DPAlphaGlu2 recombinante .....	95
<b>Figura 47.</b> Reconhecimento da DPAlphaGlu2 por western blotting .....	96
<b>Figura 48.</b> Micrografias de enterócitos de V1 imunomarcados por anti-DPAlphaGlu2 .....	97
<b>Figura 49.</b> Micrografia de enterócitos de V2 imunomarcados por anti-DPAlphaGlu2 .....	97
<b>Figura 50.</b> Expressão do RNAm que codifica a DPAPN foi avaliada por RT-PCR .....	99
<b>Figura 51.</b> Eletroforeses obtidas no processo de clonagem da DPAPN .....	100
<b>Figura 52.</b> Micrografia de enterócito de V1 imunomarcado por anti-DPAPN .....	101
<b>Figura 53.</b> Proteínas encontradas no lúmen de V1 .....	104
<b>Figura 54.</b> Modelo da distribuição das enzimas digestivas em <i>Dysdercus peruvianus</i> .....	111
<b>Figura 55.</b> Ligação entre a rota endossomo-lisossomo e a rota secretória .....	117
<b>Figura 56.</b> Modelo de secreção proposto .....	122
<b>Figura 57.</b> Eletroforese das proteínas de homogeneizados intestinais de <i>D. peruvianus</i> .....	133
<b>Tabela 1.</b> Compilação das enzimas descritas em <i>Dysdercus peruvianus</i> .....	23
<b>Tabela 2.</b> Contigs de genes que possivelmente codificam transportadores de glicose .....	47
<b>Tabela 3.</b> Expressão diferencial dos genes que codificam transportadores de glicose .....	49
<b>Tabela 4.</b> Expressão dos RNAs que codificam as diferentes catepsinas D .....	50
<b>Tabela 5.</b> Catepsinas D identificadas no lúmen de V1 por técnica proteômica .....	55
<b>Tabela 6.</b> Modulação da expressão de catepsinas D em V1 .....	55
<b>Tabela 7.</b> Predição de características obtidas a partir de modelagem computacional .....	57
<b>Tabela 8.</b> Comparação entre as DPCALs com catepsinas L de inseto ( <i>T. molitor</i> ) .....	59
<b>Tabela 9.</b> Aminopetidases obtidas no pirosequenciamento .....	98
<b>Tabela 10.</b> Resumo da produção e experimentos realizados com diferentes anticorpos .....	102
<b>Tabela suplementar 1.</b> Iniciadores usados .....	134
<b>Tabela suplementar 2.</b> Proteínas solúveis identificadas no conteúdo de V1 .....	136

## LISTA DE ABREVIATURAS E SIGLAS

**APN:** aminopeptidase N

**BSA:** Albumina sérica bovina

**cDNA:** Ácido desoxirribonucléico complementar

**DEPC:** Dietil pirocarbonato

**dNTP:** Desoxirribonucleotíde fosfatado

**DTT:** Ditioneitol

**EDTA:** Ácido etilenodiamino tetra-acético

**GLUT:** Transportador de glicose

**GPI:** Glicosil fosfatidil inositol

**IPTG:** Isopropil beta-D-1-tiogalactosídeo

**LB:** Meio de cultura Luria-Bertani

**mRNA:** RNA mensageiro

**MCA:** 7-amino-4-metilcumarina

**PAGE:** Eletroforese em gel de poliacrilamida

**PCR:** Reação em cadeia da polimerase

**pH:** cologaritmo de concentração de H<sup>+</sup>.

**RPM:** Rotação por minuto

**RT-PCR:** Reação da transcriptase reversa seguida de PCR

**SDS:** Dodecil sulfato de sódio

**SDS-PAGE:** Eletroforese em gel de poliacrilamida na presença de SDS

**SGLT:** Cotransportador sódio/glicose

**TBS:** Salina tamponada por Tris

**TBST:** TBS na presença de Tween

**Tris:** Tris-hidroximetil aminometano

**TAE:** Tris-Acetato-EDTA

## SUMÁRIO

<b>1. INTRODUÇÃO</b> .....	15
1.1. Considerações iniciais .....	15
1.2. Membranas apicais intestinais .....	16
1.3. Função geral e organização da digestão .....	18
1.4. Encaminhamento celular de enzimas .....	19
1.5. Enzimas digestivas em Hemiptera e suas origens.....	20
1.6. Enzimas digestivas de <i>D. peruvianus</i> .....	22
1.7. Catepsina D (aspártico proteinase).....	23
1.8. Catepsina L (cisteíno proteinase).....	24
1.9. Aminopeptidase N de insetos .....	25
1.10. Alfa-glicosidase na digestão em insetos .....	25
<b>2. OBJETIVOS</b> .....	27
<b>3. PROCEDIMENTOS EXPERIMENTAIS</b> .....	28
3.1. Animais .....	28
3.2. Dissecção de animais .....	28
3.3. Extração de RNA .....	29
3.4. Produção de cDNA .....	30
3.5. PCR e RT-PCR Semiquantitativo.....	31
3.6. Pirosequenciamento .....	32
3.7. Análise de sequências .....	32
3.8. Reação de sequenciamento de DNA segundo Sanger.....	33
3.9. Ligação dos fragmentos de cDNA a plasmídeos vetores.....	34
3.10. Transformação de bactérias competentes para expressão .....	34
3.11. Expressão de proteínas recombinantes com vetor pAE .....	35
3.12. Lise das células induzidas .....	35
3.13. Purificação de proteína recombinante.....	35
3.14. Eletroforeses .....	36

3.15. Centrifugação diferencial do homogeneizado .....	37
3.16. Ultracentrifugação em gradiente de sacarose.....	38
3.17. Produção de anticorpo .....	39
3.18. Tratamento do soro.....	39
3.19. Produção do anti-DPAPN .....	40
3.20. <i>Western blotting</i> .....	40
3.21. Ensaio enzimáticos .....	41
3.22. Fixação e imunomarcação para microscopia de transmissão.....	42
3.23. qPCR .....	43
3.24. Expressão diferencial por sequenciamento Illumina .....	43
3.25. Preparação de amostras para proteômica .....	44
<b>4. RESULTADOS.....</b>	<b>46</b>
4.1. Expressão diferencial entre V1 e V2 – e a absorção de glicose .....	46
4.1.1. Transportador de glicose .....	46
4.1.2. Expressão diferencial de genes em V1 e V2 .....	47
4.2. Secreção de proteinases .....	49
4.2.1. Catepsinas D de <i>Dysdercus peruvianus</i> .....	49
4.2.2. Catepsinas L de <i>Dysdercus peruvianus</i> .....	55
4.2.3. Catepsinas L de <i>Dysdercus peruvianus</i> (estudo das selecionadas).....	62
4.2.3.1. DPCAL1 .....	63
4.2.3.2. DPCAL2.....	69
4.2.3.3. DPCAL3.....	71
4.2.3.4. DPCAL5.....	73
4.2.3.5. DPCAL10.....	73
4.3. Produção de membrana perimicrovilar .....	80
4.3.1. Fracionamento celular por centrifugação diferencial.....	80
4.3.2. Fracionamento celular (resolução em gradiente de sacarose) .....	83
4.3.3. Estudo dos marcadores celulares selecionados .....	84
4.3.3.1. Alfa-glicosidases gerais de <i>D. peruvianus</i> .....	84
4.3.3.2. DPAIpMan.....	86
4.3.3.3. DPAlphaGlu1 .....	91

4.3.3.4. DPAlphaGlu2 .....	93
4.3.3.5. DPAPN .....	98
4.3.3.6. Compilação dos anticorpos produzidos .....	102
4.3.4. Proteômica do conteúdo solúvel de V1 .....	102
<b>5. DISCUSSÃO .....</b>	<b>105</b>
5.1. Fisiologia molecular da digestão em <i>Dysdercus peruvianus</i> .....	105
5.2.1. Digestão e absorção no intestino médio .....	105
5.2.2. Compreensão global da digestão .....	110
5.2. Secreção de enzimas digestivas.....	111
5.2.1. Secreção de proteinases e a rota endossomo-lisossomo.....	111
5.2.1.1. Catepsina D.....	112
5.2.1.2. Catepsina L.....	114
5.2.1.3. Secreção de proteinases .....	117
5.2.2. Secreção de enzimas digestivas no sistema de duas membranas.....	118
<b>6. CONCLUSÕES .....</b>	<b>121</b>
<b>7. BIBLIOGRAFIA .....</b>	<b>124</b>
<b>8. APÊNDICES .....</b>	<b>130</b>

## 1. Introdução

### 1.1. Considerações iniciais

A ordem Hemiptera é a mais diversa dentre os insetos não-holometábolos e compreende insetos das subordens Auchenorrhyncha (cigarras e cigarrinhas) e Sternorrhyncha (pulgões e cochonilhas), que se alimentam quase que exclusivamente de seiva vegetal, e da subordem Heteroptera (percevejos), que são predadores ou fitófagos (Grimaldi e Engel, 2005)

A ordem Hemiptera inclui os diversos percevejos que sabidamente causam prejuízos à agricultura, como dos pulgões e cigarrinhas, e são vetores de doenças, como as diferentes espécies de barbeiros. Espécies do gênero *Dysdecus* causam prejuízos à cultura do algodão, pois ao sugarem as maçãs (fruto jovem) e sementes causam queda ou retardamento no desenvolvimento do fruto além de permitir a entrada de patógenos pelo orifício perfurado. Podem também manchar as fibras do algodão em decorrência de seus dejetos (Pereira et al., 2006).

Os insetos da ordem Hemiptera apresentam membranas lipoproteicas que revestem as microvilosidades, como se fossem dedos de luva, e formam expansões para o lúmen do intestino que parecem terminar em fundo cego (membrana perimicrovilar, ver Ferreira et al., 1988).

O significado funcional das membranas perimicrovillares, segundo Terra e Ferreira (1994), seria a absorção de nutrientes essenciais da seiva sugada pelos percevejos ancestrais, que seriam semelhantes aos Auchenorrhyncha atuais. Os percevejos predadores, sugadores de sangue ou sementes (Heteroptera) teriam evoluído dos sugadores de seiva e conservado as membranas perimicrovillares para compartimentar a digestão, como substituto da membrana peritrófica perdida pelos percevejos ancestrais.

A presença das duas membranas no ápice dos enterócitos gera questões intrigantes de como se dá a formação da membrana perimicrovilar, como ocorre a absorção de nutrientes e como se dá o encaminhamento de enzimas digestivas para o lúmen intestinal.



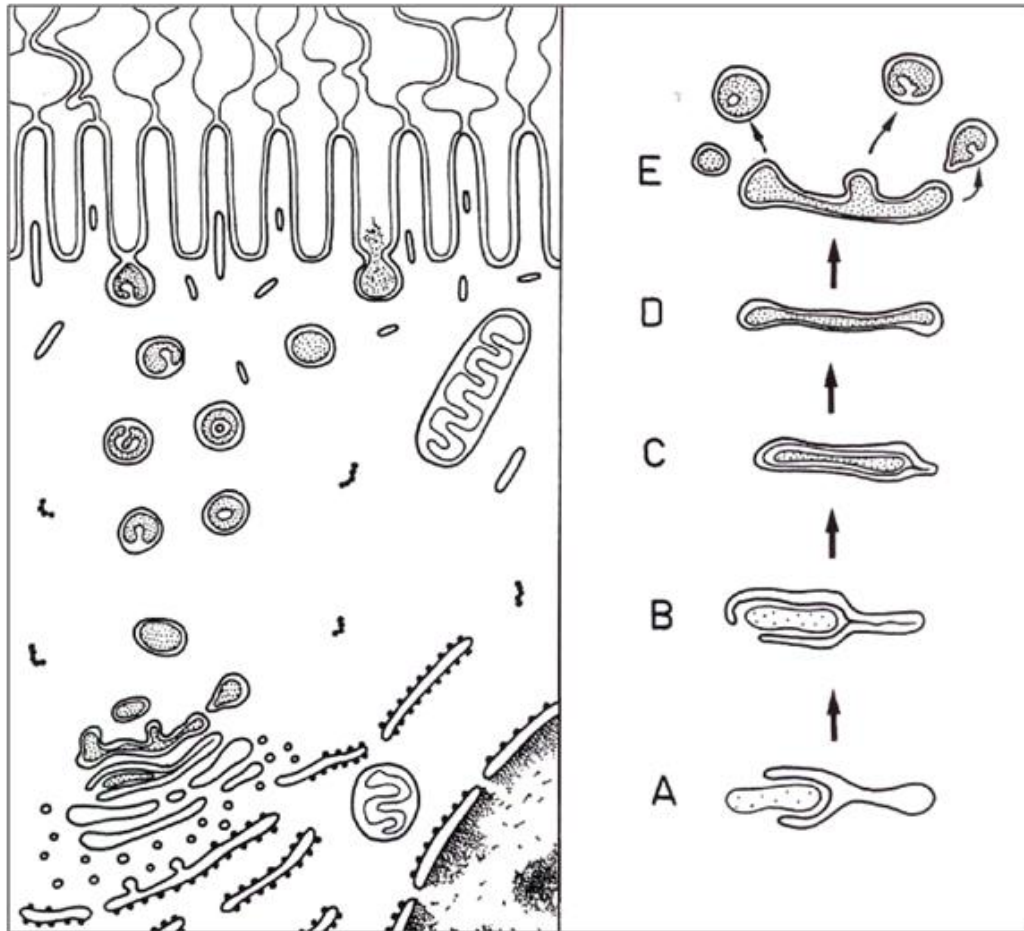
## 1.2. Membranas apicais intestinais

Células intestinais tanto de mamíferos quanto de insetos são descritas como contendo uma membrana plasmática simples com microvilosidades cobertas por glicocálix no polo apical. Há, contudo, a descrição da presença de dupla membrana do ápice de células em insetos sugadores de sangue ou seiva.

Inicialmente descrito nas células intestinais de sugadores de seiva e no sugador de sangue *Triatoma infestans*, a presença da dupla membrana cobrindo as células epiteliais do intestino médio foi mais bem caracterizada no também sugador de sangue *Rhodnius prolixus* (Lane e Harrison, 1979). Foi descrita a presença de dupla membrana no ápice celular de *Oncopeltus fasciatus*. Nele foram descritas vesículas no lado luminal da dupla membrana, formado pela membrana externa (Baerwald e Delcarpio, 1983).

A porção apical das células intestinais contém além da membrana microvilar típica uma segunda membrana, externa, que circunda a microvilosidade e se alonga em direção ao lúmen celular. Essa membrana externa se estende até lúmen intestinal formando um sistema complexo de membranas semelhantes à mielina e é caracterizada pela presença de cristas lineares compostas por partículas (Lane e Harrison, 1979).

Membranas perimicrovilares estão presentes nas diferentes fases da ninfa e nos adultos, em animais alimentados ou não. A membrana perimicrovilar é uma estrutura permanente, mas sua extensão parece depender da disponibilidade de alimento (Silva et al., 1995). É interessante notar que as membranas das microvilosidades e as membranas perimicrovilares apresentam composição de enzimas e densidades características (Ferreira et al., 1988). Como em *Rhodnius prolixus*, as membranas microvilares e perimicrovilares dos enterócitos de *Dysdercus peruvianus* diferem quanto à densidade (Silva et al., 1996).



**Figura 1.** Modelo de formação de membrana perimicrovilar. À esquerda, vesículas de dupla membrana brotando da rede trans-Golgi e transportando proteínas secretórias que são liberadas no lúmen após a fusão da membrana externa da vesícula com a membrana microvilar e da membrana interna com a membrana perimicrovilar. À direita, detalhes da origem das vesículas de dupla membrana a partir de uma cisterna trans do complexo de Golgi (A) que é envolvida por outra cisterna (B) e (C). A membrana mais interna desaparece dando origem a uma cisterna de dupla membrana da qual brotam as vesículas (retirado de Silva et al., 1995).

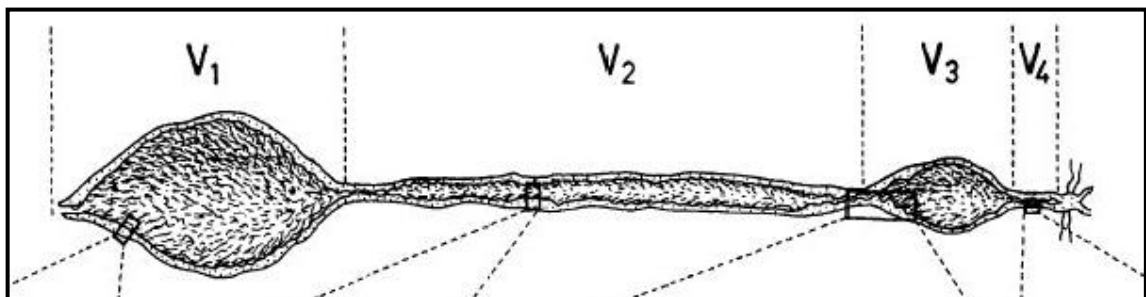
Silva e colaboradores (1995), através de experimentos de imunolocalização de uma alfa-glicosidase (marcador de MPM), indicaram que a MPM se origina da membrana interna de uma vesícula de dupla membrana após dois eventos de fusão de membrana (figura 1, esquerda). As vesículas de dupla membrana se originariam do rearranjo de duas cisternas adjacentes da rede trans-Golgi, seguido da perda da membrana interna na cisterna circundante (conforme figura 1, direita).

Tal modelo, contudo, precisa ser mais bem validado utilizando marcadores celulares dos diferentes compartimentos: conteúdo da vesícula interna, espaço entre vesículas e membrana da vesícula externa. Gerar anticorpos que identifiquem tais marcadores seria uma importante ferramenta para melhor compreensão do processo. As aminopeptidases já se apresentam como boas candidatas para

marcadores do espaço entre a vesícula interna e externa e a catepsina L como marcadora do conteúdo da vesícula interna. Isso porque dados de fracionamento celular apontam para uma localização intermembrana para as aminopeptidases (Ferreira et al., 1988) e as catepsinas são encontradas no lúmen intestinal (Silva e Terra, 1994).

### 1.3. Função geral e organização da digestão

No Hemiptera hematófago *Rhodnius prolixus* enzimas responsáveis pela digestão inicial (hidrólise de polímeros) ocorrem no lúmen do intestinal, enquanto aquelas envolvidas na digestão intermediária e final (hidrólise de oligômeros, trímeros e dímeros) ocorrem ligadas as membranas microvilares e perimicrovilares dos enterócitos (Ferreira et al., 1988). Somado a isso, a presença de estruturas semelhantes às junções de hiato (gap-junctions) na membrana externa (perimicrovilar) e sua aparente ausência na membrana microvilar (Lane e Harrison, 1979) indica diferença funcional entre as membranas.



**Figura 2.** Esquema do intestino médio de *Dysdercus peruvianus*. Retirado de Silva et al., 1995. V: ventrículo.

Considerando agora as porções intestinais, no Hemiptera sugador de semente *Dysdercus peruvianus* a digestão de carboidrato e a absorção de glicose e água se dão principalmente em V1, contrastando com a digestão de proteínas e absorção de aminoácidos que é feita especialmente em V2 e V3 (figura 2) (Silva e Terra, 1994). Baseados em experimentos fisiológicos in vivo, Silva e Terra (1994) demonstraram haver inibição da absorção de leucina em *D. peruvianus* quando adicionado  $\text{Na}^+$  na dieta. Já  $\text{K}^+$  não influenciou a absorção de leucina. Os autores propuseram haver absorção de leucina por cotransporte leucina/ $\text{K}^+$ . Transportadores de tipo GLUT e do tipo SGLT foram parcialmente sequenciados em *D. peruvianus* e

suas atividades foram demonstradas *in vivo* em V1, demonstrando também cotransporte com potássio (Bifano et al., 2010). Há, portanto, variação espacial no processo de digestão, seja pela variação proteica entre as membranas seja na variação funcional das porções intestinais, caracterizando a compartimentalização da digestão.

#### **1.4. Encaminhamento celular de enzimas**

Insetos em geral apresentam na luz de seu intestino médio uma estrutura quitino-proteica chamada membrana peritrófica. Esta membrana mantém o alimento afastado do epitélio e, dentre várias funções, garante a compartimentalização da digestão. Os Hemiptera, por sua vez, não apresentam membrana peritrófica, mas sim as membranas perimicrovilares (Terra e Ferreira, 2012). A presença da dupla barreira causada pelas duas membranas na superfície apical das células inviabilizaria a secreção de enzimas pelas vias clássicas em Hemiptera (Silva et al., 1995).

O encaminhamento de enzimas, de maneira ampla, se inicia na rede trans-Golgi que, baseada em sinais, dirige as enzimas para três vias. Proteínas marcadas com manose-6-fosfato (M6P) são encaminhadas para os lisossomos. As duas outras vias de distribuição de proteínas na rede trans-Golgi são: a via secretora constitutiva e o direcionamento para vesículas secretoras mediado por sinal (secreção regulada). Em ambos os casos o sinal que direciona as proteínas secretoras não é conhecido (Alberts et al., 2007).

Na via clássica de encaminhamento de hidrolases e proteínas de membranas ao lisossomo, proteínas sintetizadas no retículo endoplasmático rugoso são transportadas através do complexo de Golgi até a rede trans-Golgi. Neste momento, receptores de M6P reconhecem a M6P presente nas proteínas que serão encaminhadas aos lisossomos. A associação desses receptores às proteínas adaptadoras de clatrina auxilia a empacotar vesículas revestidas de clatrina que brotam da rede trans-Golgi (Alberts et al., 2007).

A vesícula encaminhada para o sistema lisossomo/endossomo sofre fusão com outras vesículas. Há então uma acidificação do lúmen tornando ativas as hidrolases enviadas. A atuação de uma fosfatase ácida remove o grupo fosfato da

manose, destruindo o sinal de classificação e contribuindo para a liberação das hidrolases lisossômicas dos receptores de M6P (Alberts et al., 2007).

Seria interessante analisar se proteínas que tipicamente utilizam o encaminhamento clássico, como hidrolases ácidas, apresentam marcação por M6P nos Hemiptera e se esse sinal está ou não presente em proteínas secretadas para o lúmen.

O estudo dos lisossomos e do mecanismo de encaminhamento de proteínas nos permitiria distinguir as enzimas tipicamente lisossômicas daquelas que agora realizam função digestiva nos insetos. Lisossomos apresentam proteínas solúveis e ligadas à membrana que são a eles encaminhadas devido à presença de um sinal. A maioria das hidrolases ácidas solúveis é modificada com resíduos de manose-6-fosfato (M6P), permitindo o reconhecimento pelos receptores de M6P no complexo de Golgi e subsequente transporte para o sistema lisossomo/endossomo. Há também o transporte realizado independentemente da M6P mediada por receptores alternativos como LIMP-2 e sortilina (Braulke e Bonifacino, 2009).

Bagshaw et al. (2005) comparou as proteínas identificadas por análise proteômica em lisossomos purificados de células de mamíferos com outros trabalhos feitos com fagossomos e melanossomos. Com isso é possível identificar candidatos para marcadores celulares com maior especificidade para lisossomos. Uma proteína que se mostrou promissora é a fosfatase ácida, que se apresenta bem expressa em lisossomos, mas não foi encontrada em fagossomos e melanossomos. Contudo, estudos realizados em *Rhodnius prolixus* (Terra et al., 1988) demonstraram haver uma fosfatase ácida com características típicas da fosfatase ácida lisossômica de mamífero (como especificidade por substrato e respostas a inibidores) no lúmen intestinal. Neste caso seria importante saber se a fosfatase ácida luminal é a mesma presente nos lisossomos ou se são distintas. Uma busca por outras enzimas tipicamente lisossômicas no interior luminal intestinal seria também interessante.

### **1.5. Enzimas digestivas em Hemiptera e suas origens**

Admite-se que o ancestral dos Hemiptera se alimentava de seiva semelhante aos atuais Auchenorrhyncha sugadores de floema. A seiva (tanto do floema quanto do xilema) é em geral pobre em polímeros como proteínas e polissacarídeos e relativamente pobre em aminoácidos livres. Em comparação com a seiva do xilema,

a seiva do floema é rica em sacarose. Portanto, exceto pela hidrólise de sacarose, nenhuma digestão de alimento é realmente necessária em sugadores de seiva. Em resposta à diluída seiva do floema e xilema, os ancestrais dos Hemiptera teriam perdido a digestão inicial e intermediária associada à falta de digestão luminal (Terra e Ferreira, 2012).

A evolução dos Heteroptera está associada ao retorno da habilidade de digerir polímeros. Devido à perda das enzimas digestivas originais, foram utilizadas enzimas derivadas dos lisossomos para digestão do alimento ingerido (Houseman e Downe, 1983a). A habilidade de digerir polímeros pode ser visto nos Heteroptera basais atuais que são majoritariamente predadores, sendo que alguns grandes e recentes grupos retornaram a fitofagia (Grimaldi e Engels, 2005).

Lisossomos são organelas celulares envolvidas na digestão intracelular com atuação de proteinases especiais chamadas de catepsinas (De Duve, 2005). A digestão nos dois maiores táxons de Heteroptera, Cimicomorpha (como o barbeiro *Rhodnius prolixus*) e Pentatomorpha (como o manchador do algodão *Dysdercus peruvianus*), é similar e apresenta proteinases digestivas do tipo catepsina. A maioria da digestão de proteínas se dá no lúmen com a ação de uma cisteíno-proteinase e termina no espaço perimicrovilar sob a ação de aminopeptidases e dipeptidases (Terra e Ferreira, 2012).

Há indícios que em certos táxons de insetos enzimas tipicamente lisossômicas passaram a participar do processo digestivo. Nos Diptera Cyclorrhapha, por exemplo, temos a atividade de catepsina D no lúmen da porção ácida de seu intestino médio (Espinoza-Fuentes e Terra, 1987; Padilha et al., 2009). A catepsina D tem ação tipicamente lisossômica. Contudo, neste grupo de insetos, um conjunto de características parece ter ocorrido como resposta a uma alimentação rica em bactérias (Espinoza-Fuentes e Terra, 1987), dentre elas a presença de uma catepsina D digestiva. Desta forma, provavelmente uma duplicação gênica possibilitou que atividade lisossômica de catepsina D se mantivesse e sua cópia seguisse outro caminho evolutivo, participando agora da digestão (Padilha et al., 2009).

No caso dos Coleoptera Cucujiformia sua alimentação provavelmente de grãos, que são ricos em inibidores, gerou uma pressão para que esses besouros apresentassem defesas contra a constante inibição de suas serino proteinases pelos inibidores produzidos pelas plantas. A resposta deste grupo de besouros foi a

expressão de uma enzima do tipo cisteino proteinase que possibilita a digestão de proteínas mesmo em uma dieta rica em inibidores de serino proteinases (Cristofolletti et al., 2005). Contudo, sabemos que as cisteino proteinases como a catepsina L e B são enzimas tipicamente lisossômicas (Kirschke, H, 2004). Assim, temos mais uma vez a presença de uma enzima tipicamente lisossômica fazendo agora parte do processo digestivo de um organismo.

Há em Heteroptera, no entanto, um grande arsenal de enzimas lisossômicas atuando na digestão como fosfatase ácida, catepsina D, catepsina B e catepsina L (Houseman e Downe, 1980; Houseman e Downe, 1982; Terra et al., 1988). A presença dessas enzimas gera a intrigante questão de como foi possível alterar o encaminhamento celular de enzimas dos lisossomos para o lúmen intestinal.

### **1.6. Enzimas digestivas de *D. peruvianus***

Acompanhar o processo de secreção de enzimas e analisar seus possíveis sinais de encaminhamento permitiria compreender melhor qual a base molecular do tráfego de proteínas em um sistema tão complexo como o dos Hemiptera.

Quando o alimento entra no V1 do *D. peruvianus*, os sacarídeos são hidrolisados pela alfa-glicosidase ligada à membrana perimicrovilar. Há em V1 a atividade de uma alfa-manosidase. A pequena atividade de amilase no lúmen parece refletir o baixo teor de amido das sementes de algodão. Grande parte da proteína é digerida em V2 sob a ação de uma cisteino proteinase e uma aminopeptidase (Silva e Terra, 1994).

Silva e colaboradores (1996) demonstraram, com experimentos de ultracentrifugação em gradiente de sacarose, a associação de alfa-glicosidases e beta-glicosidases com membranas perimicrovillares e microvillares, respectivamente.

Silva e Terra (1997) demonstraram a presença de uma alfa-galactosidase produzida pelo *D. peruvianus* que, juntamente com atividade de alfa-galactosidase presente na semente do algodão, hidrolisa rafinose (abundante nas sementes usadas) que libera galactose e sacarose. A sacarose seria posteriormente hidrolisada pela alfa-glicosidase presente na membrana perimicrovilar. A galactose, glicose e frutose derivadas das hidrólises seriam absorvidas, principalmente em V1.

A compilação de enzimas cujas atividades foram descritas no intestino médio de *Dysdercus peruvianus* e cujo mecanismo de secreção poderia ser estudado está presente na tabela 1.

**Tabela 1.** Compilação das enzimas descritas em *Dysdercus peruvianus*. MPMV: membrana perimicrovilar; EIM: espaço inter membranas; MMV: membrana microvilar.

Enzima	Estado de Purificação	Localização	Referência
Amilase	Não purificada	Luminal	Silva e Terra, 1994
Fosfatase ácida	Não purificada	Luminal	Silva e Terra, 1994
Fosfatase alcalina	Não purificada	Luminal + celular em MPMV	Silva e Terra, 1994; Silva <i>et al.</i> , 1996
Aminopectidase	Purificada	Luminal + celular em EIM	Silva e Terra, 1994; Silva <i>et al.</i> , 1996; Costa <i>et al.</i> , 2011
alfa-Galactosidase	Purificação Parcial	Luminal + celular em EIM	Silva e Terra, 1997
alfa-Glicosidase	Purificada	Celular em MPMV	Silva e Terra, 1995; Silva <i>et al.</i> , 1996
beta-Glicosidase	Não purificada	Luminal + celular em MMV	Silva e Terra, 1994; Silva <i>et al.</i> , 1996
alfa-Manosidase	Não purificada	Luminal + celular em MMV	Silva e Terra, 1994
Catepsina L	Purificada	Luminal	Silva e Terra, 1994; Bifano <i>et al.</i> , não publ.

A tabela 1 mostra que várias enzimas descritas no intestino médio não foram ainda isoladas e, portanto, não tiveram suas caracterizações cinéticas realizadas. Das enzimas já isoladas a alfa-glicosidase teve caracterizada sua atividade sobre diferentes substratos (sob a ação de diferentes inibidores), pH ótimo, massa molecular e presença de glicosilação (Silva e Terra, 1995) e, exceto pela glicosilação, a aminopectidase teve os mesmos parâmetros calculados (Costa *et al.*, 2011). Para a melhor compreensão da atuação das enzimas presentes no intestino médio de *D. peruvianus* seria importante estender os estudos cinéticos para as outras enzimas citadas na tabela 2.

### 1.7. Catepsina D (aspártico proteinase)

Inicialmente isolada do baço bovino, a catepsina D é uma enzima tipicamente lisossômica (Press *et al.*, 1960; De Duve, 2005). Essa enzima é uma endopeptidase com pH ótimo entre 3,5 e 5,0, dependendo do substrato e das condições de ensaio. A especificidade da quebra da cadeia polipeptídica do substrato pela catepsina D é



similar à pepsina ao preferir aminoácidos hidrofóbicos na ligação a ser hidrolisada. Índícios experimentais indicam preferência por quebra entre dois resíduos hidrofóbicos. É inibida por pepstatina A, assim como outros membros da família. A enzima precursora tem peptídeo sinal N-terminal que é clivado durante a translocação pela membrana para gerar a procathepsina D inativa (similar ao pepsinogênio). A procathepsina D sofre hidrólise autocatalítica para a forma ativada (Conner, 2004).

Considerando os mamíferos, a catepsina D é uma proteinase lisossômica que constitui em torno de 10% das proteínas lisossômicas solúveis do fígado de rato e pode ser encontrada em praticamente todos os tecidos animais. É codificada por apenas um gene em humanos e ratos e expressa de maneira constitutiva na maioria das células. Pode haver regulação positiva em alguns casos (Conner, 2004). Além da atividade lisossômica, há descrição de catepsina D secretada em alguns insetos (Padilha *et al.*, 2009, Balczun *et al.*, 2012).

Há modificação por manose-6-fosfato que serve para direcionar a enzima para lisossomos através da associação específica com receptores manose-6-fosfato. A glicosilação não é necessária para o dobramento ou para atividade enzimática, mas sim para o direcionamento da enzima (Fortenberry *et al.*, 1995).

### **1.8. Catepsina L (cisteíno proteinase)**

Cisteíno peptidases fazem a hidrólise de ligações peptídicas. O grupo sulfidrila do resíduo de cisteína hidrolisa a ligação peptídica por ataque nucleofílico. Nesse processo há participação de um resíduo de histidina na formação de um par iônico tiolato-imidazólico. Um resíduo de glutamina ajuda na formação da cavidade do oxianion e o resíduo de asparagina ajuda na orientação do anel imidazólico, mantendo assim a estabilidade do sítio ativo (Kirschke, 2004). As cisteíno proteinases são ativadas por grupos sulfidrilas (DTT, cisteína) e por agentes quelantes e são inibidas por E-64.

As catepsinas L são classificadas dentro do grupo das cisteíno peptidases na subfamília C1A que inclui tanto enzimas secretadas como enzimas lisossômicas. A catepsina L tem atividade predominantemente de endopeptidase, preferindo hidrolisar ligações peptídicas com resíduos aromáticos em P2 e hidrofóbico na posição P3. Catepsinas L são cataliticamente ativas em pHs 3,0-6,5 na presença de

tiol e são sintetizadas como preproenzimas (Kirschke, 2004). Catepsinas L são glicosiladas e têm a manose fosforilada no Golgi, sendo o propeptídeo fundamental para o correto dobramento da proteína, estabilidade e transporte para o complexo de Golgi (Tao, *et al.*, 1994).

### 1.9. Aminopeptidase N de insetos

Aminopeptidase N (APN) é uma família de exopeptidases. Aminopeptidases que estão presentes nos insetos são, de maneira geral, solúveis em insetos basais e ligadas a membrana microvilar em insetos derivados (Terra e Ferreira, 2012). A atividade das APN intestinal de insetos se dá pela clivagem do aminoácido N-terminal de peptídeos, atividade necessária para a liberação de aminoácidos para o cotransporte nas células epiteliais (Adang, 2004). Aminopetidases purificadas de *Tenebrio molitor* e *Acyrtosiphon pisum* apresentam uma ampla especificidade por diferentes aminoácidos N-terminais, mas não hidrolisam peptídeos com aminoácidos ácidos. Têm em comum também a preferência na hidrólise de tripeptídeos (além de tetrapeptídeos no caso de *T. molitor*) em relação à di e pentapeptídeos (Cristofolletti e Terra, 1999; Cristofolletti *et al.*, 2003).

As APN de Lepidoptera foram bastante estudadas por ligar a toxina inseticida Cry1 do *Bacillus thuringiensis*. A toxina, quando ingerida, se liga ao intestino e se insere na membrana celular, formando um poro capaz de matar o inseto (Sangadala *et al.*, 1994). Essa toxina é usada na engenharia genética de plantas para o controle de pragas.

Considerando os Hemiptera, o Sternorrhyncha *A. pisum* apresenta atividade de aminopeptidase na fração de membranas correspondente às membranas lamelares (homólogas às membranas microvillares) (Cristofolletti *et al.*, 2003). Já nos Heteroptera *D. peruvianus* e *R. prolixus*, a atividade de aminopeptidase estaria na forma solúvel no espaço entre a membrana microvilar e a perimicrovilar (Silva *et al.*, 1996; Ferreira *et al.*, 1988).

### 1.10. Alfa-glicosidase na digestão em insetos

As alfa-glicosidases (EC 3.2.1.20) são descritas como enzimas que catalizam hidrólise de resíduos de glicoses terminais não-redutoras em ligações alfa1-4 com a

liberação de alfa-D-glicose (<http://www.chem.qmul.ac.uk/iubmb/enzyme/EC3/2/1/20.html>). Alfa-glicosidases em insetos fazem parte na digestão como enzima presa à membrana microvilar ou solúvel. A enzima solúvel foi descrita no lúmen intestinal de coleópteros e himenópteros e adsorvida ao glicocálix em lepidópteros e dípteros. Alfa-glicosidases integrantes de membrana foram descritas em membrana microvilar de dípteros (Terra e Ferreira, 2012). Em estudos de fracionamento celular, a atividade de alfa glicosidase foi encontrada em *Rhodnius prolixus* associada às grandes estruturas membranosas (de baixa densidade), indicando fazer parte da membrana perimicrovilar (Ferreira et al., 1988) . Foi descrita para *D. peruvianus* atividade de alfa-glicosidase presa à membrana perimicrovilar (Silva e Terra, 1994; Silva et al., 1996). Esse importante marcador celular foi isolado cromatograficamente e suas características cinéticas descritas. A alfa glicosidase de membrana isolada tem pH ótimo de 5,0, apresenta maior preferência por maltose do que por sacarose e a eficiência catalítica sobre maltose, maltotriose e maltotetraose são similares, diminuindo a eficiência para maltopentaose, maltohexaose e maltoheptaose (Silva e Terra, 1995).

## 2. Objetivos

Sabendo que membrana perimicrovilar, característica dos Hemiptera, é formada pela secreção de vesículas de membrana dupla temos como objetivo, usando *Dysdercus peruvianus*:

(a) Verificar o mecanismo de secreção de catepsina L e catepsina D digestivas e avaliar a possibilidade da secreção incluir outras enzimas tipicamente lisossômicas.

(b) Analisar a formação de vesículas intracelulares de dupla membrana utilizando marcadores celulares.

(c) Testar e ampliar o modelo existente de absorção de água e nutrientes, incluindo a identificação das moléculas envolvidas.

(d) Caracterizar *in silico* e cineticamente enzimas digestivas e lisossômicas que se demonstrarem favoráveis.

### **3. Procedimentos experimentais**

#### **3.1. Animais**

Foram utilizados insetos adultos do percevejo *Dysdercus peruvianus* (Hemiptera: Pentatomorpha: Pyrrhocoridae) obtidos a partir de uma colônia mantida em nosso laboratório desde 1990. A manutenção dos insetos é feita mantendo-os em viveiros plásticos contendo areia no fundo, coberto por tecido e com livre acesso a água e sementes de algodão (*Gossypium hirsutum*). A colônia é mantida sob fotoperíodo natural com umidade relativa de 50-70% e temperatura de 22-26°C.

#### **3.2. Dissecção de animais**

Na obtenção de homogeneizados foram utilizados insetos adultos, fêmeas, com de 2 a 9 dias após a última ecdise e alimentados normalmente com semente de algodão. Os animais foram imobilizados em gelo e dissecados em solução salina (NaCl 215 mM). Foram retirados os diferentes ventrículos que eram usados com seus conteúdos, para obtenção de proteínas, ou apenas o epitélio, para experimentos de centrifugação diferencial.

Inicialmente, para obtenção de amostra proteica os tecidos eram imediatamente colocados em tubos plásticos em contato com gelo. Em experimentos, que serão apresentados no material suplementar II, foi testada a imersão imediata do tecido em água contendo ou não inibidor de protease, em tubos em contato com gelo ou gelo seco.

Para o processo de homogeneização para obtenção de amostra proteica foi feito em homogeneizador Potter-Elvehjem com o pistilo sob rotação contínua e dez movimentos de sobe e desce. Isso foi feito em água contendo inibidores de protease (10  $\mu$ M de pepstatina e 10  $\mu$ M de E-64). O processo de homogeneização para obtenção das frações subcelulares foi feito com o homogeneizador Potter-Elvehjem com o pistilo de Teflon sem rotação, em quinze movimentos de sobe e desce. A amostra foi homogeneizada em salina (NaCl 215 mM) com inibidores de protease (10  $\mu$ M de pepstatina + 10  $\mu$ M de E-64 ou MTS 1mM). Em ambos os casos a homogeneização era feita com o homogeneizador em gelo.

### 3.3. Extração de RNA

Para a obtenção de RNA dos tecidos da *D. peruvianus* foi preciso dissecar os insetos adultos utilizando material que permaneceu imerso em água DEPC (dietil pirocarbonato) 0,1% por 1h a 37°C, foi autoclavado e ficou por pelo menos 4h na estufa a 150°C. Outros materiais necessários no processo de extração sofreram o mesmo processo para que diminuíssem as possibilidades de contaminação por RNAses.

O tecido correspondente ao intestino médio completo foi utilizado para produção do cDNA molde para as clonagens. Já para o experimento de RT-PCR, foram extraídos de aproximadamente 2-6 adultos (dependendo do tecido) o corpo gorduroso, glândulas salivares, túbulos de Malpighi e frações do intestino médio de acordo com o experimento. Em experimentos de qPCR, 3 adultos foram dissecados para obtenção das frações do intestino médio e o restante (corpo total menos a intestino médio) foi chamado de carcaça.

Cada tecido foi homogeneizado separadamente em 1 ml de Trizol (Invitrogen) à temperatura ambiente e transferidos para tubos livres de RNAses. No caso da carcaça, é utilizado almofariz resfriado em gelo seco e o tecido é moído junto com o volume correspondente de Trizol. Ao resfriar, todo o material foi drenado para tubos livre de RNAses. O material foi centrifugado a 11.000 rpm por 10 min, sendo descartado o sedimento. Ao sobrenadante foram acrescentados 0,2 ml de clorofórmio, mantido por 15 s sob agitação e mais 3 min de incubação a temperatura ambiente. O material foi centrifugado a 11.000 rpm por 15 min. Novamente o sedimento foi descartado. Ao sobrenadante foram acrescentados 250 µl de isopropanol e 250 µl de solução de precipitação (NaCl 1,2M e citrato de sódio 0,8M em água DEPC 0,1%), incubando por 10 min a temperatura ambiente. O material foi centrifugado a 11.000 rpm por 10 min e, descartou-se, desta vez, o sobrenadante. O próximo passo foi acrescentar ao precipitado 500 µl de etanol 75%, e centrifugado a 11.000 rpm por 5 min, sendo seu sobrenadante desprezado. Esse passo foi repetido mais uma vez. O sedimento foi seco por 3 min no fluxo laminar e por fim acrescentaram-se 100 µl de água DEPC 0,1% que permaneceu por 5 min incubando a 65°C. A quantificação do RNA foi feita pela leitura da absorbância do material em um comprimento de onda de 260 nm. Para tal medição foi necessário o uso de um espectrofotômetro (uma unidade de absorbância = 40 µg de RNA). Para a

verificação da pureza e integridade do material extraído utilizou-se a razão da absorvância obtida em 260 nm pela absorvância obtida em 280 nm (valores ideais encontram-se no intervalo 1,9 – 2,1).

### 3.4. Produção de cDNA

O RNA extraído foi tratado com DNase I (Invitrogen) na proporção 1µg de RNA total para 1 µl de DNase (1U/µl) e incubado conforme protocolo.

A produção do cDNA a partir do RNA total tratado com DNase, obtido da extração de diferentes tecidos da *D. peruvianus*, foi feita com o uso do kit SuperScript™ (Invitrogen). Esse kit possuía uma transcriptase reversa que transcreveu o RNA, formando a primeira fita de cDNA na presença de iniciadores aleatórios (no caso de experimentos de RT-PCR) ou iniciador poliT (no caso de experimentos de clonagem). Em experimentos de RT-PCR, para que as quantidades de cDNA correspondessem ao nível de expressão dos genes em um dado tecido, foi necessário iniciar cada reação (correspondente a cada tecido) com concentrações equivalentes de RNA total.

O protocolo se iniciou pela mistura de quantidades iguais de RNA para os diversos tecidos (conforme foi explicado acima) a 1 µL de dNTPmix, 5 µL de hexameros randômicos (como iniciadores) e 10 µL de água DEPC. Essa mistura foi incubada por 5 min a 65°C e posteriormente incubada 1 min em gelo.

Foi preparada a mistura de reação contendo (para cada reação): 2 µL de RT Buffer 10X, 4 µL MgCl<sub>2</sub> 25 mM, 2 µL DTT 0,1 M e 1 µL de RNaseOUT™ Recombinant Ribonuclease Inhibitor. Após, 9µL da mistura de reação foram adicionados ao volume contendo o RNA e iniciadores e o total foi incubado por 2 min a 25°C. Houve a adição de 1µL de enzima transcriptase reversa SuperScript™ sendo, a reação, incubada por 10 min a 25°C. Após, foi incubada por 50 min a 42°C. A reação foi finalizada incubando os tubos contendo a reação por 15 min a 70°C. Os tubos eram postos em gelo e tinham 1 µL de RNase H (Invitrogen) adicionado, para então serem incubados por 20 min a 37°C.

Com todo o processo descrito, obteve-se o cDNA que seria usado na reação de PCR (reação em cadeia da polimerase).

### 3.5. PCR e RT-PCR Semiquantitativo

O RT-PCR teve neste experimento o objetivo de amplificar o cDNA obtido com o uso de uma transcriptase reversa, a partir de diferentes tecidos de adultos de *D. peruvianus*. O tecido usado variou conforme o experimento. As diferentes formas de fracionar o intestino médio foram escolhidas conforme necessidades específicas. As glândulas salivares, Túbulos de Malpighi e corpo gorduroso são tecidos de origens embriológicas distintas do intestino médio, sendo útil para verificar o grau de abrangência na expressão de um gene.

Diversos genes foram testados. Teve como controle positivo do material utilizado, uma sequência de gene para produção de RNA ribossomal (como gene constitutivo). Os iniciadores correspondentes às sequências utilizadas neste trabalho estão alistados na tabela suplementar 1.

Os fragmentos de cDNA codificando as respectivas sequências foram amplificados a partir do material produzido na reação da transcriptase reversa sobre o RNA dos diferentes tecidos. Tanto para experimentos de RT-PCR quanto para reação de PCR usuais, cada reação foi montada da seguinte forma:

1 µL	cDNA (RT)/ DNA (amostra)
0,2 µl	iniciador 1 (100 µM)
0,2 µl	iniciador 2 (100 µM)
0,25 µl	TAQ DNA Polymerase (Invitrogen)
5 µl	Tampão 10 X
1 µl	dNTP (10 mM)
2 µl	MgCl <sub>2</sub> (50 mM)
40,35 µl	H <sub>2</sub> O milliQ

A reação de PCR foi realizada em termociclador, com o programa abaixo:

5 min a 94°C.

Em seguida foram realizados 30 ciclos conforme especificados:

30 s a 94°C (desnaturação das fitas molde)

30 s a 55°C (ou gradiente de 50 a 60°C) (pareamento dos iniciadores)

0,5-1,5 min a 72°C (alongamento das fitas filhas)

No caso de RT-PCR, a partir do 20º ou do 25º ciclo, amostras de 5 µL foram coletadas a cada 5 ciclos. Isso foi feito para obter todas as reações em fase logarítmica. As amostras retiradas foram aplicadas em gel de agarose 0,1% para



eletroforese (ver eletroforese em gel de agarose). Para reação de PCR tradicional foram utilizados 30 ciclos.

### **3.6. Pirosequenciamento**

Foi extraído RNA total do intestino médio de adultos machos e fêmeas de *D. peruvianus* conforme especificado no item Extração de RNA. Este material, depois de verificado os parâmetros de qualidade, foi enviado para o Centro Avançado de Tecnologias em Genômica (CATG) do Instituto de Química da USP. Foi utilizada a plataforma 454 da Roche (454 GS FLX) para sequenciamento do cDNA de acordo com os procedimentos padrões definidos pela Roche. O arquivo gerado (.sff) contendo os reads foi processado pelo GS De Novo Assembler (Newbler, Roche) que gerou os contigs que formaram a base de dados para anotação e busca de sequências.

### **3.7. Análise de sequências**

As sequências de nucleotídeos e aminoácidos (por tradução) obtidas neste trabalho foram comparadas com outras sequências depositadas em bancos de dados. Para isso foi feita uma consulta ao “National Center for Biotechnology Information” (<http://www.ncbi.nlm.nih.gov>) onde foi utilizado o programa de alinhamento de sequências Blast X, quando era analisada a sequência de nucleotídeos, e o programa Blast P, quando a sequência analisada era de aminoácidos.

A análise de grande volume de sequências provenientes do pirosequenciamento foi possível com o uso do software dCAS – Desktop cDNA Annotation System ([HTTP://exon.niaid.nih.gov](http://exon.niaid.nih.gov)). Com essa ferramenta foi possível excluir sequências de vetores, montar contigs e anotar usando BLAST X contra vários bancos de dados (por ex. nr, pfam, GO). As sequências de interesse passaram por um tratamento manual que envolve correções de janela de leitura (ou *frame*). Mesmo com a montagem das sequências com o programa Newbler foi necessária essa curadoria manual das sequências, pois muitas delas apresentavam deleções e inserções. Essas deleções e inserções são típicas do processo de

sequenciamento utilizado, necessitando assim de um extenso trabalho após a montagem automática das sequencias.

Para predição de peptídeo sinal, foi utilizada a ferramenta SignalP 3.0 (<http://www.cbs.dtu.dk/services/SignalP/>). A previsão de sítios de N-glicosilação foi realizado com o programa NetNGlyc 1.0 (<http://www.cbs.dtu.dk/services/NetNGlyc/>) com limiar de 0,5 (de valores de 0 a 1) para o potencial de N-glicosilação calculado pelo algoritmo. Para a previsão de hélices transmembranas das sequencias das preproteínas foi utilizado o programa SOSUI engine (<http://bp.nuap.nagoya-u.ac.jp/sosui/>). Previsões de sítios de modificação por âncora de glicosilfosfatidilinositol (GPI) foram feitas com o bigPI Predictor Server ([http://mendel.imp.ac.at/sat/gpi/gpi\\_server.html](http://mendel.imp.ac.at/sat/gpi/gpi_server.html)), selecionando a opção Metazoa.

O alinhamento de sequencias foi realizado com o programa BioEdit (v. 7.0.5.3) (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>), utilizando sua interface interna com o programa de alinhamento múltiplo ClustalW.

A produção de arquivos com a modelagem tridimensional por homologia da DPCAL1 foi realizada com o uso do programa SWISS-MODEL (<http://swissmodel.expasy.org/>), sendo a modelagem feita no modo automático. O modelo estrutural foi visualizado e modificado para identificação dos resíduos de interesse pelo uso do programa PyMOL v0.99 (<http://www.pymol.org/>).

A geração de dendrogramas das sequencias foram feitas com o programa MEGA 4 (<http://www.megasoftware.net/>), utilizando o algoritmo Neighbor-Joining, avaliação estatística por Bootstrap a partir de 1000 réplicas. Os ramos do dendrograma com Bootstrap abaixo de 50 foram colapsados.

### **3.8. Reação de sequenciamento de DNA segundo Sanger**

O sequenciamento de DNA segundo Sanger foi realizado na central de sequenciamento do Departamento de Bioquímica do Instituto de Química da USP – São Paulo. Para isso foi utilizado um aparelho ABI 377 DNA Sequencer (Applied Biosystems) e a técnica foi executada conforme recomendação do fabricante. A reação de sequenciamento foi realizada utilizando 4 µl da mistura de reação Big Dye terminator mix (Perkin Elmer), 200 – 500 ng de DNA, 20 pmols de oligonucleotídeos, tampão Tris 40 mM, pH 9,0, MgCl<sub>2</sub> 1 mM em um volume final de 20 µl. A reação foi incubada por 1 min a 96°C e em seguida submetida a 35 ciclos (96°C por 30 s, 50°C

por 30 s e 60°C por 4 min). Ao final, a reação foi precipitada com etanol de acordo com a orientação do fabricante, e ressuspendido em tampão para aplicação no sequenciador automático.

### **3.9. Ligação dos fragmentos de cDNA a plasmídeos vetores**

Para clonagem (processo de subclonagem) de fragmentos de cDNAs obtidos de amplificação por PCR, foi utilizado o Kit de ligação “pGEM-T Easy vector System” (Promega). Foi empregada uma mistura de reação que consistia de 1 U de T4 DNA ligase, tampão apropriado, plasmídeo (pGEM-T) e o produto de PCR (purificado do gel de agarose). O plasmídeo e o produto de PCR foram utilizados em uma concentração de 1:3. Essa mistura foi incubada por 16 h a 4°C. O produto de ligação foi utilizado para transformar bactérias competentes.

Para inserir fragmentos de cDNAs no vetor de expressão pAE (Ramos *et al.*, 2004), tanto o inserto (ou construção da subclonagem) quanto o vetor foram digeridos com as enzimas de restrição, purificados do gel de agarose e a seguir foram misturados numa proporção de 3 volumes de inserto por 1 volume de plasmídeo, adicionando 1U de T4 DNA ligase (Invitrogen) e tampão apropriado. A amostra foi incubada a 16°C por 16 h e a seguir usada para transformar bactérias competentes. A clonagem do inserto foi realizada de forma que a proteína seja expressa com sete histidinas na porção N-terminal da proteína recombinante. Esse procedimento facilita, em muitos casos, o processo de purificação em colunas de afinidade.

Os plasmídeos foram purificados utilizando-se o protocolo Wizard Miniprep (Promega) a partir de 3,0 ml de cultura de bactérias cultivadas durante 16 h.

### **3.10. Transformação de bactérias competentes para expressão**

A construção de plasmídeo pAE contendo como inserto sequências de catepsinas L foram adicionadas a 50 µL de células competentes (BL21 STAR<sup>tm</sup>-DE3) e mantidas em gelo por 30 min. Em seguida as células foram submetidas a um choque térmico a 42°C por 45 s e novamente transferidas para o gelo por mais 2 min. Adicionamos então, 0,95 ml de meio LB (10 g NaCl, 10 g bactotripton e 5 g de extrato de levedura, e água até 1,0 l) e essa solução foi incubada por 1 h a 37°C,

para recuperação das bactérias e expressão das proteínas envolvidas na resistência ao antibiótico de seleção. Em seguida, a mistura era plaqueada em LB ágar com carbenicilina (50 µg/ml) e incubadas “overnight” a 37°C, obtendo-se dessa forma colônias isoladas.

### **3.11. Expressão de proteínas recombinantes com vetor pAE**

Bactérias *Escherichia coli* BL21 STAR foram transformadas com os plasmídeos pAE (Ramos *et al.*, 2004) com insertos codificantes das DPCAL1 e 3. As colônias transformadas eram crescidas em meio LB contendo ampicilina (50 µg/ml) sob agitação de 250 rpm a 37°C até atingirem  $A_{600nm} = 0,6 - 1,0$ , quando eram induzidas com IPTG 1.0 mM, para produzir as proteínas recombinantes, por uma noite a 37°C sob agitação de 250 rpm. Após esse procedimento, as células eram sedimentadas por centrifugação a 7.000 g por 30 min a 4°C e então armazenadas a -20°C. Foram coletadas alíquotas de 20 µl de bactérias não-induzidas e induzidas a fim de verificar através de eletroforese desnaturante (SDS-PAGE, vide abaixo) se houve produção da proteína recombinante.

### **3.12. Lise das células induzidas**

Após o processo de indução, 50 ml do meio contendo as células era centrifugado a 4-6°C e o sedimento era ressuspensão em 1 ml de tampão de lise (tampão fosfato de sódio 20 mM, NaCl 150 mM, pH 7,5). Após esse procedimento, as células eram sonicadas com 3 pulsos de 15 s e intervalos de 1 min em gelo. O “output” usado foi 3, em um sonicador Branson com “microtip”. Em seguida, as amostras eram centrifugadas a 13000 rpm por 30 min a 4°C. O sobrenadante e o precipitado eram armazenados a -20°C separadamente para posterior purificação da proteína recombinante ou análise por SDS-PAGE.

### **3.13. Purificação de proteína recombinante**

O produto solubilizado em ureia 8 M passou por um processo de purificação, utilizando cromatografia de afinidade Ni-NTA agarose (Qiagen). A purificação foi possível devido à interação da cadeia lateral das histidinas (adicionadas a sequencia

proteica) com a resina de níquel, sendo que a presença de imidazol no meio compete pela ligação à resina, liberando a proteína ligada.

Inicialmente 300 µl de suspensão de resina Ni-NTA agarose foi colocado em um microtubo (1,5 ml) e centrifugado por 1 min a 13000 rpm a 4°C. Essa condição de centrifugação e temperatura foi mantida em todo o experimento. Em seguida, o sobrenadante foi descartado e a resina foi ressuspendida em 600 µl, centrifugada e teve o sobrenadante descartado. Esse procedimento foi repetido mais duas vezes. A solução usada neste procedimento corresponde ao tampão de lavagem (tampão fosfato 50 mM, NaCl 100 mM, ureia 8 M, imidazol 10 mM em pH 7,5).

Cerca de 150 µl de amostra foram adicionados à resina com mais 450 µl de tampão de amostra e mantido em gelo por 10 min. O material foi centrifugado e o sobrenadante foi coletado. Em seguida, 600 µl de tampão de lavagem foram adicionados ao material que foi ressuspendido, centrifugado e teve o sobrenadante coletado. Esse procedimento foi repetido mais duas vezes.

Cerca de 600 µl de tampão de eluição (tampão fosfato 50 mM, NaCl 100 mM, ureia 8 M, imidazol 500 mM em pH 7,5) foram adicionados ao material que foi ressuspendido, centrifugado e teve o sobrenadante coletado. Esse procedimento foi repetido mais duas vezes.

Os materiais coletados seguiam para diálise 1/500 (v:v) por uma noite a 4-6°C com uma troca. Após a diálise, o material era liofilizado, misturado a 50 µl tampão de amostra para SDS-PAGE e fervido por 5 min.

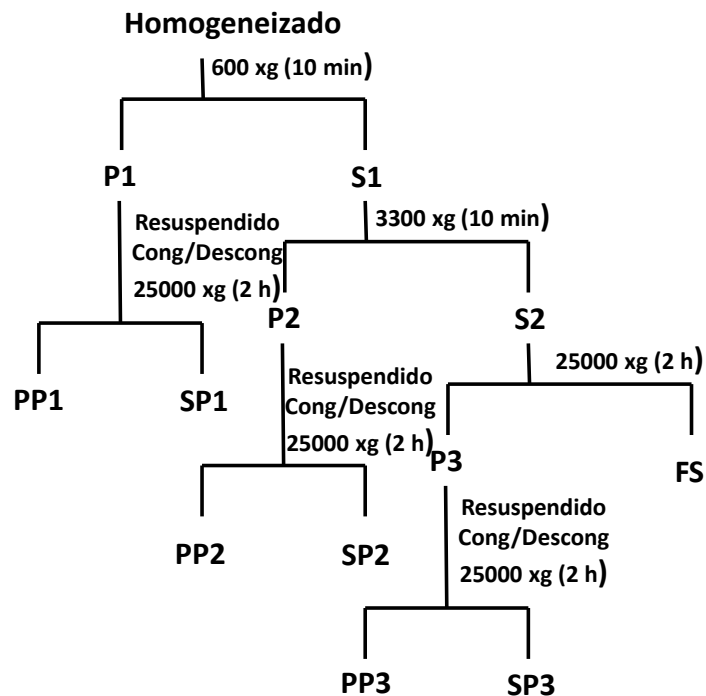
### **3.14. Eletroforeses**

Nas eletroforeses desnaturantes, isso é eletroforese em gel de poliacrilamida na presença de dodecil sulfato de sódio (SDS-PAGE), as amostras de proteínas foram diluídas com tampão de corrida (Tris-HCl 120 mM pH 6,8, SDS 5% (m/v), β-mercaptoetanol 1,3 mM, EDTA 1,0 mM, glicerol 20% (v/v) e azul de bromofenol 0,01% (m/v). Antes da aplicação, as amostras foram fervidas junto com o tampão de amostra por 5 min. As amostras foram corridas sob voltagem constante de 200 V. O gel de poliacrilamida 12% e SDS 1% foi preparado segundo Laemmli (1970). A coloração dos géis foi feita com Comassie Blue R-250 0,1% (p/v) preparado em uma solução de metanol 40% (v/v) e ácido acético 10% (v/v).

A preparação de eletroforese em gel de agarose foi realizada em géis em uma concentração final de agarose 1% (m/v) em tampão TAE (Tris 40 mM, Acetato 20mM, EDTA 1mM, pH 8,0). As amostras de DNA foram misturadas com tampão de amostra (azul de bromofenol 0.25%, xileno cianol 0.25%, glicerol 30% em água) na proporção de 1:10, respectivamente. O azul de bromofenol, devido a sua coloração, auxilia a estabelecer visualmente o final da eletroforese. Após a eletroforese, o gel foi tratado com brometo de etídeo (0,5 mg/ml) para visualização das amostras de DNA, pois o DNA complexado com o brometo de etídeo emite luz laranja na presença da luz ultravioleta (312 nm). Os produtos amplificados em reação de PCR foram recuperados diretamente dos géis de agarose usando o método padrão para extração com uso do produto QIAquick (Qiagen), o qual baseia-se na dissolução da agarose em um tampão de alta força iônica e ligação do DNA em uma resina de troca iônica.

### **3.15. Centrifugação diferencial do homogeneizado**

O homogeneizado foi filtrado em um tecido de Nylon com malha de 45 µm de poro e o filtrado foi submetido a centrifugações a 4°C. As seguintes frações foram coletadas: P1, sedimento resultante da centrifugação a 600 g por 10 min; P2, sedimento de 3300 g por 10 min; P3 sedimento de 25000 g por 2 h; FS, sobrenadante final. Cada sedimento foi resuspenso em solução salina (NaCl 215 mM) na presença dos inibidores de proteases e submetido a três ciclos de congelamento e descongelamento e, finalmente, esse material foi centrifugado a 25000 g por 2 h. O processo foi sumarizado na figura 3. O sedimento e o sobrenadante correspondente a cada fração foi ensaiado para diversas enzimas e a determinação de proteínas realizada com método de Bradford tendo ovoalbumina como padrão.



**Figura 3.** Fluxograma da preparação das frações subcelulares. FS, sobrenadante final; P, sedimento; Cong/Descong, três ciclos de congelamento e descongelamento; PP1, sedimento resultado da centrifugação após a resuspensão e os ciclos de congelamento e descongelamento. SP, sobrenadante resultado da centrifugação após a resuspensão e os ciclos de congelamento e descongelamento.

### 3.16. Ultracentrifugação em gradiente de sacarose

As frações subcelulares P2 e P3, obtidas como já descrito, foram reunidas para o processo de ultracentrifugação em gradiente de sacarose. O sedimento (P2+P3) foi obtido pela centrifugação de homogeneizados de quatro fêmeas e foi ressuspenso em 200  $\mu$ L de salina (250 mM com MTS). O material ressuspenso foi aplicado na parte superior de um gradiente de sacarose (10-50%) feito com tampão imidazol 5 mM, pH 7,0, KCl 40 mM e EDTA 5mM. O volume total de gradiente é de 9 mL.

Após centrifugação de aproximadamente 100000 g por 4 h a 4-6°C, amostras de cerca de 250  $\mu$ L foram coletadas da parte inferior do gradiente e congeladas (-20°C). No dia seguinte as amostras foram congeladas e descongeladas três vezes e seguiram para medidas de atividade enzimática.

### 3.17. Produção de anticorpo

O certificado da Comissão de Ética em Uso de Animais Nº 05/2013 foi aprovado em 26/03/2013. A manutenção do coelho, manuseamento, aplicações e coletas foram realizados pela equipe do Biotério do Instituto de Química da USP.

O coelho (*Oryctolagus cuniculus*) foi utilizado na obtenção de soro policlonais contra as proteínas recombinantes produzidas em *Escherichia coli* (*E. coli*). O coelho passou por quarentena, de 2 a 4 semanas. Uma amostra de 3 mL de sangue (pré-imune) foi coletada dos vasos auriculares após a quarentena. O material inoculado consiste na proteína recombinante semi-purificada que foi recuperada de géis SDS-PAGE. Quinhentos µg da proteína recuperada foram homogeneizados em aproximadamente 1 ml de água, por fim, misturada na proporção 1:1 com adjuvante de Freund completo na primeira inoculação, e de adjuvante de Freund incompleto nas subsequentes. As inoculações têm de 3-4 semanas de intervalo entre elas.

A aplicação do inóculo foi feita pela via subcutânea, utilizando 4 a 5 pontos do dorso do animal, dependendo do volume inoculado, que varia entre 0,5 a 1,5 ml. Após cerca de duas semanas da segunda inoculação, uma nova amostra de sangue de 3 ml foi retirada consistindo o soro imune.

### 3.18. Tratamento do soro

O sangue coletado do coelho foi mantido por 1 h a 37°C e, posteriormente, mantido 16 h a 4°C. Foi feita uma breve centrifugação e retirada do soro (porção aquosa) e esse soro foi misturado em um volume igual de solução saturada de sulfato de amônio ( $\geq 3,4$  M) em tampão fosfato de sódio (100 mM) pH 6,5. A mistura foi mantida por 16 h a temperatura ambiente e, após, centrifugada. O sedimento foi ressuspenso na solução de sulfato de amônio diluída uma vez em água e novamente centrifugado.

O novo sedimento obtido foi diluído em uma solução de NaCl 0,1M e esse material foi submetido a uma diálise contra 1000X o volume da amostra com uma troca após 4 h. O soro foi então alíquotado e congelado até o uso.



### 3.19. Produção do anti-DPAPN (anti-aminopeptidase de *D. peruvianus*)

A estratégia para a produção do anticorpo policlonal anti-DPAPN foi através da imunização de coelho com o peptídeo sintético baseado na sequência proteica encontrada. O peptídeo escolhido KRDFINVPSDITNDC foi sintetizado e conjugado com carreador KLH (hemocianina de molusco) para o processo de imunização do coelho. O soro imune gerado foi purificado com coluna de afinidade com o peptídeo sintético imobilizado. Os procedimentos foram realizados pela empresa FastBio (Ribeirão Preto, São Paulo).

### 3.20. *Western blotting*

#### Western colorimétrico

O processo de *western blotting* se inicia com a realização de uma SDS-PAGE 12% conforme protocolo já mencionado, sendo aplicado 10-15 $\mu$ l de padrão de massa molecular pré-corado. Esse gel, assim como a membrana de transferência (nitrocelulose 0,45  $\mu$ m), foi incubado por 20 minutos em tampão de transferência (Tris 25 mM, glicina 192 mM, metanol 20% em pH 8,3).

A transferência das proteínas do gel de poliacrilamida para a membrana de nitrocelulose é feita por transferência semi-seca. Foi montado papel de filtro sobre gel de poliacrilamida sobre membrana de nitrocelulose sobre papel de filtro. A transferência se deu a 15 V por 15-30 min e foi utilizado o padrão de massa molecular pré-corado para indicar a passagem das proteínas.

Após a transferência, a membrana foi incubada por 1 h a temperatura ambiente com TBST (Tris 50 mM, NaCl 150 mM, Tween 0,1% em pH 8,0) + 5% de leite em pó. A membrana foi então lavada com TBST por 30 min com três trocas.

A membrana foi incubada com anticorpo anti-DPCAL1 (ou soro pré-imune) diluídos nas proporções 1:200 ou 1:600 por 2 h a temperatura ambiente sob agitação. A membrana foi então lavada com TBST por 30 min com três trocas.

A membrana foi incubada com anticorpo secundário, que apresenta atividade de peroxidase, diluído em TBST na proporção 1:1000 por 2h a temperatura ambiente sob agitação. A membrana foi então lavada com TBS (Tris 50 mM, NaCl 150 mM em pH 8,0) por 30 min com três trocas.

A revelação do *western blot* foi feita com a mistura de 20 mg de cloronaftol dissolvido em 4 ml de metanol, 20 ml de TBS aquecido a 37°C e 15 ml de H<sub>2</sub>O<sub>2</sub>.

### **Western fluorimétrico**

Devido à necessidade de aumentar a eficiência e sensibilidade da técnica, foram utilizados procedimentos de *western blotting* com uso de reagente fluorimétrico com marcação em filme fotográfico. As técnicas utilizadas na obtenção da membrana de nitrocelulose com as proteínas a serem marcadas são comuns aos dois métodos.

Após a transferência, a membrana foi incubada por 1 h a temperatura ambiente com TBST (Tris 50 mM, NaCl 150 mM, Tween 0,1% em pH 8,0) + 5% de leite em pó. A membrana foi então lavada com TBST por 30 min com três trocas. A membrana foi incubada com anticorpo (ou soro pré-imune) diluído em TBST + 2,5% de leite em pó nas proporções 1:1000 até 1:5000 por 2 h à temperatura ambiente sob agitação. A membrana foi então lavada com TBST por 30 min com três trocas. A membrana foi incubada com anticorpo secundário, que apresenta atividade de peroxidase, diluído em TBST + 2,5% de leite em pó na proporção 1:10000 por 2 h à temperatura ambiente sob agitação. A membrana foi então lavada com TBS (Tris 50 mM, NaCl 150 mM em pH 8,0) por 30 min com três trocas.

A revelação do *western blot* foi feita pela adição do substrato Luminata™ Forte (Western HRP Substrate) à membrana. A membrana foi então usada para sensibilizar o filme (CL-XPosure™ Film) que, por fim, foi revelado.

### **3.21. Ensaio enzimáticos**

As atividades de alfa-glicosidase e beta-glicosidase foram ensaiadas com os substratos p-nitrofenil-alfa-D-glicosídeo e p-nitrofenil-beta-D-glicosídeo, respectivamente. Vinte e cinco µl do substrato 10 mM em tampão citrato-fosfato 50 mM pH 5,0 foi adicionado a 25 µl da amostra (enzima). Após o tempo determinado, a reação de hidrólise foi interrompida pela adição de 200 µl de tampão carbonato-bicarbonato e SDS 0,1%. Após o ensaio era feita a leitura das absorvâncias a 420nm.

Ensaio de atividade de catepsina L foram realizados adicionando 100  $\mu$ l de tampão citrato-fosfato 50 mM pH 5,0 com EDTA 3 mM, cisteína 3 mM e o substrato Z-FR-MCA 10  $\mu$ M a 25  $\mu$ l de amostra (enzima). Após o tempo determinado, foram adicionados 25  $\mu$ l de ácido acético 30% ao ensaio para interrupção da reação. A leitura de fluorescência foi feita com excitação em 380 nm e emissão em 460 nm.

Em todos os casos, exceto quando mencionado, a determinação de atividade foi feita com quatro ensaios a 30°C com diferentes durações de tempo, de forma a determinar as velocidades iniciais de hidrólise.

### **3.22. Fixação e imunomarcação para microscopia de transmissão**

Para imunomarcação, fêmeas adultas de *Dysdercus peruvianus* foram dissecadas na própria hemolinfa para a remoção de V1 e V2. Os tecidos foram fixados em tampão fosfato 0,1 M, pH 7,4 com 4% de paraformaldeído e 0,3% glutaraldeído por 2 h a 4 °C. O tecido foi então lavado com tampão fosfato e, após ser desidratado em diferentes graus de etanol, foi embebido em resina acrílica (L.R. White). O bloco acrílico contendo o tecido foi cortado em ultramicrotomo e os cortes ultrafinos foram coletados sobre as telas de níquel cobertas por parlodio. Os cortes passaram, então, pelo processo de marcação. Duas lavagens com água bidestilada por 5 min. Bloqueio com Solução 2 (TBS + 1% BSA + 0,05% azida sódica pH 7,2) por 5 min. Incubação com NGS (soro normal de cabra) diluído 30 X em Solução 2 por 30 min a 4°C. Incubação com anticorpo primário diluído 1:200 em Solução 2 por aproximadamente 16 h a 4°C. Lavagem com Solução 1 (TBS + 0,2% BSA + 0,05% azida sódica + 0,1% Tween 20 pH 7,2) repetida quatro vezes sob agitação. Lavagem em Solução 3 (TBS + 1,0 BSA + 0,05% azida sódica pH 8,2) por 30 min. Incubação com anticorpo secundário diluído 15 X em Solução 3 por 1 h. Lavagem com Solução 1 repetida quatro vezes sob agitação. Lavagem com Solução 4 (TBS + 0,1% Tween 20) por 5 min. O anticorpo secundário anti-IgG de coelho produzido por cabra (Jackson Immune Research, EUA) é associado à partículas de ouro de 12 nm. A fixação foi feita com glutaraldeído 2,5% em tampão cacodilato de sódio 0,1 M pH 7,4 por 10 min em temperatura ambiente. Três lavagens com água bidestilada por 2 min sob agitação. Foi contrastado com acetato de uranila e citrato de chumbo, mantido em câmara com hidróxido de sódio para evitar a umidade. Por fim, o material foi

analisado em um microscópio eletrônico de transmissão (Zeiss EM 900) operado a 80 kV.

Para microscopia de transmissão, secções do intestino foram fixadas em 3% glutaraldeído em tampão cacodilato 0,1 M, pH 7,4 por 2 h a 4 °C. Após lavagem com 0,2 M de sacarose em tampão cacodilato 0,1 m (pH 7,4), os tecidos foram pós-fixados em 1% de tetróxido de ósmio no mesmo tampão por 1 h a 4 °C. O contraste “en-bloc” foi feito com 1% de acetato de uranila “overnight” a 4 °C. Após desidratação com etanol, o material foi embebido em resina de Spurr (Eletron Microscopy Sciences, USA). Cortes ultrafinos foram contrastados com citrato de chumbo e examinado em microscópio eletrônico de transmissão.

### **3.23. qPCR (PCR em tempo real)**

As reações foram realizadas com o reagente FastStart Essencial DNA Green Master (Roche). A concentração final do primer foi de 1,6 pmol e o volume de cDNA foi de 2 µl com diluições de 10 X ou 100 X de acordo com a gene em estudo. O volume final das reações foi de 15 µl. Foram realizados os controles utilizando reação sem cDNA, e sem cDNA e primer.

A reação e leitura foram feitas no aparelho Lightcycler Nano (Roche) nas seguintes condições: incubação da reação por dez minutos a 95°C; 45 ciclos de 15s a 95°C e 45s a 60°C. O produto de amplificação foi submetido a análises de fusão (ou “melting”). A curva resultante (curva de fusão) foi analisada como padrão de qualidade, sendo que reações com mais de um pico ou amplificação em controles sem cDNA tiveram seus primers refeitos até a solução destes efeitos (de inespecificidade e dimerização de primer, respectivamente). Os iniciadores foram listados na tabela suplementar 1.

### **3.24. Expressão diferencial por sequenciamento Illumina**

O RNA total foi extraído de três porções intestinais (V1, V2 e V3+4) e da carcaça (corpo total – intestino médio), ou seja, extraído pelo método do regente Trizol (Invitrogen) e tratando o RNA total resultante com DNase I (Invitrogen). Segue-se a isso, passos de quantificação final por RiboGreen (Invitrogen) e análise de integridade do RNA no Bioanalyzer (kit Agilent RNA 6000 Nano).

A confecção da biblioteca de cDNA e corrida no HiScanSq foi conduzida no serviço do Laboratório de Biotecnologia Animal da Escola Superior de Agricultura “Luiz de Queiroz” (Esalq/USP), utilizando procedimentos padrões estabelecidos pela Illumina no processamento das amostras (“sample prep”).

A extração de dados e análise bioinformática foi conduzida na mesma instituição de acordo com a descrição no material suplementar I.

Os genes diferencialmente expressos passaram por curagem manual. Nesta etapa, genes que expressam proteínas ribossomais foram excluídos, por não comporem o escopo da nossa pesquisa, assim como genes de plantas, que são contaminantes das amostras. Segundo a anotação do transcrito, foi feita a classificação de acordo com a função prevista para a possível proteína traduzida.

### **3.25. Preparação de amostras para proteômica**

Adultos de *D. peruvianus* foram mantidos por 10 dias sem alimento (sementes de algodão) e com acesso à água. Após esse período, os insetos foram alimentados em barra de sacarose e fibras de algodão por 24 h. Os animais foram dissecados em salina (215 mM NaCl) e V1 foi isolado e mantido submerso em salina por alguns minutos para eliminar contaminação da amostra por hemolinfa. A seguir, as frações intestinais V1 tiveram seus conteúdos drenados para uma solução salina de volume conhecido os epitélios correspondentes foram descartados. O material foi centrifugado a 25000 g a 4°C por 4 h e a fração solúvel foi utilizado como fonte de proteína.

A fração solúvel teve a salina trocada por água deionizada (tipo 1; MilliQ; Merck Millipore) em filtro de centrifuga Amicon com limite de exclusão de 10 KDa (Merck Millipore). Um volume contendo 50 µg de proteína foi seco em microtubos de centrifuga. O sedimento foi ressuspenso em 25 µL de ureia 8 M, aos quais foram adicionados 2,5 µL de bicarbonato de amônio 500 mM e em seguida 2,5 µL de DTT 120 mM, sendo então incubado por 1h. Ao final desse tempo, foram adicionados 2,5 µL de iodoacetamida 520 mM ao meio que foi incubado por 40 min no escuro. Seguiram-se a adição de 2,5 µL de DTT 14 mM, incubação por 30 min e nova adição agora de 215 µL de bicarbonato de amônio 50 mM. O material foi, então, tratado com 2 µg de tripsina com incubação noturna a 37°C. Finalmente, 5 µL de ácido

trifluoracético 50% (concentração final de aproximadamente 1%) foi reunido ao meio. A amostra final foi seca até um volume aproximado de 100  $\mu$ L e os peptídeos gerados foram purificados com ZipTip C18 (Merck Millipore) conforme instrução do fabricante.

A preparação final foi encaminhada para o Centro de Facilidades de Apoio à Pesquisa (CEFAP) do Instituto de Ciências Biomédicas da Universidade de São Paulo (ICB-USP), com o apoio do Dr. Giuseppe Palmisano para resolução dos peptídeos por espectrometria de massa na Plataforma nanoLC Easy-LTQ Orbitrap Velos-ETD.

## 4. Resultados

### 4.1. Expressão diferencial entre V1 e V2 – e a absorção de glicose

#### 4.1.1. Transportador de glicose

O estudo do transporte de glicose foi feito com a busca no banco de dados local por sequências que codifiquem transportadores de glicose (Sugar\_tr; pfam00083). A escolha desta família gênica, que inclui os GLUT, foi feita segundo o trabalho de Bifano *et al.* (2010), que indica um dos genes responsáveis pelo transporte de glicose pertencer a essa família. A escolha da família dos GLUT (facilitadores do transporte de glicose) é também apoiada pelo transporte de glicose (e água) ser independente de sais (Silva e Terra, 1994).

Com o transcriptoma do intestino médio completo, havia expectativa de encontrar a sequência do referido trabalho como majoritária (em número de *reads*) dentre outras poucas sequências da família dos transportadores de glicose. A tabela 2 demonstra que, na verdade, há pelo menos 26 sequências de proteínas relacionadas ao transporte de glicose (pfam00083). Dentre essas sequências, a que se destaca com maior número de *reads* é o contig\_834, não sendo a sequência ACY03364.1 (número de acesso) citada por Bifano *et al.* (2010). A sequência ACY03364.1 apresenta expressão (medida em *reads*) comparável à pelo menos outros quatro genes (contigs: 1763, 1842, 1999 e 2297).

**Tabela 2.** Contigs de genes que possivelmente codificam transportadores de glicose. A medida dos *reads* indica o nível de expressão de cada gene tecido. Sequências com mais de 80 *reads* foram verificadas se estavam completas.

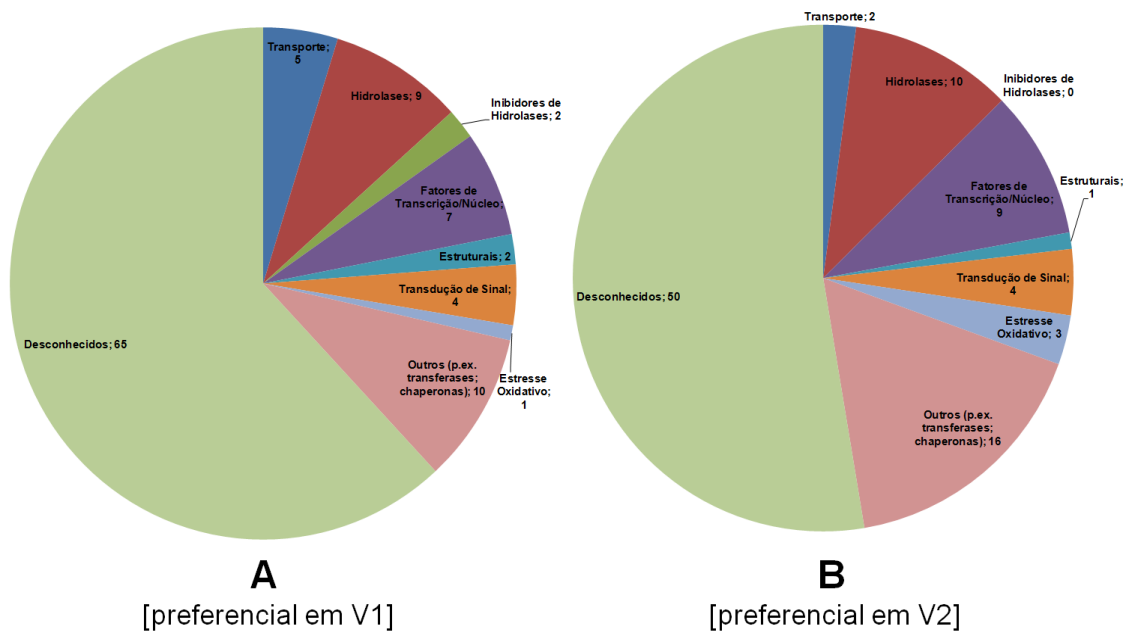
GenBank	Contigs (Intestino médio)	reads	completa
	Contig_834	274	sim
	Contig_1575	96	sim
	Contig_1621	98	sim
	Contig_1694	77	sim
	Contig_1755	89	sim
	Contig_1763	193	sim
	Contig_1842	171	sim
ACY03364.1	Contig_1905	158	sim
	Contig_1942	40	
	Contig_1959	41	
	Contig_1971	45	
	Contig_1999	141	sim
	Contig_2297	101	sim
	Contig_2382	26	
	Contig_2510	13	
	Contig_2643	29	
	Contig_2703	14	
	Contig_2943	11	
	Contig_3795	8	
	Contig_4143	9	
	Contig_4145	8	
	Contig_4202	10	
	Contig_4330	8	
	Contig_5021	8	
	Contig_5156	6	
	Contig_5458	24	

#### 4.1.2. Expressão diferencial de genes em V1 e V2

Como forma de análise complementar às já realizadas e com o objetivo de identificar os transportadores responsáveis pela absorção de nutrientes em *D. peruvianus*, foi realizado transcriptoma na plataforma Illumina com comparação da expressão gênica entre as frações intestinais V1 e V2. Os genes mais expressos em cada porção intestinal foram anotados e categorizados (figura 4). A maioria dos genes diferencialmente expressos tem função desconhecida e não é possível, ainda, atribuir qualquer anotação a eles. Os genes relacionados ao transporte que são diferencialmente expressos na comparação entre V1 e V2 serão tratados com mais detalhes abaixo, visto ser o foco principal desta abordagem. V1 expressa preferencialmente hidrolases do tipo aspártico proteinases. Já V2 expressa preferencialmente hidrolases do tipo cisteíno proteínases e hidrolases de lipídios. Devido a expressão diferencial entre essas porções intestinais, há uma fração razoável de fatores de transcrição e outros componentes nucleares diferencialmente



expressos, assim com genes que transcrevem proteínas relacionada à transdução de sinal. A categoria *Outros* tem composição heterogênea e levemente distinta entre os tecidos, sendo que V1 tem genes relacionados ao processamento de hormônios e V2 tem genes de chaperonas.



**Figura 4.** Análise dos genes diferencialmente expressos entre as porções V1 e V2 do intestino médio de *Dysdercus peruvianus*. **A.** Categorias dos genes diferencialmente mais expressos em V1 quando comparados à expressão de genes em V2. **B.** Categorias dos genes diferencialmente mais expressos em V2 quando comparada à expressão de genes em V1. As categorias estão identificadas e o valor ao lado corresponde ao valor absoluto de genes preferencialmente expressos daquela categoria.

Como base para as análises foi necessário considerar informações fisiológicas prévias da absorção diferencial entre porções do intestino médio (Silva e Terra, 1994; Bifano et al., 2010). Segundo a figura 4, há transportadores diferencialmente expressos quando comparamos os dois compartimentos (V1 e V2) entre si. Dentre esses transportadores, há três pertencentes à família dos facilitadores de transporte de glicose (ou transportadores do tipo GLUT). Dois deles são mais expressos em V1, sendo que um deles já foi relatado (contig\_834, Tabela 2) como sendo o de maior número de reads (que é uma forma de medida de expressão). O segundo transportador de glicose ainda não havia sido verificado e sua presença neste novo experimento está muito provavelmente associada à maior cobertura dada na plataforma Illumina quando comparada com a plataforma 454 (Roche) usada até então. Foi encontrada a expressão diferencial para um outro transportador de glicose em V2 (tabela 3).

**Tabela 3.** Expressão diferencial dos genes que codificam transportadores de glicose. A medida dos *reads* indica o nível de expressão de cada gene. Sequências com menos de 50 *reads* foram excluídas.

GenBank	Contigs (Intestino médio)	Reads	Completa	Preferencial em
	Contig_834	274	sim	V1
	Contig_1575	96	sim	
	Contig_1621	98	sim	
	Contig_1694	77	sim	V2
	Contig_1755	89	sim	
	Contig_1763	193	sim	
	Contig_1842	171	sim	
ACY03364.1	Contig_1905	158	sim	
	Contig_1999	141	sim	
	Contig_2297	101	sim	

## 4.2. Secreção de proteinases

### 4.2.1. Catepsinas D de *Dysdercus peruvianus*

Segundo os dados apresentados na Tabela 4 há pelo menos dez catepsinas D expressas no intestino, sendo que delas apenas uma é expressa também em outros tecidos. A DpCatD1 é a única proteína com expressão considerável nas duas amostras intestinais, uma do intestino médio completo e outra contendo apenas V2, V3 e V4. DpCatD5-7 apresentam grande expressão quando a amostra contém V1, mas não tem (ou tem muito pouca) expressão na amostra sem V1.

**Tabela 4.** Expressão dos RNAs que codificam as diferentes catepsinas D avaliadas segundo o número de *reads* obtidos no pirosequenciamento de diferentes tecidos. A primeira coluna indica a previsão do peptídeo sinal (prevista pelo SignalP). A segunda coluna indica o número de N-glicosilações (previstas por NetNGlyc 1.0).

	Pep. Sinal	N-glicosilação	Reads		
			Intestino Médio	V2/V3/V4	Carcça
DPCATD_1	sim	3	151	241	-
DPCATD_2	sim	6	41	67	-
DPCATD_3	sim	0	51	63	-
DPCATD_4	sim	4	84	-	-
DPCATD_5	sim	1	541	-	-
DPCATD_6	sim	1	142	36	-
DPCATD_7	sim	2	674	21	-
DPCATD_8	sim	6	39	50	-
DPCATD_9	?	0(?)	4	6	-
DPCATD_10	sim	2(1*)	3	4	65

?, dado ausente ou impreciso devido o falta da região n-terminal. \*, uma prolina após a asparagina torna improvável a glicosilação prevista.

Isso sugere que a expressão de genes como DpCatD5-7 seria predominantemente em V1 (contradizendo a baixa atividade proteolítica neste tecido).

```

CAT_D_1 1 -----MRYQFFIYVFLFG-----SMEALERTIHWKRPVP-----KSQEILKNIKF 42
CAT_D_2 1 -----MVLKVFQFVLLAS-----CAHTDAFIOVQKSLDD-----SWMILKEAVAKA 46
CAT_D_3 1 -----MRLIWMIEILYFIKLIIS-----LLITGSMFFFRVNLHKEAH-----SLDPNKLKLLHD 49
CAT_D_4 1 -----MGYKIFLFLCMV-----VFFVNGEIRVPLKREA-----KRD 32
CAT_D_5 1 -----MKTAVLVVALFA-----AFYEASARIALHYSNGFRSLRDFRNNSVSYKNN 48
CAT_D_6 1 -----GGVAKCLLLAFA-----LCSDALHRVQKRFPR--PELKKFSQHI TDFRRT 47
CAT_D_7 1 -----VALRNCLLVLSVV-----LVGCEALFRVPLKRFKKT-PEIKDFS NHLTFNKVQ 47
CAT_D_8 1 -----MFCKSLFVLSVLAS-----VQVYAQVTVQLYRNTIS-----DLIQNAHSH 40
CAT_D_10 1 MLNLKSETFLNSCLVLFCLISYIEHVS AENKLI RVPLETKS---AHQTFRDVDTEVRHV 57
*
CAT_D_1 43 AKSNNTNQOQRVEIDIFNFONVOYGT VYVCSGKSPFOVLEDTGSDI FWFPEKDVHG--T 100
CAT_D_2 47 AHRFNSSLSTHLTPRFEEFKYLQHYTKIGICDPEQTFDFTFNPIYRMVLVPS--VAHKK 104
CAT_D_3 50 LINNFNKKSVGT--KISSYLDGNYYG EISIGSPPOTFKVFDFTS SSDLWVPSKKCSKFN 107
CAT_D_4 33 TNPAPKTGRVN---LKNLRNEEYWG AISIGTPPOHFFVDFDTGSHLLWVPSHRCI--- 84
CAT_D_5 49 LHRYFLLKLGIPKILKNDENMDY G EISIGTPPOHFFVDFDTGSSDLWVPSOSCSLTS- 107
CAT_D_6 48 KPQALSTGAI---HLSNYQNAQY G EISIGTPPOHFFVDFDTGSSNLWVPSKKCSILNL 103
CAT_D_7 48 KANSILTGTV---SLSNYMNAQY G EISIGTPPOHFFVDFDTGSSNLWVPSOCSFFNL 103
CAT_D_8 41 PAKIDTTERVNG--LNFNLLAYWY YSTISVGSNKOKLNVSEFA TEDES IYIPTGGDK---- 94
CAT_D_10 58 TLKYNFNGPIPE--PLSNMMDAQY G DITIGTPPOHFFVDFDTGSSNLWVPSKKCSIVNI 115
*
CAT_D_1 101 CGVSHNKYDCSTSSCTLLKKDDEIN YLKGSTKYHLAKDILTITG DVSIKQMEVGLVYEG- 159
CAT_D_2 105 CGGTFKAYDEPSSHTYQDIHESFVS NEG-FKLNKAVTDFTHLCKLNVSDQSEGLITEG- 162
CAT_D_3 108 AOWTLKLYDSSKSATPYEEGSTS FSVS VIECETTFVSRDILVKIGSLELINOTFGEAVHQ 167
CAT_D_4 85 TDONKMLFYEEKSATYVNLNRSVD INYI GYMNGFAGRDTLSVGGITVONQCFAAALEG- 143
CAT_D_5 108 TCFKHNTFKGPLSSTYVPNSRRLS IEYGTGSLVGNVAMDTLITG DVPVNVOTFIBAIYIS 167
CAT_D_6 104 ACWTHHTYNSKKSSTYVRNDTKMSV OY GSGSFEGLSTDTLEVGDLRVEHOTFABAVQEP 163
CAT_D_7 104 ACWTHRTYKKNKKSSTYTPNGTEI SLCYVSGSLTEFLSTDTLELGLTITNOTFVEAVKPE 163
CAT_D_8 94 ---NVTTFNPSESSSFIRTTTKI PNVHCNVSMYG---NDIVQIGNLSAASQELALYEVG- 147
CAT_D_10 116 AGLIHRKYNSQKSSTYKQDGRKFAI OY GSGSLSGFLSODTLSMAGVDIKKQIFABAVNEP 175
*
CAT_D_1 159 --DCEGSQFDGILGMSLQ---MREDYQPMFRMIDQHIVSHKVFSERLN-NKDTGMSGGE 213
CAT_D_2 162 --PCRKSVYDCVIGLGFSS----KNHLTLLDNMSRLGTINSTAGLYLDKDVRTT-IGGE 214
CAT_D_3 168 SKTFYSTKLSGIVGLAFP--YSTNTGV PPIRRIINEMLEKNI FSVVLNRFNGNSE-SGGE 224
CAT_D_4 143 --AFQNLFDGIMCSFTEENS YDNLTTVIDNMVKQNLERQILFSFHMNRNGTSD-AGGE 200
CAT_D_5 168 EYFPERVKADGILGLGFP-ELASTL ATPPLYSMVTEDLLPKPIVSEYLNRPDSAE-RGGE 225
CAT_D_6 164 GVSFAFGKEDGILGGLGYS-SIAVLG VATPVKNMARONLA-QPIFSFELSRDETNESEGG 221
CAT_D_7 164 GLAEVFGKEDGILGLIAYP-AISV NNVLPVQNMALONLA-KPIESFELNRRNADDE-RGGE 220
CAT_D_8 147 ---TCMNLNLSVGLGFT---VENQL PTFVDNLFLQGLKQVVGLELNRDWSH-FAGE 200
CAT_D_10 176 GMTFVAAKEDGILGLGYD-TIAVDG IRPPFYLMVDQKLVSEPVFSFYLNRNPSAS-PGGE 233
*
CAT_D_1 214 LFIQGANP-VYSEEDFKYIEVNSEHTGFWSVAMNRVNFV-SAVIAGLGHAFEDTCTSLII 271
CAT_D_2 215 VALGDWNRKLVKNHDLIDFIPIVNTQI PAWQIHPMSFGETQISFDRYALDDTTSYFID 274
CAT_D_3 225 IIFGGWDDTKFNASSIKYIPELMDNTEWIFEDLDFISAGDG-EELCIK-COAVASTGTPYII 282
CAT_D_4 201 LFLGQNNEDFDPNLTLYLPELTD SHFVKKVQSIHMDDG-DWCKN-CLADTADTGTMTIT 258
CAT_D_5 226 IMFGETDESLEYKETAVPIEITERGYWQFSIDSVSIGGNTVLCCEDDCBAIADTGTSVVA 285
CAT_D_6 222 LIFGGDPELVDESTLHYVWVSKKGYWQFNMGGISLGGQ-SVACKGCCOAIADTGTSLIV 280
CAT_D_7 221 LILGGDDHNLQEETLNYIEVNKKAYWQFTMDSVSDKI-ESACSGTCOAVADTGTSLII 279
CAT_D_8 201 ISICHDDTRVDDNIKYIKATPVVPY WSVTLNIVKFNVD---LIATNVEAVLSSYLAFIL 257
CAT_D_10 234 IIFGGSDPEKYLGN-FTVWVSKKGYWQFSMDGLEINN--KKLCKGCCOAIADTGTSLII 290
*
CAT_D_1 272 TSYSVANKIAIQINAAIISQD--IMFV ECSRMDVPSLVESTHCTDYELTADPELLFEMEDG 329
CAT_D_2 275 VPAYEAEIHAKEIQAWNYDG--NLI VDCDRTGILPNI TLVINQRNYTLQAKDYVQHVNNY 332
CAT_D_3 283 GHTKLVDELYEKI CAKKHSN--MATVDCRETEKLSVTSINGNKYSIKGKDYVILKKTKE 340
CAT_D_4 259 GFINQLERIFKKFOAEDLDNNGYFVKCDKLEPELDIKFVINGRDYVLRSDYTFONSTI 318
CAT_D_5 286 VPAGKLDPLKSGAISDST--LGLTDCEKRLS LPPVTEKINGLN YTLSEKDYVIQTKED 343
CAT_D_6 281 GQDQVQAVDSDA IYSVS-GITMDCAKVDSLPEL TENINGRDY TLEGKDYVILKFSSE 339
CAT_D_7 280 GPTKVDAILSKIPGVSTFF-GISTVDC AVDSLPEIVFRINGDY TLEGKDYVILKMTAF 338
CAT_D_8 258 LPNDNFRTVSKLMKGSILVD-GFLLD KSSVDKLEPALEITVSENKTYLSPKDYVDTVHDF 316
CAT_D_10 291 GPTVEISAINSILGGTPIAG-GEYLVNCDLIPKLEPIDFINGEKETLKGKDYVILRVSAM 349
*
CAT_D_1 329 ----LCVLELGASSQLQDN--EVLG DVFLSRFYSVHDYQNRRIALIRINGTSPTAEPHS 383
CAT_D_2 333 KGS--CIVLIR----STKDKVWTLG AFLCNFYTFHYETKKLAFABLLPKNQGNIAQI 385
CAT_D_3 341 LEFWEQVVGIVGMPLYSQ--EVLG DFLQNYTFHDIENYQIAFNDLQI----- 389
CAT_D_4 319 PG--YQMLGMS---PNED-TEVVLGNVFLVKYYSVFD MENKRICEGTLK----- 361
CAT_D_5 344 GKT-FCFVCFIEN--ADSS---EILG DVFLLKGYTVDFDIGSQTVTFARKKEQK----- 390
CAT_D_6 340 LST-SCALCQAMELPHIQ---WILG DVFLLKGYTFDFGVNDQVGFALK----- 385
CAT_D_7 339 FSS-ACVGLTGMPLMIQ---WILG DVFLLKGYVSYDAENNOVGLAKL----- 384
CAT_D_8 317 NGLNLLIGSTE---DSK---WOLG SFLGKYV MELDYQNRRIE GSELKSSGGPSSRSL 369
CAT_D_10 350 GKT-VLCSFLGLDIPAPHGPI WILG DVFLLKGYTFDFLGDQDRVGFQSRDFYY----- 402
*
CAT_D_1 384 TTPPRNGSSTMSFSICSI FSLSEFMVLYLYIFNYR 418
CAT_D_2 386 TFWLISLSIFGFLQLIKW----- 404
CAT_D_3 389 ----- 389
CAT_D_4 361 ----- 361
CAT_D_5 390 ----- 390
CAT_D_6 385 ----- 385
CAT_D_7 384 ----- 384
CAT_D_8 370 FVPELLLITALINVM----- 383
CAT_D_10 402 ----- 402

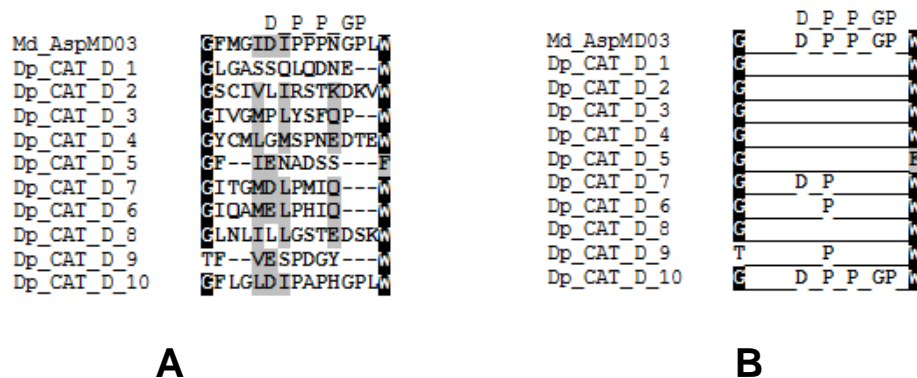
```

**Figura 5.** Alinhamento das sequencias de DPCATD (exceto a DPCATD9). Os resíduos de aminoácidos sublinhados correspondem ao peptídeo sinal. As colunas com asterisco acima correspondem à posição dos aminoácidos responsáveis pela catálise.

As catepsinas D, quando comparadas por alinhamento (figura 5), apresentam os resíduos responsáveis pela catálise formados por dois ácidos aspárticos conservados. Três sequencias fogem deste padrão. A sequencia DpCatD2, que não apresenta o primeiro ácido aspártico; DpCatD3, que não apresenta o segundo ácido aspártico e, por fim, DPCatD8, que não apresenta ambos os resíduos de aminoácido. A DpCatD9 por não estar completa, não pode ser analisada.

Há indícios, portanto, de que seis de um total de dez sequencias apresentam atividade proteolítica. A presença de ambos os ácidos aspárticos na região conservada (correspondente a resíduos do sítio catalítico) fazem com que DpCatD1, DpCatD4-7 e DpCatD10 sejam, provavelmente, enzimas com atividade proteolítica. Por outro lado, DpCatD2, DpCatD3 e DpCatD8 são proteínas que compartilham muitos resíduos com as proteinases, mas que não apresentam atividade proteolítica.

Como forma de discriminar a função lisossômica de uma catepsina D, é possível analisar a presença de uma região consenso (DxPxPx(G/A)P) que permanece conservada apenas nas catepsinas D que mantiveram sua ação intracelular (Padilha *et al.*, 2009). Das 10 sequencias de DpCatD apenas a DpCatD10 apresenta todos os cinco resíduos de aminoácidos da região consenso devidamente conservados (figura 6). As outras catepsinas D por não apresentarem a região consenso podem ser relacionadas à outra função, como a de digestão extracelular.

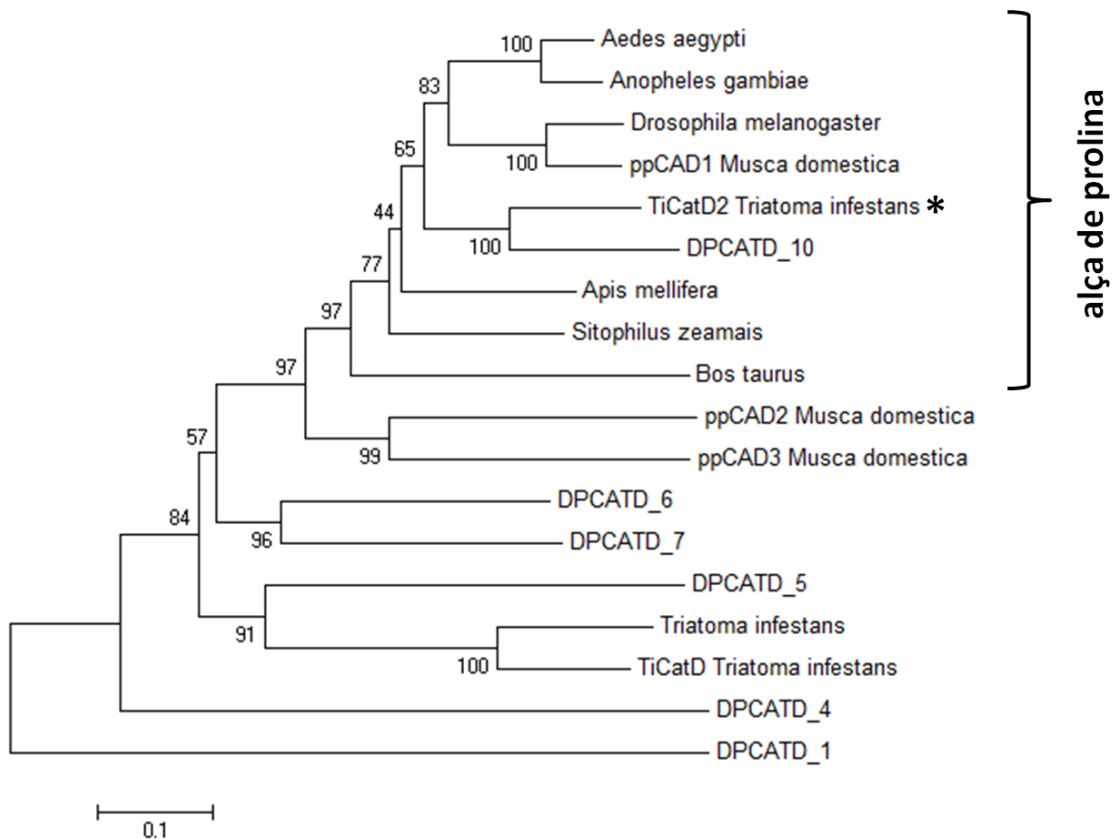


**A**

**B**

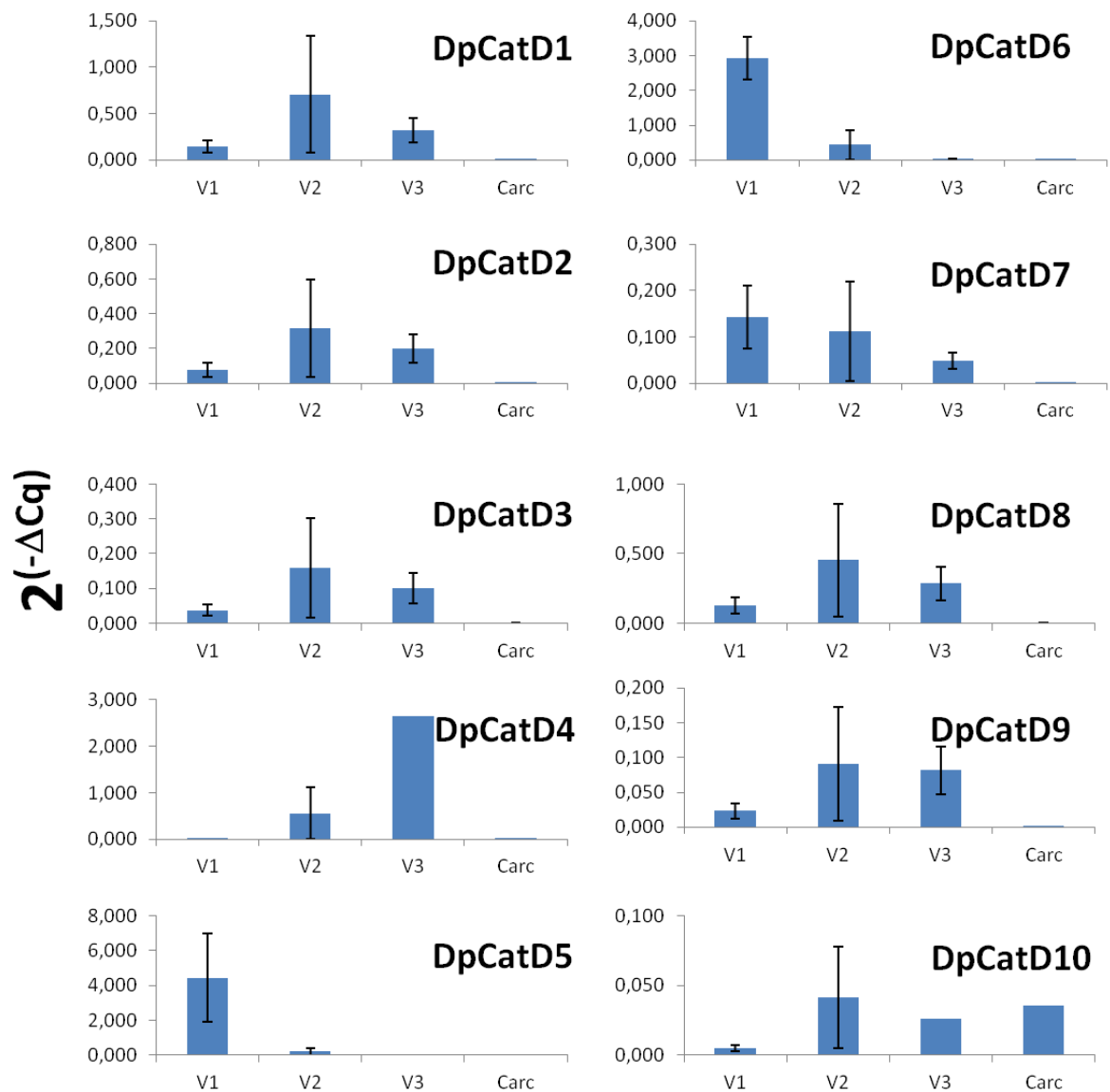
**Figura 6.** Comparação da posição da alça de prolina nas diferentes sequencias de catepsina D. AspMD03 foi usada como exemplo de catepsina D com a sequencia consenso completa. **A.** alinhamento de todos os resíduos de aminoácidos. **B.** Alinhamento mostrando apenas os resíduos que obedecem a sequencia consenso (além dos resíduos conservados que flanqueiam a região).

Em uma comparação de sequências a DpCatD10 forma um grupo monofilético com enzimas sabidamente lisossômicas, como a catepsina D do vertebrado *Bos taurus* (figura 7). Já as sequências que não apresentam a alça de prolina (DpCatD1-8) ficam em posições externas a esse grupo.



**Figura 7.** Cladograma de sequências de catepsina D obtido por Neighbor Joining. A chave indica as sequências que apresentam a alça de prolina característica das catepsinas D lisossômicas. \*, sequência com alça de prolina isolada no lúmen intestinal.

O estudo de expressão dos genes nos diferentes tecidos deixa claro a expressão de todas as catepsinas D no intestino médio, mostrando ainda a grande expressão da DpCatD5 e 6 em V1 (figura 8). Na mesma figura, a DpCatD10 mostra ser o gene com maior expressão em tecidos não intestinais, possivelmente reflexo da sua atuação lisossômica. Essa análise é complementar ao demonstrado na tabela 4 e responde a origem variação abrupta na expressão de genes como DpCatD5-7 entre os tecidos.



**Figura 8.** Expressão dos genes ( $2^{-\Delta Cq}$ ) de cathepsina D em diferentes tecidos de *Dysdercus peruvianus*. A expressão foi quantificada por qPCR e é mostrada como  $2^{-\Delta Cq}$  do gene de interesse tendo 18S como controle interno. V1, primeira porção do intestino médio; V2, segunda porção do intestino médio; V3, terceira e quarta porção do intestino médio; Carc, corpo total menos o intestino médio. +/- SEM. Barra sem desvio são a média de duas réplicas biológicas.

A importante expressão da DpCatD5 e 6 em V1 foi confirmado pelo proteoma do conteúdo de V1 (tabela 5). Foram identificadas também as cathepsinas DpCatD1, 4, 7 e 8 em um dos dois experimentos realizados.

**Tabela 5.** Catepsinas D identificadas no lúmen de V1 por técnica proteômica. Cada experimento corresponde a uma réplica biológica do conteúdo solúvel de V1.

	Cobertura total	NSAF* experimento 1 (%)	NSAF* experimento 2 (%)
DpCatD1	? %	0,041 (0,004)	-
DpCatD4	11 %	0,207 (0,023)	-
DpCatD5	52 %	1,436 (0,159)	0,434 (0,061)
DpCatD6	43 %	0,334 (0,037)	0,698 (0,098)
DpCatD7	32 %	0,286 (0,032)	-
DpCatD8	9 %	-	0,440 (0,062)

\*NSAF, fator de abundância normalizada de espectros (normalized spectral abundance factor)

Como os dados de expressão indicam ser V1 a principal porção do intestino na secreção de catepsina D, o tecido foi selecionado para estudos de modulação da transcrição pelo alimento. Cinco dos dez genes de catepsina D são modulados pela alimentação em V1 (tabela 6). E, como previsto, a DpCatD10 não teve sua expressão modulada pela alimentação neste mesmo tecido, indicando mais uma vez não ser responsiva às funções de digestão de alimentos no intestino.

**Tabela 6.** Modulação da expressão de catepsinas D em V1 de *Dysdercus peruvianus*. A expressão foi quantificada por qPCR e é mostrada como a média de  $2^{(-\Delta Cq)}$  do gene de interesse tendo 18S como controle interno. A comparação é feita entre animais em jejum com animais após um dia e meio (36 h) de alimentação. Em negrito  $p < 0,05$  do teste t.

	$2^{(-\Delta Cq)}$ médio		p (studentT)
	Alimentado (n4)	Jejum (n4)	
DpCatD1	0,0412	0,0205	<b>0,0036</b>
DpCatD2	0,0336	0,0212	0,0501
DpCatD3	0,0170	0,0111	<b>0,0269</b>
DpCatD4	0,0037	0,0026	0,2793
DpCatD5	2,2587	1,7976	0,0770
DpCatD6	8,6806	1,6943	<b>0,0001</b>
DpCatD7	0,0141	0,0118	0,3103
DpCatD8	0,0546	0,0159	<b>0,0001</b>
DpCatD9	0,0023	0,0015	<b>0,0147</b>
DPCatD10	0,0007	0,0005	0,1637

#### 4.2.2. Catepsinas L de *Dysdercus peruvianus*

Com as análises de bioinformática sobre os dados obtidos no pirosequenciamento do cDNA do intestino de *D. peruvianus* foi possível obter as





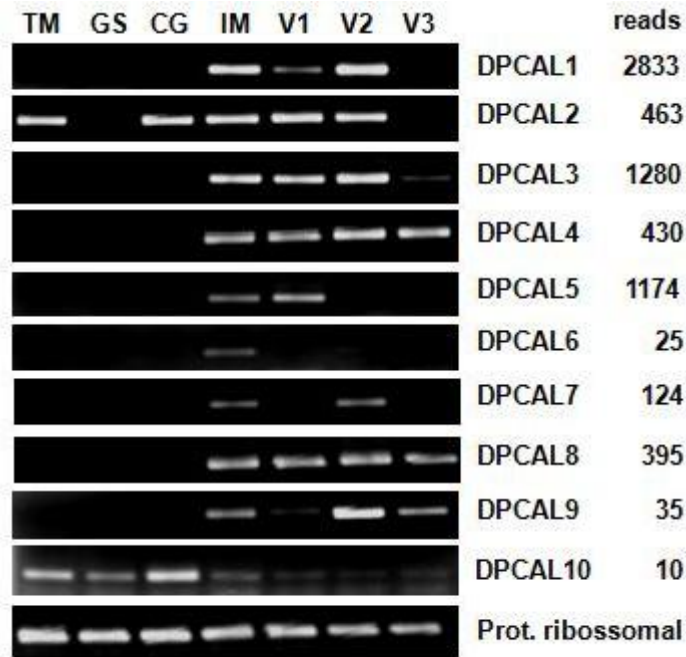
A DPCAL2 foi suprimida do alinhamento por conter uma longa sequência N-terminal que não tem homologia com as outras DPCALs, mas mantém conservados os resíduos catalíticos, bem como é possível prever um peptídeo sinal, como está sintetizado na tabela 7. Para confirmar que o N-terminal da DPCAL2 não é simplesmente um artefato da montagem do contig, esse trecho não homólogo as outras DPCALs foi isolado e usado no programa BlastP. Desta forma, é possível identificar o trecho N-terminal em cisteína proteinases de diversos animais, sendo aquele que mais se aproxima da DPCAL2 é a cisteína proteinase do também Heteroptera *Plautia stali* (BAF94153.1).

**Tabela 7.** Predição da presença de peptídeo sinal, sinais de N-glicosilação nas DPCALs e medidas entre lisinas obtidas a partir de modelagem computacional.

	Peptídeo sinal	N-glicosilação	Distância (Å)
DpCaL1_A1	Sim	1 (N20)*	35.76 (K16-K194); 36.52 (K17-K194); 38.93 (K33-K79); 40.12 (K17-K137)
DpCaL3_1	Sim	1 (N156)	33.36 (K57-K181); 35.64 (K57-K138); 38.15 (K45-K210)
DpCaL4	Sim	1 (N270)	32.58 (K104-K194); 39.43 (K57-K194)
DpCaL5_1	Sim	1 (N76)	33.63 (K12-K77); 35.10 (K45-K85)
DpCaL6	Sim	3 (N44; N191; N265)	32.50 (K101-K184); 36.89 (K184-K293); 36.96 (K184-K292)
DpCaL7	Sim	1 (N137)	33.97 (K39-K216); 34.44 (K38-K213); 34.53 (K60-K196); 34.92 (K15-K192)
DpCaL8	Sim	1 (N70)	36.51 (K18-K104); 37.23 (K37-K104)
DpCaL9	Sim	1 (N132)	35.41 (K13-K187); 36.51 (K13-K100); 36.59 (K36-K100)
DpCaL10	Sim	1 (N79)	33.37 (K33-K80); 34.08 (K44-K80); 37.32 (K59-K145)

Todas as sequências foram submetidas à predição de hélice transmembrana e nenhuma apresentou essa característica na preproteína. Não houve predição de sítios de modificação por âncora de GPI em nenhuma das sequências. \*, uma prolina após a asparagina torna improvável a glicosilação prevista.

O estudo do alinhamento nos permite inferir que todas as DPCALs apresentam peptídeo sinal, tendo então condições de seguir a rota lisossômica ou o encaminhamento para via secretória. A previsão da N-glicosilação em todas as DPCALs (Tabela 7), por sua vez, indica que todas as proteínas consideradas teriam condições de seguir a rota lisossômica pela via do transportador de manose-6-fosfato, ao menos sob o aspecto do sítio de glicosilação (discutiremos outros aspectos abaixo). Os motivos típicos de catepsina L estão presentes nas sequências com algumas variações que já foram descritas para sequências de catepsina L de outros Hemiptera (Waniek, *et al.*, 2012).



**Figura 10.** Expressão do RNAm que codificam as diferentes catepsinas L foi avaliada por RT-PCR. Como forma de comparação relativa, todas as amostras de DPCALs são mostradas no ciclo 25, mas houve acompanhamento de mais e menos ciclos para certificar que as ampliações estavam na fase logarítmica. A quantidade de reads obtidas no pirosequenciamento está à direita e indica o quanto cada DPCAL é expressa em termos de RNAm.

A figura 10 dá base para se inferir a importância relativa dos genes que codificam as catepsinas L. Todos os genes de catepsinas L estudados são expressos no intestino médio do inseto, sendo que apenas DPCAL10 apresenta expressão em todos os tecidos. Exceto por DPCAL10 e DPCAL2, todos os genes têm expressão restrita ao epitélio intestinal, sendo preferencialmente expressos em V2. Com a análise da quantidade de reads obtidos no pirosequenciamento do cDNA do intestino do *D. peruvianus* é possível dizer que a sequência DPCAL1 é a mais expressa, sendo DPCAL3 e DPCAL5 ainda comparativamente muito expressas (figura 10).

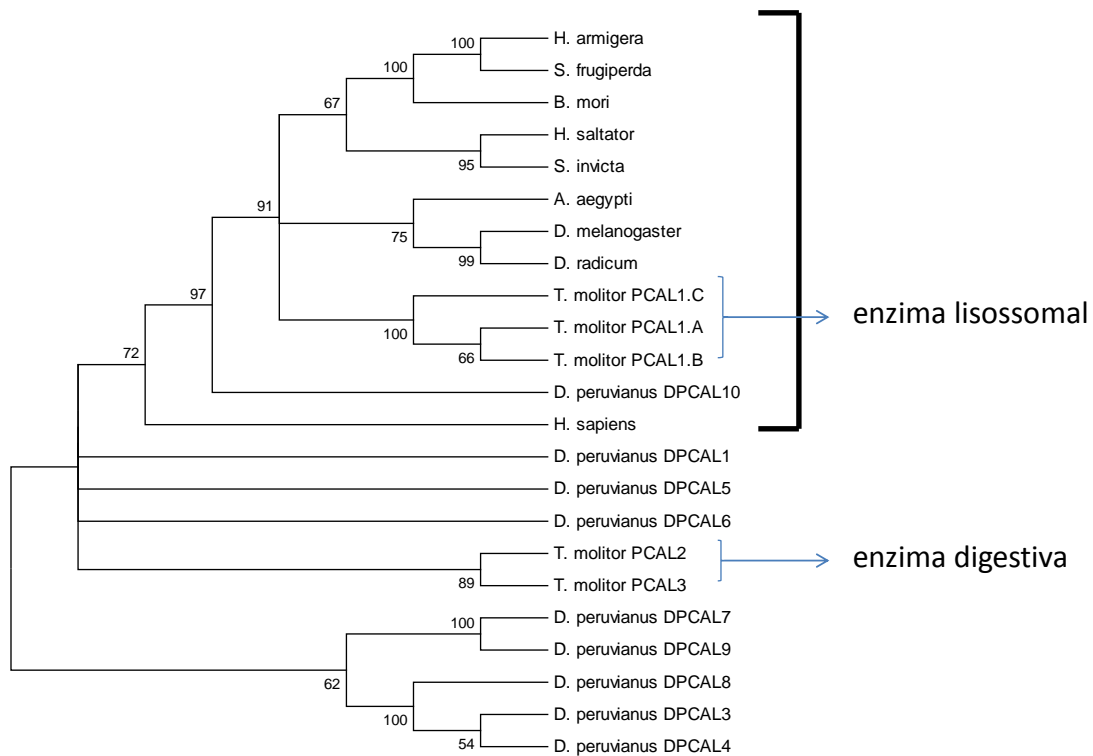
**Tabela 8.** Comparação entre as DPCALs com catepsinas L de inseto (*T. molitor*), sabidamente lisossômicas (PCAL1) e digestivas (PCAL2 e PCAL3) (Cristofolletti *et al.*, 2005; Beton, *et al.*, 2012).

	PCAL1		PCAL2		PCAL3	
	ident. (%)	simil. (%)	ident. (%)	simil. (%)	ident. (%)	simil. (%)
DPCAL1	58	73	44	59	50	67
DPCAL2	44	60	36	51	39	55
DPCAL3	55	69	40	58	47	64
DPCAL4	54	68	40	55	46	60
DPCAL5	54	67	44	59	47	61
DPCAL6	49	66	42	59	47	63
DPCAL7	47	64	41	57	47	63
DPCAL8	54	69	41	59	47	63
DPCAL9	44	58	37	50	42	59
DPCAL10	61	79	43	58	44	64

ident., corresponde a porcentagem de resíduos de aminoácidos idênticos entre as sequencias; simil., corresponde a porcentagem de aminoácidos similares, quanto a sua classificação, entre as sequencias.

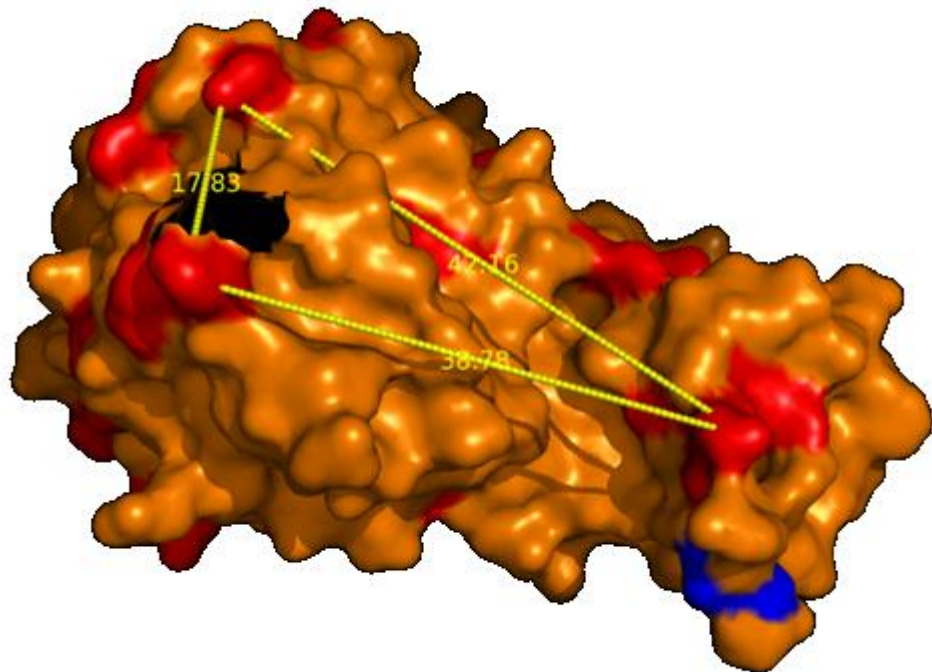
Relacionar as sequencias obtidas com sequencias de insetos que apresentem o mesmo processo de duplicação e divergência gênicas nos traria mais indícios sobre a função lisossômica ou digestiva das DPCALs. A PCAL1 foi imunocitocalizada em todos os tecidos de *T. molitor* dentro de vesículas semelhantes a lisossomos (Cristofolletti, *et al.*, 2005). Já a imunocitocalização da PCAL3 em tecidos de *T. molitor* marca apenas vesículas de secreção presentes em células do intestino médio deste inseto, sendo que PCAL2 seria também uma enzima secretada pelas células intestinais (Beton, *et al.*, 2012).

A simples comparação entre os aminoácidos conservados (idênticos ou similares) na tabela 8 já permite verificar que a DPCAL10 tem a maior identidade (e similaridade) com a enzima lisossômica de *T. molitor*, quando comparado às demais DPCALs. A mesma DPCAL10, quando comparada com as duas enzimas digestiva (PCAL2 e 3) de *T. molitor* mostra ter valores muito próximos de quando comparamos as demais DPCALs.



**Figura 11.** Cladograma de sequências de catepsina L obtido por Neighbor Joining. O grupo monofilético com bootstrap 72 inclui sequências de genes expressos apenas em lisossomos e foi demarcado com colchete.

A hipótese de que DPCAL 10 seria a enzima responsável pela digestão intracelular (ou lisossômica) é reforçada por um cladograma de sequências de catepsina L (figura 11). Nele é possível perceber que sequência de enzimas sabidamente lisossômicas, como a PCAL1 de *T. molitor* (Cristofolletti *et al.*, 2005) e a catepsina L de *H. sapiens* (Schröder, *et al.*, 2010), se agrupam junto com DPCAL10 em um grupo monofilético (demarcado pelo colchete). Por outro lado, existem enzimas que na análise ficam fora do grupo dito das enzimas lisossômicas, como as sabidamente digestivas PCAL2 e PCAL3 de *T. molitor* (Cristofolletti *et al.*, 2005; Beton, *et al.*, 2012). Finalmente, um grupo monofilético com bootstrap 62 é externo a todos as outras catepsinas L, incluindo as DPCAL 3 e 4 e DPCAL7-9, que deve refletir fenômenos de duplicação gênica e divergência dentro da espécie em estudo.



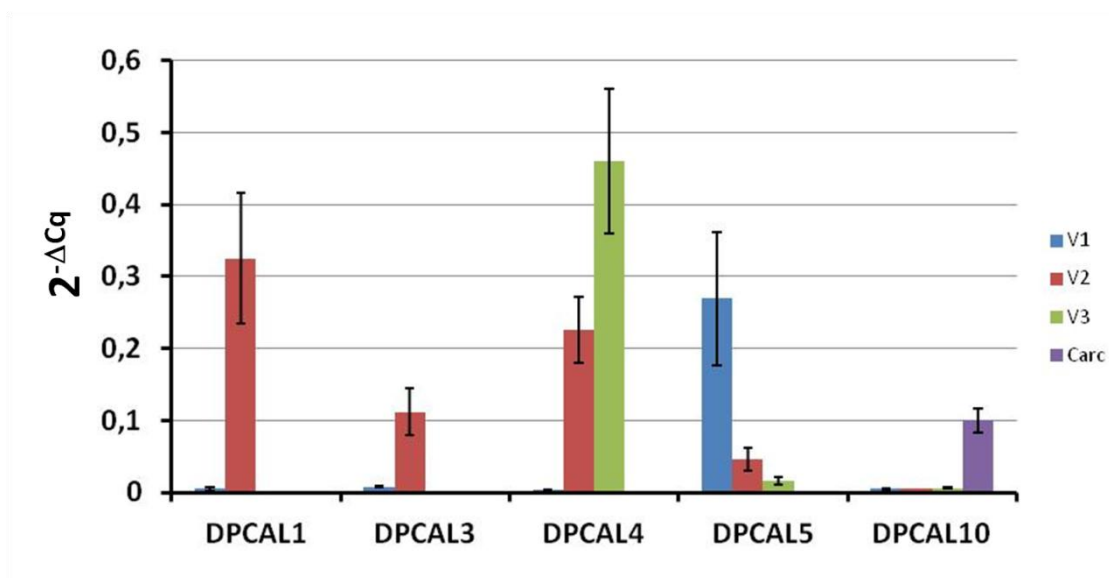
**Figura 12.** Modelo computacional por homologia da DPCAL1. As lisinas expostas estão marcadas em vermelho e o resíduo (de asparagina) que recebe a N-glicosilação está marcada em azul.

A presença do peptídeo sinal em todas as enzimas estudadas indica que elas são destinadas para a rota secretória (que inclui tanto enzimas secretadas como residentes intracelulares) (figura 9 e tabela 7).

A figura 12 e os valores da tabela 7 mostram as medidas entre lisinas selecionadas com base na sua topologia na superfície do modelo molecular. Há pelo menos uma medida favorável em cada modelo analisado. A posição do sítio de N-glicosilação e as medidas entre lisinas são muito próximas (se não idênticos) aos requeridos para o reconhecimento pela fosfotransferase no aparato de Golgi.

A expressão tecidual das catepsinas L, além de mostrar variação tecido-específica, indicava também variação do nível de expressão relativa entre os genes. Genes que codificam enzimas digestivas (DPCAL1; 3-9) indicavam maior transcrição quando comparados ao gene que codifica a catepsina L lisossômica (DPCAL10). Além disso, a catepsina L lisossômica mostrava indícios de repressão em tecidos digestivos (ver reads na figura 10 e figura 13).

Para verificar a expressão dos genes de maneira mais direta, a realização de experimentos de qPCR foram realizados (figura 13).



**Figura 13.** Expressão relativa das catepsinas L de *Dysdercus peruvianus* por qPCR. A expressão foi quantificada por qPCR e é mostrada como  $2^{(-\Delta Cq)}$  do gene de interesse tendo 18S como controle interno. V1, primeira porção do intestino médio; V2, segunda porção do intestino médio; V3, terceira e quarta porção do intestino médio; Carc, corpo total menos o intestino médio. +/- SEM.

Consideremos que o nível de expressão da DPCAL10 no tecido carcaça como a expressão basal de catepsina L de um tecido, ou seja, com a catepsina L atuando intracelularmente. A figura 13 mostra que a expressão da DPCAL10 é menor do que expressão de outras DPCALs intestinais (como DPCAL1 e 4 em V2). Isso indica que genes de catepsina L são mais expressos em tecidos intestinais. Considerando o conjunto dos oito genes que expressam catepsina L no intestino do *D. peruvianus*, há maior expressão global de catepsinas L (ou superexpressão).

Considerando agora a conservação da região carboxi-terminal (ver figura 9), que permite a entrada na rota secretória alternativa à dependente de manose fosfato, podemos dizer que DpCal1, 5, 7, 9 e 10 (e provavelmente DpCaL3 e 4) seriam capazes de seguir a rota secretória (Chauhan et al., 1998).

#### 4.2.3. Catepsinas L de *Dysdercus peruvianus* (estudo das selecionadas)

Os dados até então parecem indicar uma rota de secreção das catepsinas L que inclui a rota lisossômica. Experimentos de imunocitocalização trariam mais dados para essa análise, permitindo a comparação da localização celular das proteínas que seguem a rota secretória ou, neste caso, a rota de encaminhamento lisossômico. Para tal, é necessária a produção de anticorpos das catepsinas L.

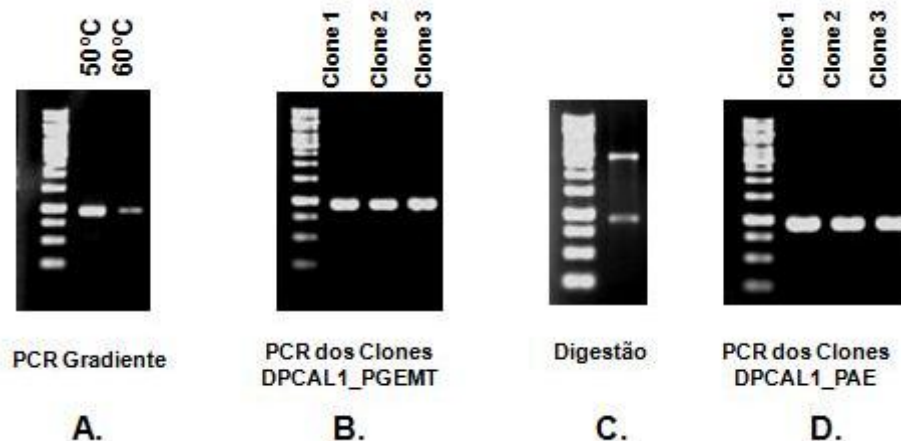


Foram selecionados para um estudo mais aprofundado as catepsinas L DPCAL1, 3 e 5 por seu provável papel no processo digestivo, DPCAL2, por seu intrigante padrão de expressão e DPCAL10, por ser o melhor candidato de catepsina L tipicamente lisossômica em *D. peruvianus*. Seus genes foram submetidos ao processo de clonagem para a expressão heteróloga da proteína recombinante, como será individualmente apresentado a seguir.

#### 4.2.3.1. DPCAL1

A clonagem e expressão heteróloga da DPCAL1 se mostra bastante promissora sob dois aspectos. O primeiro diz respeito à produção de uma proteína em quantidade suficiente para produção de anticorpos policlonais para a marcação do conteúdo das vesículas internas das células do epitélio intestinal. O segundo fator que torna seu estudo atraente é a possibilidade de conhecer a atividade enzimática da enzima muito expressa no intestino.

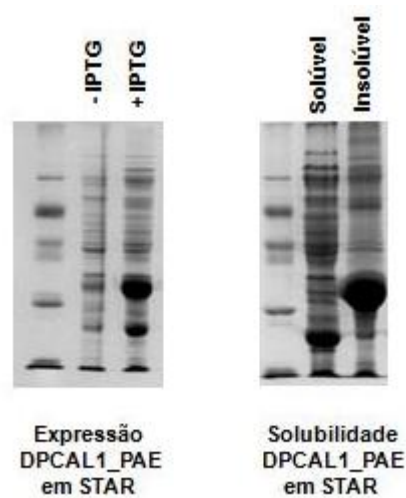
A figura 14 mostra de maneira resumida o processo de clonagem bem-sucedido desta proteína. O processo inclui a subclonagem no vetor PGEMT e a clonagem final no vetor de expressão PAE.



**Figura 14.** Eletroforeses (em gel de agarose) dos produtos das reações de amplificação. Processo de clonagem DPCAL1. A. PCR realizado sobre cDNA do intestino, foram utilizadas duas temperaturas de pareamento, sendo 50°C considerado padrão para os próximos experimentos. B. PCR de três clones de DPCAL1\_PGEMT. C. Digestão da construção DPCAL1\_PGEMT Clone 1 com as enzimas de restrição BamHI e PstI. D. PCR de três clones de DPCAL1\_PAE.

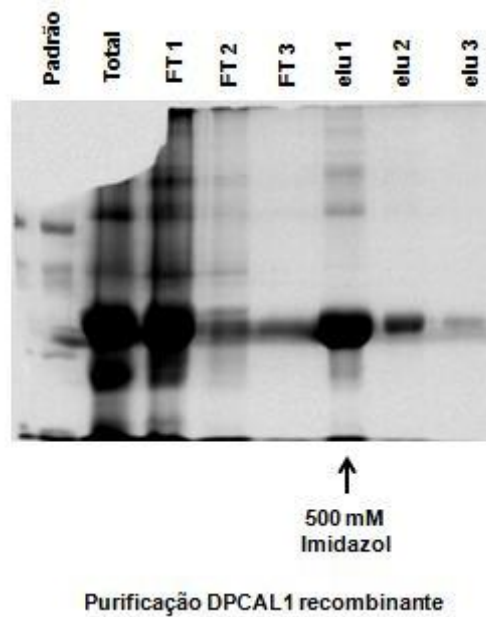
A construção DPCAL1\_PAE foi inserida em uma bactéria *Escherichia coli* da cepa BL21\_STAR e após adição do indutor IPTG a expressão foi verificada (figura 15).





**Figura 15.** Eletroforese (SDS-PAGE) das proteínas expressas em *E. coli* durante o processo produção da DPCAL1 recombinante. O primeiro gel apresenta na segunda raia uma amostra das bactérias antes da adição de IPTG no meio, e a terceira raia apresenta outra amostra do mesmo conjunto de bactérias horas após a adição de IPTG. O segundo gel apresenta o produto da lise das bactérias induzidas (ou após a adição de IPTG) após centrifugação, sendo a segunda raia a fração solúvel e a terceira raia a fração insolúvel.

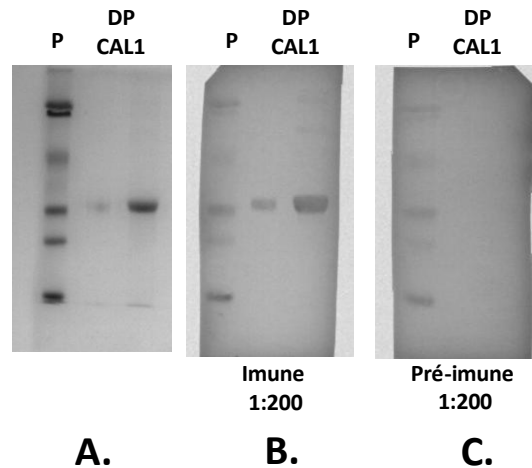
Com a expressão realizada nas condições descritas em Materiais e Métodos foi possível conseguir apenas a proteína recombinante na forma insolúvel. Com isso foi utilizado um procedimento de purificação que envolvia o uso da ureia como agente capaz de solubilizar a proteína recombinante e assim realizar a purificação por afinidade. O resultado desse procedimento está representado na figura 16. Nela é possível ver que a proteína recombinante está bastante enriquecida na fração chamada de “elu1”.



**Figura 16.** Eletroforese (SDS-PAGE) de proteínas dos diferentes passos do processo de purificação da DPCAL1 recombinante. O processo de purificação foi realizado em resina de afinidade. O lisado celular bruto solubilizado em ureia 8M (raia “Total”) é adicionado à resina que é então lavada (raias “FT”), sendo que a adição de imidazol na concentração 500mM libera a proteína recombinante ligada à resina (frações “elu”).

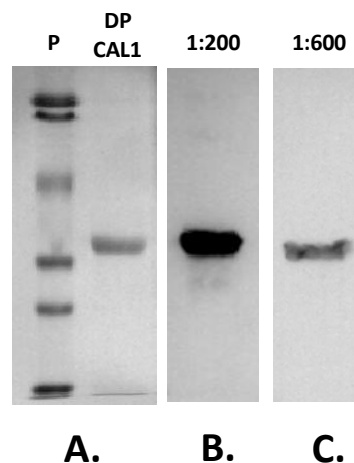
A proteína recombinante purificada foi, então, aplicada em coelho para a produção de anticorpos policlonais anti-DPCAL1.

A proteína recombinante DPCAL1 produziu uma resposta imune em coelho que, por sua vez, produziu anticorpos capazes de reconhecer a proteína recombinante. A DPCAL1 semi-purificada pode ser vista em duas concentrações distintas no gel (SDS-PAGE) corado com Coomassie (figura 17 A). Essas proteínas, assim como o padrão de massa molecular, foram transferidas para uma membrana de nitrocelulose e por *western blotting* se verificou a marcação da proteína recombinante quando aplicado o primeiro soro imune (figura 17 B), mas essa marcação não foi vista quando aplicado o soro pré-imune (figura 17 C). Exceto pelo soro utilizado, as condições do experimento foram idênticas, até mesmo quanto à diluição dos soros. Sendo assim, consideramos que o anticorpo anti-DPCAL1 foi capaz de reconhecer a proteína recombinante.



**Figura 17.** Eletroforese (SDS-PAGE) e *western blot* da DPCAL1 recombinante. **A.** Eletroforese da DPCAL1 recombinante semi-purificada (segunda e terceira raia), corada por Coomassie Blue. **B.** *Western blot* com o primeiro soro imune anti-DPCAL1 com título de 1:200. **C.** *Western blot* com soro pré-imune com título de 1:200.

A partir de uma nova inoculação da proteína recombinante no coelho, foi obtido um novo soro imune, que tende a ter maior título que o primeiro. Para testar a eficiência deste soro imune, se realizou um experimento no qual a proteína recombinante (figura 18 A) foi reconhecida com o soro sob duas diluições 1:200 (figura 18 B) e 1:600 (figura 18 C.) Nos dois casos houve reconhecimento da proteína recombinante DPCAL1.



**Figura 18.** Eletroforese (SDS-PAGE) e *western blot* da DPCAL1 recombinante – teste de diluição do anticorpo. **A.** Eletroforese da DPCAL1 recombinante semi-purificada, corada por Coomassie Blue. **B.** *Western blot* com anticorpos policlonais anti-DPCAL1 com título de 1:200. **C.** *Western blot* com anticorpos policlonais anti-DPCAL1 com título de 1:600.

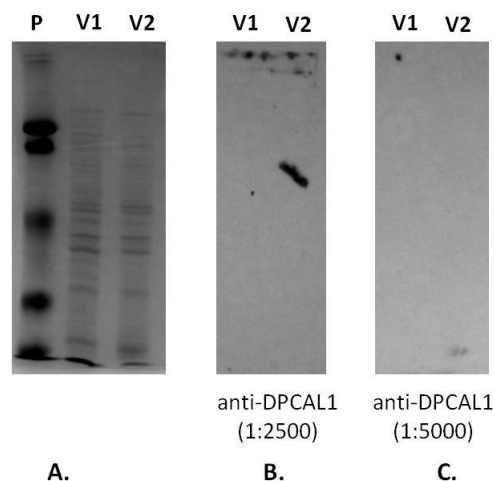
O passo lógico seguinte seria realizar experimento que identificasse a DPCAL1 no conjunto de proteína extraídas de diferentes tecidos do *D. peruvianus*.

Para isso foram feitas diversas extrações de proteínas dos ventrículos intestinais e, depois de quantificadas, as proteínas seguiram o mesmo procedimento feito com a proteína recombinante no *western blotting*. Contudo, resultados negativos se seguiam e com isso não foi possível identificar a DPCAL1 entre as proteínas intestinais.

Buscando a causa dessa falha foi feita uma análise mais cuidadosa dos controles e foram estabelecidos alguns métodos para controlar os possíveis problemas, conforme detalhado no material suplementar II.

O anticorpo anti-DPCAL1, possível marcador de enzima secretada, não é capaz de reconhecer especificamente uma proteína dentre as proteínas solúveis de V1 e V2 (figura 19). Quando o título do anticorpo é aumentado o reconhecimento passa a ser inespecífico, reconhecendo até proteínas padrão de massa molecular.

Esse experimento contou com uma dificuldade adicional. A presença de proteinases na amostra é capaz de degradar as proteínas mesmo em condições rigorosas de manutenção das amostras em gelo e homogeneização e fervura imediatas. O material suplementar II trata especificamente da solução deste problema.

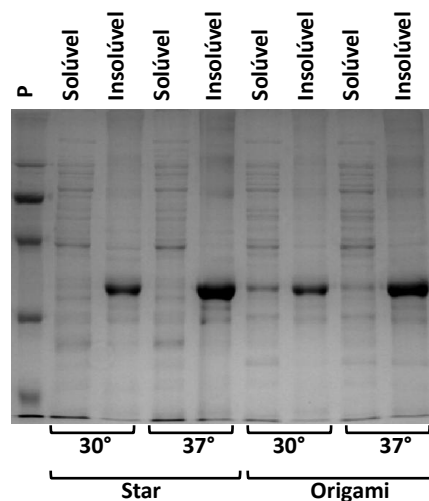


**Figura 19.** Reconhecimento da DPCAL1 de *Dysdercus peruvianus* por *western blotting*. **A.** SDS-PAGE das proteínas solúveis de V1 e V2 coradas por Coomassie Blue. **B.** *Western blot* com anticorpos policlonais anti-DPCAL1 com título de 1:2500. A forte marcação em V2 pode ser considerada artefato quando comparada ao padrão de bandas do SDS-PAGE. **C.** *Western blot* com anticorpos policlonais anti-DPCAL1 com título de 1:5000. (Marcação em B. raia V2 é artefato)

Esses resultados indicam que o anticorpo anti-DPCAL1 não apresenta título e especificidade adequados para os experimentos subsequentes de imunomarcação

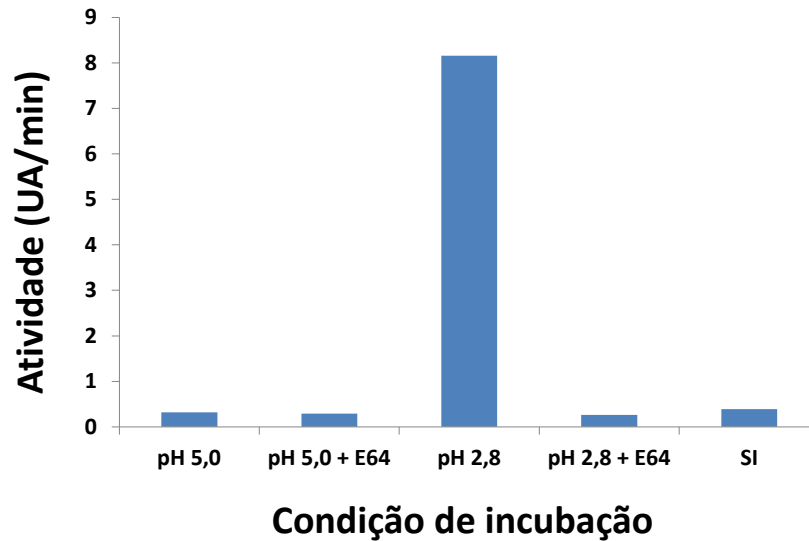
histológica. Este resultado e o resultado dos outros anticorpos testados estão resumidos na tabela 10 na seção que trata dos marcadores celulares.

A produção da DPCAL1 recombinante foi feita na forma insolúvel e, portanto, sem possibilidade de verificar sua atividade proteolítica. Na tentativa de se obter a proteína da forma solúvel o plasmídeo de expressão DPCAL1\_PAE foi inserido em duas cepas de *E. coli* (STAR e Origami) e sob duas temperaturas após a adição de IPTG (30° e 37°C). Com isso seria possível verificar se em alguma condição a proteína recombinante seria produzida na forma solúvel.



**Figura 20.** Eletroforese (SDS-PAGE) das proteínas expressas em *E. coli* durante o processo de expressão da DPCAL1 recombinante. O gel mostra um experimento de teste de solubilidade. Nele está o produto solúvel e insolúvel da lise de bactérias de duas cepas (Star e Origami) após a produção da DPCAL1 sob duas temperaturas distintas (30° e 37°).

A cepa STAR não foi capaz de produzir a proteína recombinante na forma solúvel em nenhuma das condições (figura 20). Já a cepa Origami produziu a proteína parcialmente solúvel em ambas as temperaturas, mas mais marcadamente a 30°C. Apesar de grande parte da proteína ainda estar presente na porção insolúvel, a fração de proteína recombinante presente na forma solúvel foi suficiente para testes preliminares de ativação e atividade proteolítica.



**Figura 21.** Atividade sobre Z-FR-MCA do extrato solúvel bruto da *E. coli* Origami expressando DPCAL1 recombinante a 30°C (vide figura 20). O extrato solúvel foi submetido a 30 minutos de incubação em diferentes pHs (tampão citrato-fosfato pH 5,0 e tampão acetato pH 2,8), na presença ou ausência de inibidor (E-64 5 $\mu$ M) e, após a incubação, teve a atividade enzimática mensurada em pH 5,0. SI (sem incubação) permaneceu 30 minutos em gelo antes de ter a atividade enzimática mensurada. A atividade foi medida por ensaios com seis variações de tempo e foi aqui representada por unidades arbitrárias (UA) por minuto (min).

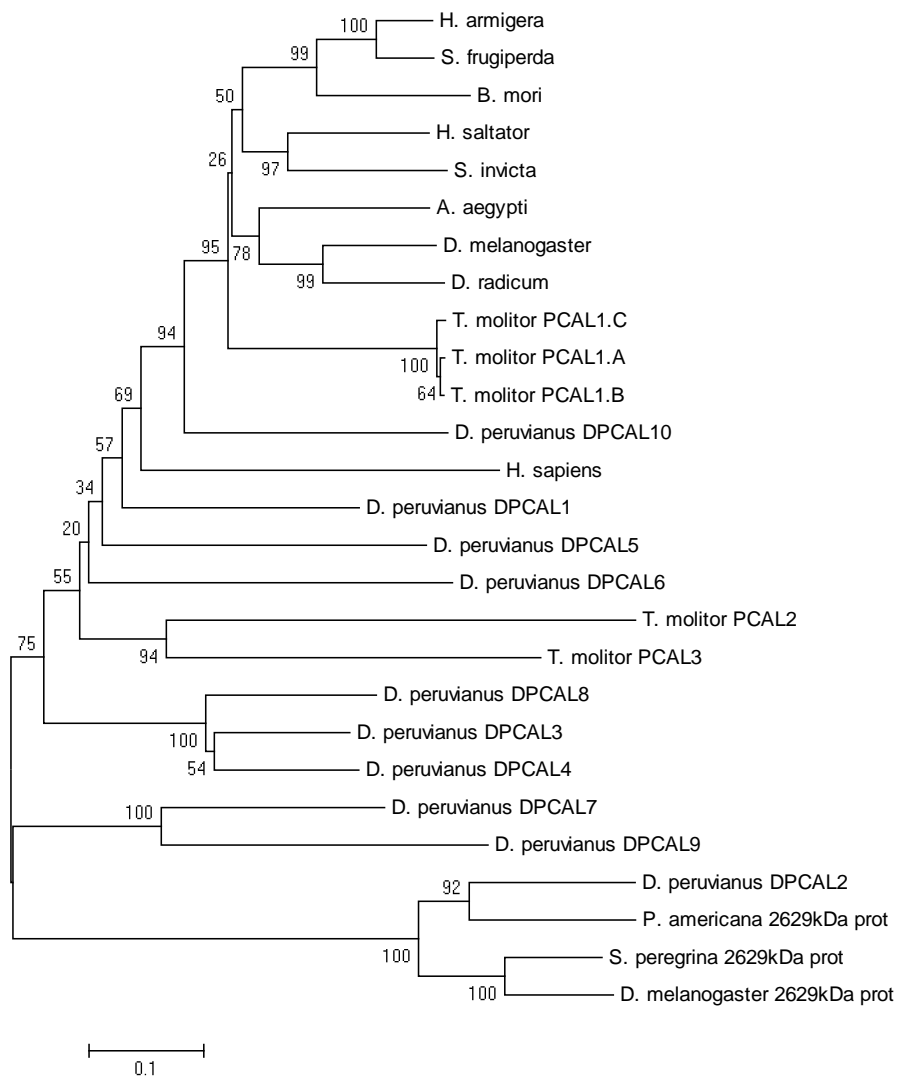
Como a proteína clonada foi expressa na forma de zimógeno, contendo o propeptídeo, é necessário antes da verificação da atividade realizar a ativação *in vitro* desta enzima recombinante. O experimento que variou as condições de ativação mostrou a necessidade de um pH baixo (pH 2,8) para ocorrer a auto ativação (figura 21). Portanto, o processo de ativação não se deu em pH 5,0. A inibição da atividade, ou do processo de ativação, pelo inibidor E-64 é um indício de se tratar da cisteína peptidase clonada (DPCAL1). Na ausência de incubação (SI da figura 21), a atividade corresponderia à atividade intrínseca das proteinases bacterianas ativas, sendo que, neste experimento a atividade pode ser considerada insignificante.

#### 4.2.3.2. DPCAL2

O estudo da DPCAL2 nos daria a possibilidade de conhecer melhor essa enzima de expressão pouco usual. Conhecer sua especificidade e localização celular nos permitiria propor sua ação. Contudo, sua clonagem não foi conseguida mesmo após variações experimentais.

Os resultados até agora mostram duas características intrigantes: a presença de um N-terminal maior que as outras catepsinas L e sua expressão em tecidos específicos como túbulos de Malpighi e corpo gorduroso (figura 22).

Para compreender melhor a função desta enzima, foi feita a análise de domínios conservados (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) que identificou, além dos domínios típicos das cisteino peptidase (Peptidase\_C1 superfamily), o domínio DUF4412 na região N-terminal. Esse domínio não foi caracterizado e, portanto, tem função desconhecida (*Domain of unknown function*). Esse domínio é encontrado em bactérias, arqueias e eucariotos e tem entre 75 e 104 aminoácidos de extensão. Há dentre das cisteino proteinases com o mesmo domínio N-terminal aquelas que apresentam anotação “26/29-kDa proteinase”. Quando comparada conjuntamente, a DPCAL2 se agrupa com as 26/29-kDa proteinases e não com as outras catepsinas L (figura 22).

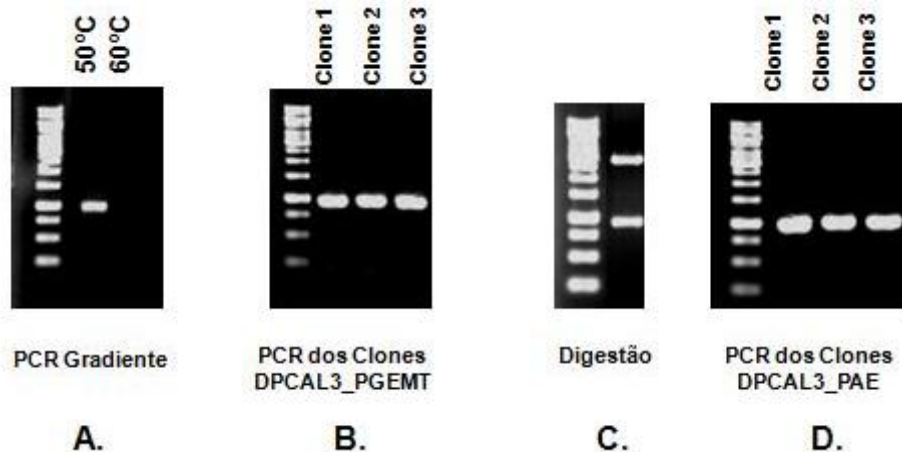


**Figura 22.** Cladograma de diferentes catepsinas L e 26/29-kDa proteinases. DPCAL2 forma um grupo monofilético com outras 26/29-kDa proteinases.

#### 4.2.3.3. DPCAL3

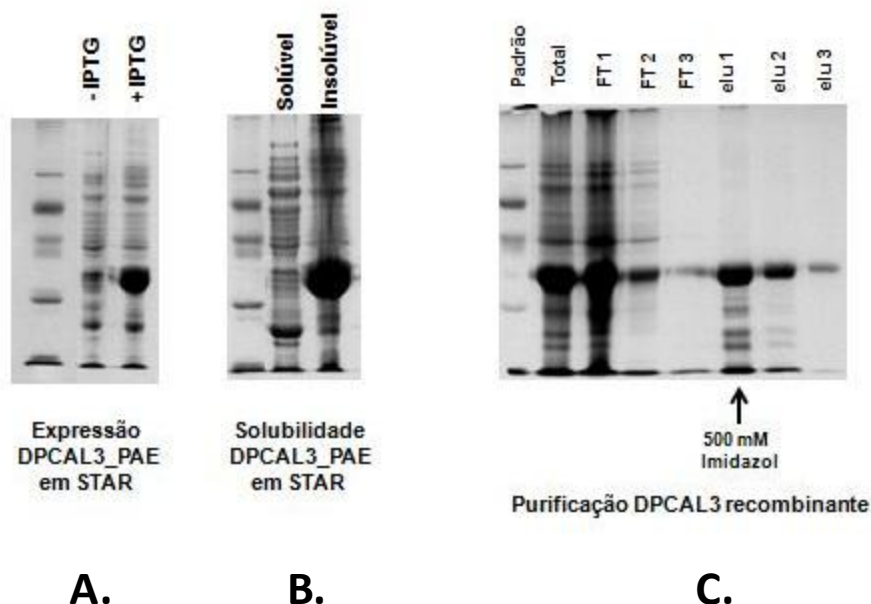
O estudo mais aprofundado da DPCAL3 nos permitiria acessar informações como a especificidade desta enzima e mesmo garantir mais um marcador do conteúdo interno das vesículas de dupla membrana. Foi possível clonar o gene da DPCAL3 e, assim, através da técnica de subclonagem, obter a construção DPCAL3\_PAE (figura 23).





**Figura 23.** Eletroforese (em gel de agarose) dos produtos de amplificação do processo de clonagem da DPCAL3. A. PCR realizado sobre cDNA dos túbulos de Malpighi, foram utilizadas duas temperaturas de pareamento, sendo 50°C considerado padrão para os próximos experimentos. B. PCR de três clones de DPCAL3\_PGEMT. C. Digestão da construção DPCAL3\_PGEMT Clone 1 com as enzimas de restrição XhoI e HindIII. D. PCR de três clones de DPCAL3\_PAE.

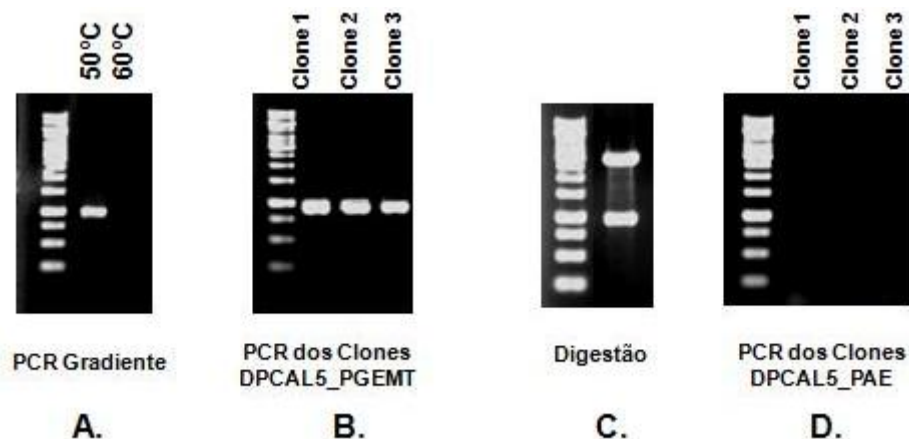
Com a construção foi possível realizar o experimento de expressão heteróloga. A proteína recombinante foi conseguida na forma insolúvel, sendo assim, sua purificação se deu em solução com alta concentração de ureia (figura 24). O processo de purificação, por enquanto gerou apenas frações bastante enriquecidas da proteína recombinante, mas ainda não pura.



**Figura 24.** Eletroforese (SDS-PAGE) das proteínas expressas em *E. coli* durante processo de expressão e purificação da DPCAL3 recombinante. A. Segunda raia uma amostra das bactérias antes da adição de IPTG no meio, e a terceira raia apresenta outra amostra de bactérias do mesmo experimento horas após a adição de IPTG. B. produto da lise das bactérias induzidas (ou após a adição de IPTG) após centrifugação, sendo a segunda raia a fração solúvel e a terceira raia a fração insolúvel. C. Processo de purificação como já descrito. O lisado celular bruto solubilizado em ureia 8M (raia "Total") é adicionado à resina que é então lavada (raias "FT"), sendo que a adição de imidazol na concentração 500mM libera a proteína recombinante ligada à resina (frações "elu").

#### 4.2.3.4. DPCAL5

Foi possível obter a construção DPCAL5\_PGEMT (figura 25) com seus respectivos clones que, tanto pelo sequenciamento do tipo Sanger quanto pelo processo de digestão, correspondem a sequência esperada, contudo sua clonagem no vetor de expressão PAE não foi obtida.



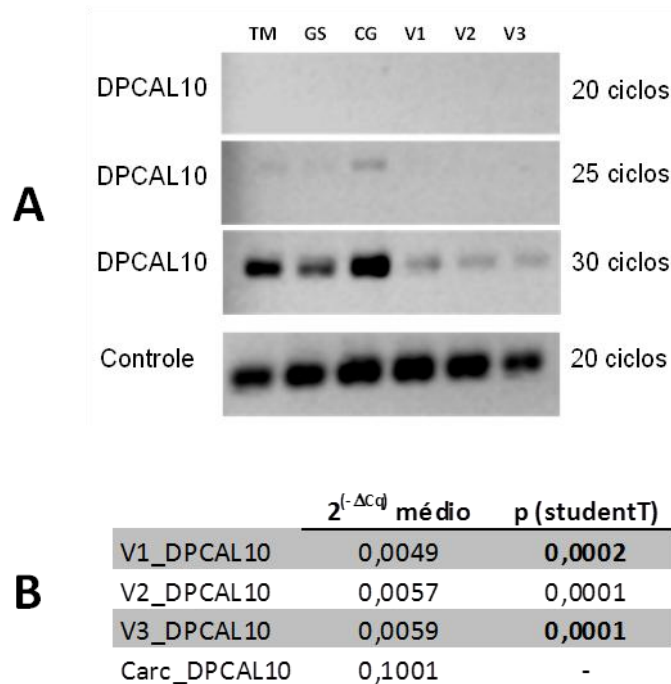
**Figura 25.** Eletroforese (em gel de agarose) do produto de amplificação durante o processo de clonagem da DPCAL5. A. PCR realizado sobre cDNA do intestino, foram utilizadas duas temperaturas de pareamento, sendo 50°C considerado padrão para os próximos experimentos. B. PCR de três clones de DPCAL5\_PGEMT. C. Digestão da construção DPCAL5\_PGEMT Clone 1 com as enzimas de restrição BamHI e EcoRI. D. PCR de três clones de DPCAL5\_PAE.

#### 4.2.3.5. DPCAL10

Como discutimos anteriormente, a DPCAL10 é a mais forte candidata de enzima lisossômica. Isso quer dizer que temos possivelmente uma enzima que conservou sua localização celular nos lisossomos nos diversos tecidos, podendo, como as outras catepsinas L, atuar na digestão no lúmen intestinal. Portanto, ter essa proteína expressa nos permitiria acessar sua localização subcelular (ao se produzir anticorpos e imunocitocalizar) e caracterizar a rota lisossômica e secretória.

A clonagem da DPCAL10 a partir do cDNA intestinal já indicava a real expressão desse gene na intestino médio. Considerando a importância dessa informação e as fracas bandas da figura 10, foi necessário nos certificar da expressão desse gene no intestino médio. Foi realizado, então, um novo

experimento usando um novo par de iniciadores (os mesmos utilizados no processo de clonagem da DPCAL10) além de um novo cDNA e o acompanhamento de vários ciclos de amplificação (20, 25 e 30 ciclos). Na figura 26 A vemos o mesmo padrão e podemos confirmar a expressão da DPCAL10 nos seis tecidos estudados.

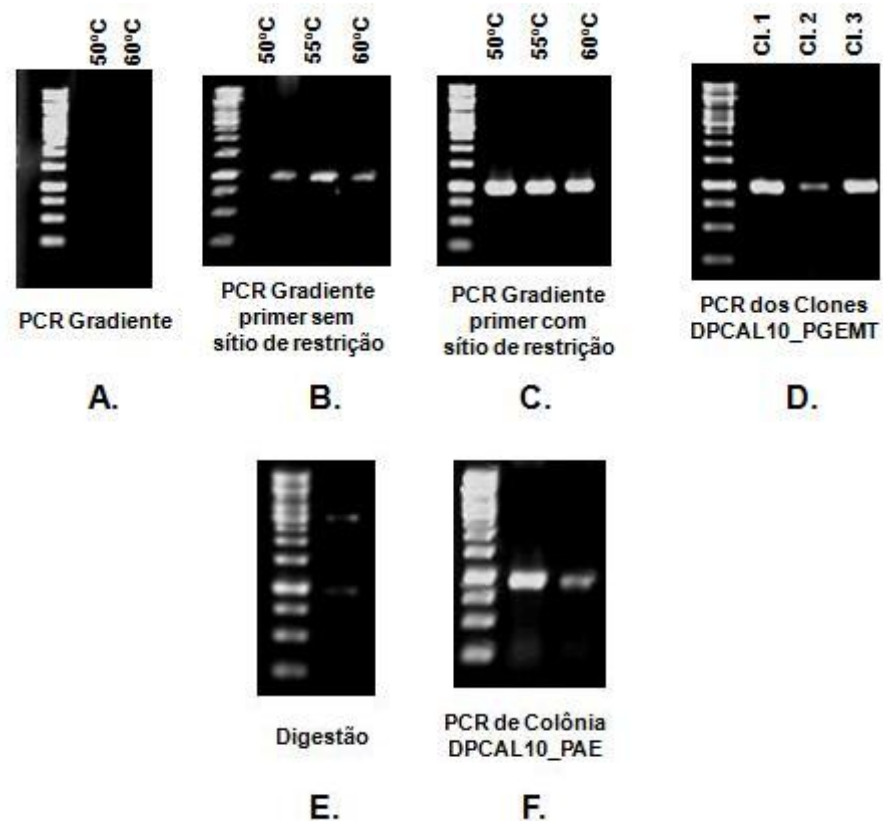


**Figura 26.** Expressão do RNAm que codifica DPCAL10. **A.** Avaliada por RT-PCR em diferentes ciclos de amplificação tendo 18S como controle. **B.** Avaliada por qPCR e é mostrada como  $2^{(-\Delta Cq)}$  do gene de interesse tendo 18S como controle interno. O teste t foi realizado comparando a expressão de DPCAL10 em cada tecido do intestino médio (individualmente) com a expressão de DPCAL10 na carcaça.

A medida da expressão por qPCR confirma a expressão nos tecidos do intestino médio (figura 26 B) e deixa claro que a expressão no intestino médio é menor (*t*-teste,  $p < 0,01$ ) quando comparada com a os demais tecidos (carcaça). Essa informação já era sugerida pela intensidade das bandas no experimento anterior. A vantagem desta abordagem (por qPCR) é a possibilidade de confirmar a amplificação específica pela análise da curva de fusão (*melting*). De maneira geral, podemos concluir que a catepsina em questão é, mesmo que em menor grau, expressa no intestino médio de *D. peruvianus*.

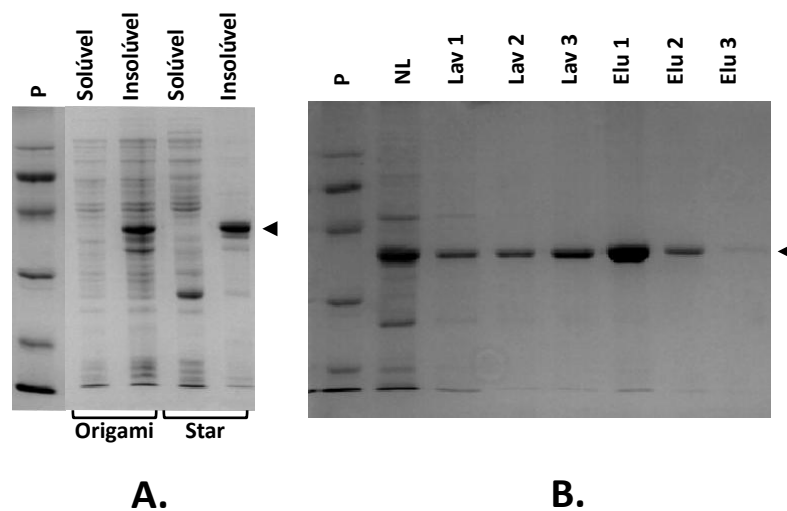
Inicialmente, não houve amplificação com os iniciadores de clonagem (que continham os sítios de restrição), por isso, novos iniciadores foram construídos sem o sítio de restrição. Com o produto dessa primeira amplificação foi realizada uma

segunda amplificação agora com os sítios de restrição. Assim foi possível fazer a primeira construção DPCAL10\_PGEMT que teve sua sequência confirmada por sequenciamento Sanger e pela digestão (figura 27). A construção obtida a partir da ligação do inserto do gene ao vetor de expressão gerou a construção DPCAL10\_PAE que foi, subsequentemente, usada na transformação de bactérias competentes.



**Figura 27.** Eletroforese (em gel de agarose) dos produtos de amplificação durante o processo de clonagem da DPCAL10. **A.** PCR realizado sobre cDNA do corpo gorduroso, foram utilizadas duas temperaturas de pareamento no primeiro experimento que teve o uso dos iniciadores de clonagem (com sítio de restrição). **B.** PCR com iniciadores sem o sítio de restrição com três diferentes temperaturas de pareamento. **C.** PCR com iniciadores de clonagem sobre DNA purificado do passo anterior em três diferentes temperaturas. **D.** PCR de três clones de DPCAL10\_PGEMT. **E.** Digestão da construção DPCAL10\_PGEMT Clone 1 com as enzimas de restrição BamHI e PstI. **F.** PCR de dois clones de DPCAL10\_PAE.

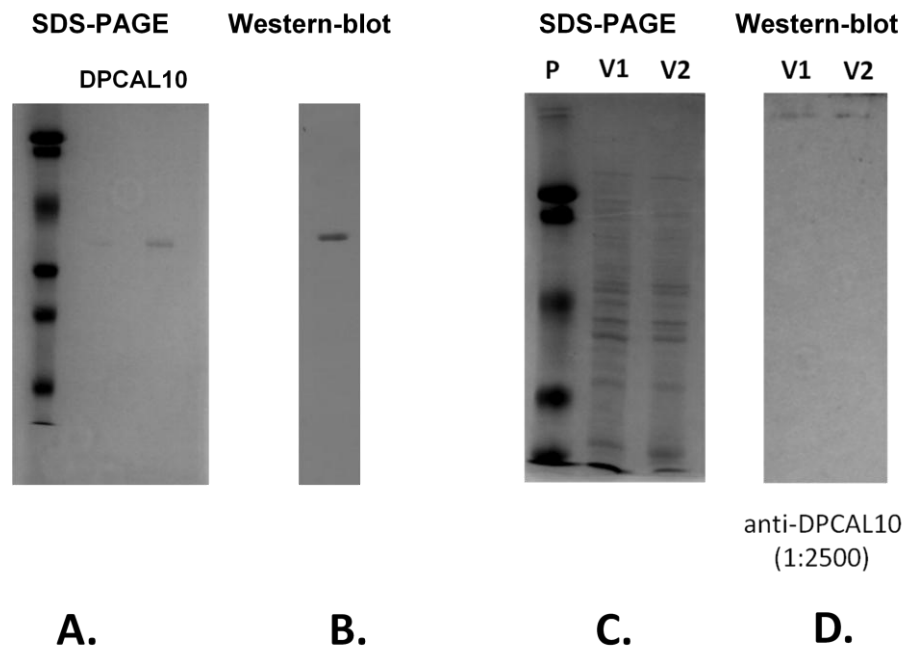
Foi realizado um experimento de expressão heteróloga, sendo que para controle se realizou PCRs do plasmídeo de expressão (DPCAL10\_PAE) bem como da colônia utilizada no experimento. Após a adição de IPTG foi gerada a proteína recombinante. A proteína recombinante foi produzida nas cepas de *E. coli* STAR e Origami na forma insolúvel (figura 28 A). Um passo de purificação foi feito após solubilização da proteína recombinante em ureia 8M (figura 28 B).



**Figura 28.** Eletroforeses (SDS-PAGE) das proteínas expressas em duas cepas de *E. coli* (Origami e Star) durante o processo de purificação da DPCAL10 recombinante. **A.** Produto solúvel e insolúvel da lise de bactérias após a indução da produção da proteína recombinante. **B.** Processo de purificação da proteína recombinante sendo: NL, proteínas não ligadas à resina de purificação. Lav1-3, proteínas liberadas em lavagens consecutivas. Elu1-3, proteínas liberadas com a adição de 500mM de imidazol ao tampão de lavagem.

A produção da catepsina L possivelmente lisossômica (DPCAL10) possibilitou a produção de anticorpos policlonais em coelho. A produção da enzima na forma insolúvel nas condições testadas não permite a análise da atividade proteolítica da enzima, como foi conseguido com a DPCAL1.

O anticorpo policlonal anti-DPCAL10 é capaz de reconhecer a proteína recombinante (figura 29 A e B) e o mesmo experimento realizado com o soro pré-imune não apresenta marcação. Mesmo reconhecendo a proteína recombinante, o anticorpo não foi capaz de reconhecer uma banda dentre as proteínas solúveis de V1 e V2 de *D. peruvianus* nas condições testadas (Figura 29 C e D). Porém, diferentemente da anti-DPCAL1, o anti-DPCAL10 parece não marcar inespecificamente quando aumentada a concentração (avaliando a marcação dos padrões de peso molecular).

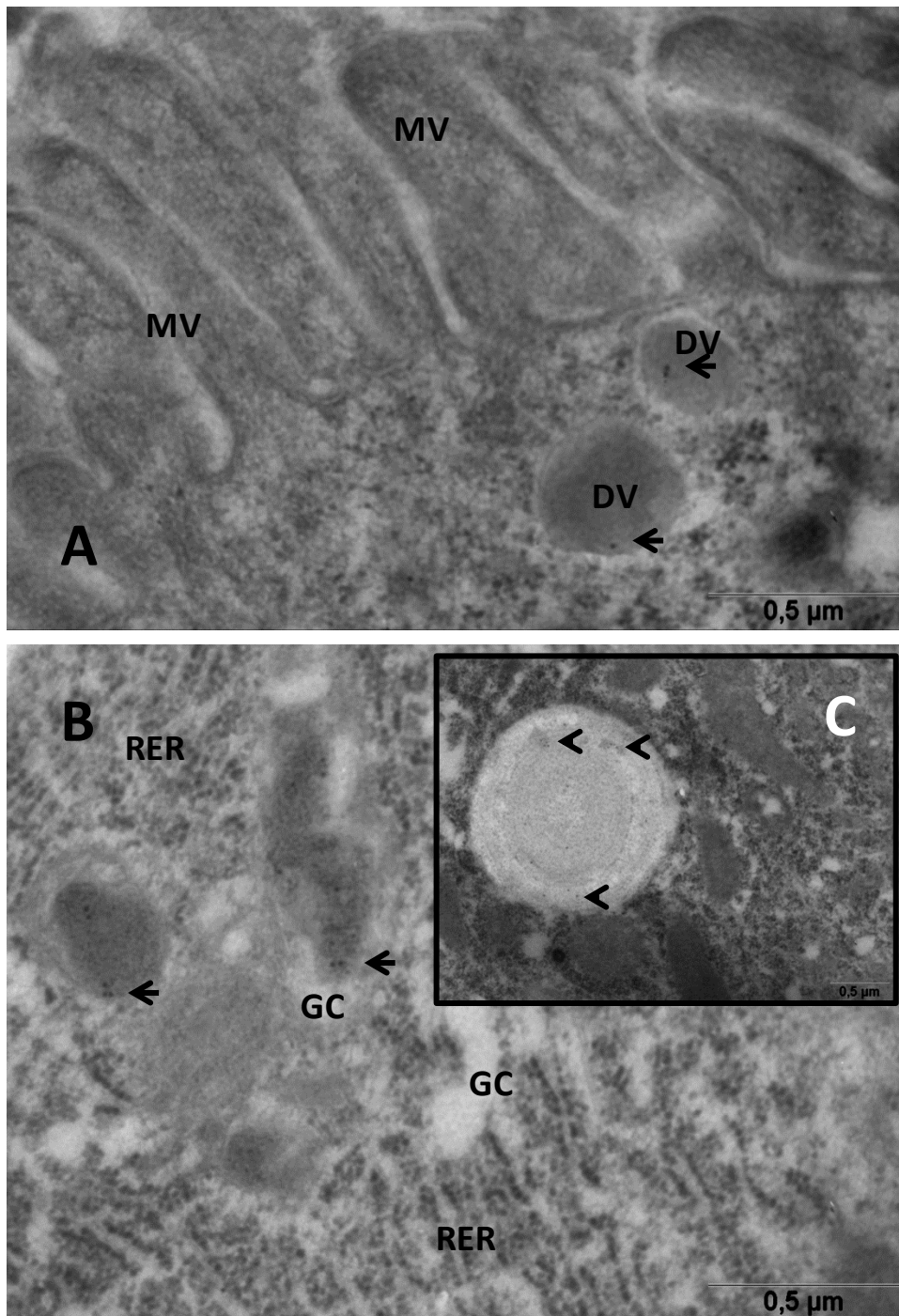


**Figura 29.** Eletroforese (SDS-PAGE) e *western blot* da DPCAL10 **A.** SDS-PAGE da DPCAL10 recombinante corado por Coomassie Blue **B.** *Western blotting* com anticorpo policlonal anti-DPCAL10 sobre proteína recombinante **C.** SDS-PAGE das proteínas solúveis de V1 e V2 coradas por Coomassie Blue. **D.** *Western blotting* com anticorpos policlonais anti-DPCAL10 com título de 1:2500 sobre proteínas solúveis de V1 e V2.

Como forma de superar as dificuldades experimentais desta fase, foram feitos testes nos quais a concentração de anticorpo foi igualada no pré-imune e no soro imune. Em diluições pareadas, os soros foram empregados na marcação de cortes ultrafinos. Em condições nas quais o pré-imune não gerava marcação o soro imune de mesma concentração era avaliado. Quando houve marcação no soro pré-imune, a figura é mostrada para fins de comparação. Esse resultado e de outros anticorpos testados estão resumidos na tabela 10 na seção que trata dos marcadores celulares.

O anticorpo policlonal anti-DPCAL10 foi capaz de marcar o interior de vesículas de secreção assim como vesículas relacionadas ao complexo de Golgi (figura 30 A e B). Houve marcação em organelas ainda não identificadas (figuras 32C e 33). O mesmo experimento realizado com soro pré-imune de mesma diluição não gerou marcações significativas.

A organela da figura 30 C é pouco eletrondensa e seu conteúdo forma estruturas concêntricas provalmente membranosas. Esta organela só permanece visível em fixações mais brandas, como as usada nas marcações por anticorpos. As marcações estão preferencialmente no conteúdo mais eletrondenso desta organela (Figura 30 C).

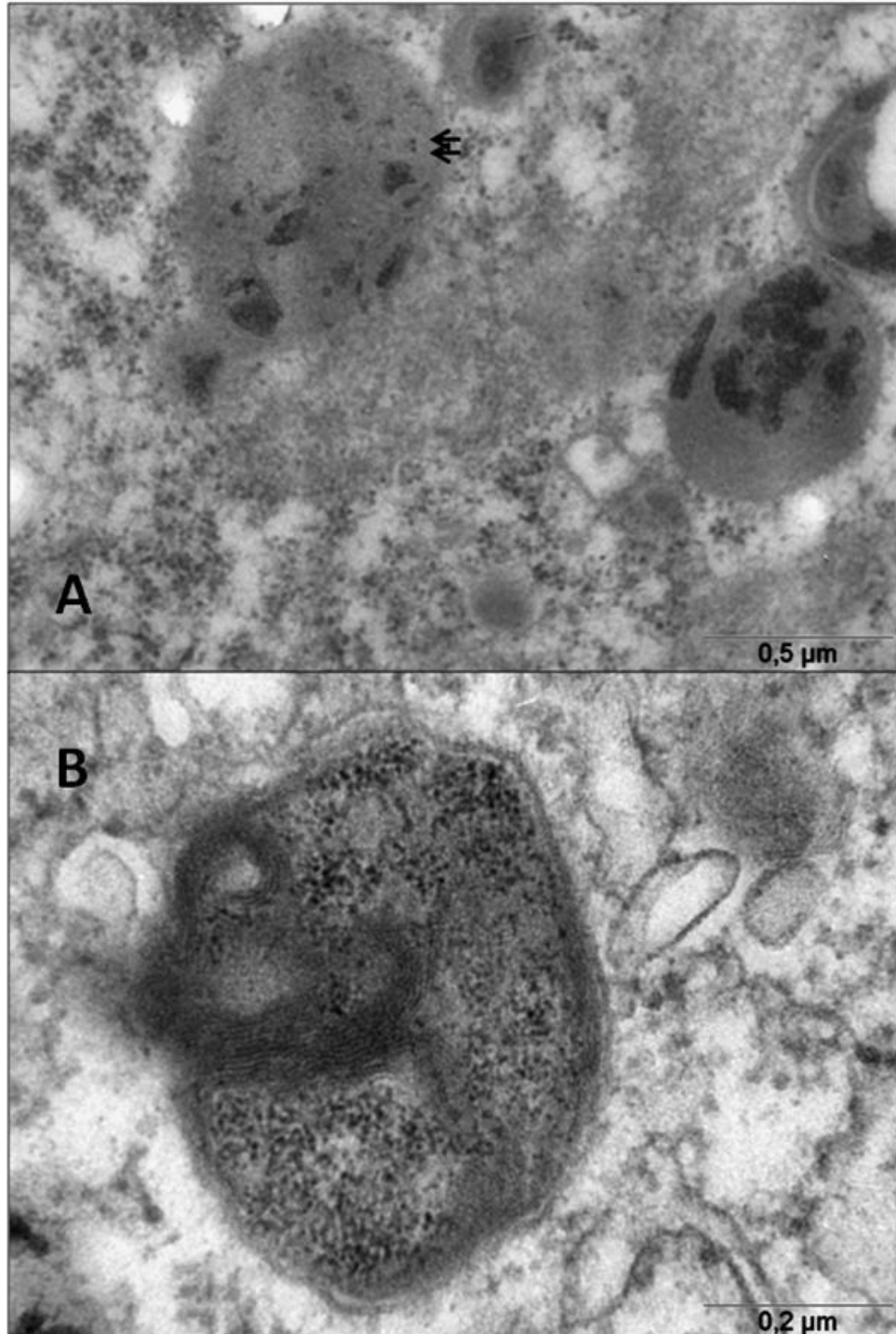


**Figura 30.** Micrografia de enterócito de V1 de *Dysdercus peruvianus* imunomarcados por anti-DPCAL10. **A.** Região do ápice celular. **B.** Região que inclui o retículo endoplasmático rugoso e complexo de Golgi. **C.** Organela pouco eletrondensa não identificada. Em todos os casos a diluição do anticorpo foi 1:500. As setas indicam as regiões marcadas. MV, microvilosidade. DV, vesícula de secreção (de dupla membrana). GC, complexo de Golgi. RER, retículo endoplasmático rugoso.

O segundo tipo de organela marcada é caracterizada pela presença de conjuntos concêntricos de membranas alternado com um conteúdo mais eletrondenso (figura 31). Essa organela é mais resistente a fixação e permite a



confirmação do conjunto de membranas entremeando o conteúdo mais eletrondenso (figura 31 B).



**Figura 31.** Micrografia de enterócito de V1 de *Dysdercus peruvianus* – organela marcada. **A.** Organela relacionada ao lisossomo. Corte imunomarcado por anti-DPCAL10 com setas indicando a região marcada. **B.** Organela relacionada ao lisossomo. Corte com fixação tradicional não marcado.

A marcação parece indicar que a catepsina L tipicamente lisossômica seria realmente encaminhada para a rota secretora. Devido as dificuldades experimentais

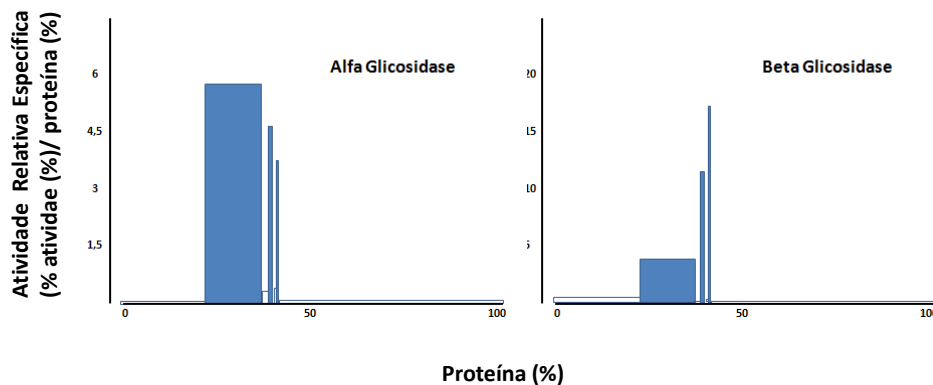


já citadas, não foi possível confirmar que a marcação da DPCAL10 é específica, ou seja, se o anti-DPCAL10 não marcaria também as outras DPCALs.

### 4.3. Produção de membrana perimicrovilar

#### 4.3.1. Fracionamento celular por centrifugação diferencial

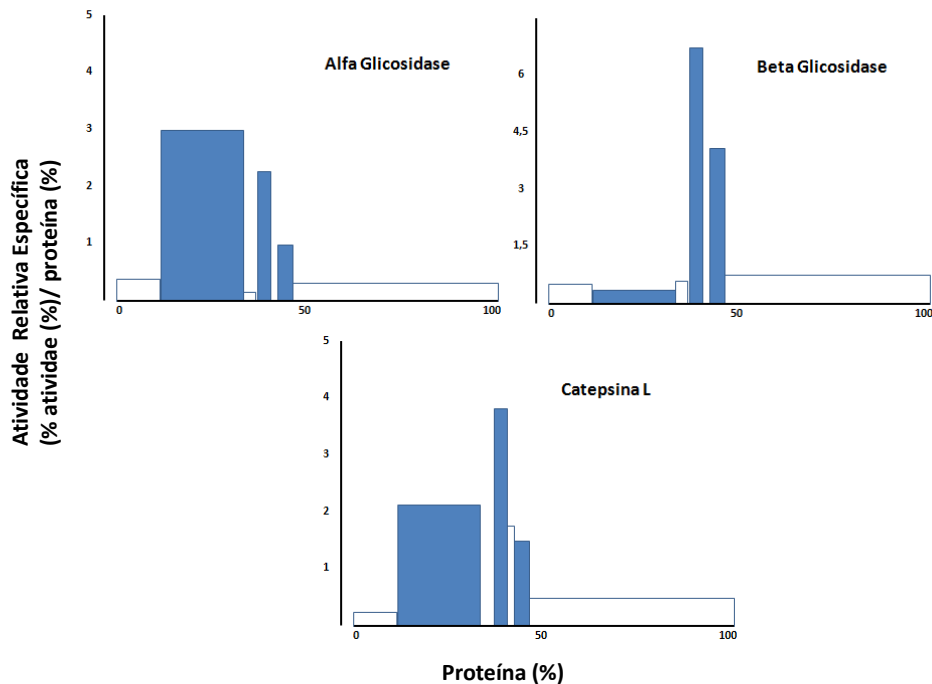
Para definir corretamente os parâmetros experimentais foram realizados experimentos de centrifugação diferencial em condições similares às de Silva et al. (1996). Nesse experimento foi utilizada a porção V1 do intestino de *D. peruvianus*. Essa porção, por não apresentar grande atividade de proteinases quando comparado com as outras porções, seria mais estável aos experimentos e foi então selecionada para os testes iniciais. A figura 32 mostra o gráfico de De Duve para duas atividades, a atividade de alfa-glicosidase (marcador de membrana perimicrovilar) e a atividade de beta-glicosidase (marcador de membrana microvilar conforme Silva et al. (1996)). Os experimentos aqui expostos tem pelo menos uma réplica biológica que reproduz o resultado demonstrado.



**Figura 32.** Centrifugação diferencial de homogeneizados do tecido de V1 em meio isotônico. Após a centrifugação diferencial, cada fração foi congelada e descongelada três vezes e novamente centrifugada gerando uma porção solúvel (SP) e uma insolúvel (PP). Solúvel (representado por barras brancas) e insolúvel (representado por preenchidas) de cada fração são ensaiados e foram aqui representados na ordem de obtenção, da esquerda para a direita (P1, P2, P3 e FS). A altura do histograma é proporcional à atividade específica relativa e sua área é proporcional à porcentagem da atividade enzimática total.

O passo seguinte seria utilizar a porção V2 do intestino de *D. peruvianus*, pois nessa porção poderíamos seguir todos os marcadores celulares já estudados e também as catepsinas L, que são os marcadores selecionados para este estudo.

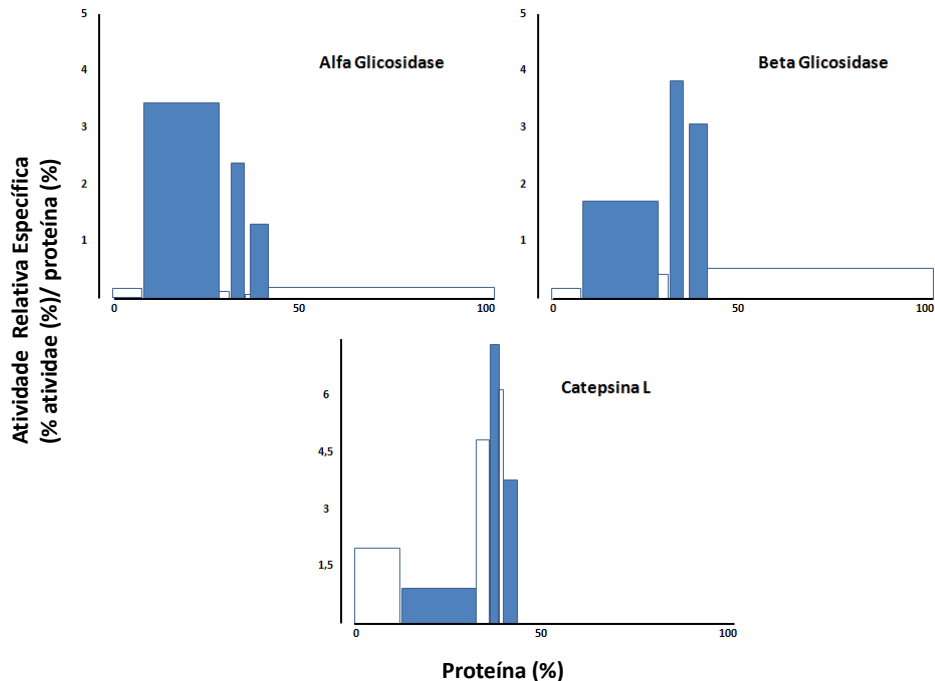
A marcante presença de proteinases em V2 faz com que seja necessária a adição de inibidores que contêm a atividade especialmente da cisteíno peptidase. Para isso se testou duas abordagens, uma que envolve o uso do modificador reversível MTS para cisteína proteinases e outra que utiliza inibidores específicos para as proteinases cisteínicas e aspárticas (E-64 e pepstatina).



**Figura 33.** Centrifugação diferencial de homogeneizados do tecido de V2 em meio isotônico na presença dos inibidores pepstatina 10 $\mu$ M e E-64 10 $\mu$ M. Após a centrifugação diferencial, cada fração foi congelada e descongelada três vezes e novamente centrifugada gerando uma porção solúvel (SP) e uma insolúvel (PP). Solúvel (representado por barras brancas) e insolúvel (representado por barras preenchidas) de cada fração são ensaiados e foram aqui representados na ordem de obtenção, da esquerda para a direita (P1, P2, P3 e FS). A altura do histograma é proporcional à atividade específica relativa e sua área é proporcional à porcentagem da atividade enzimática total.

A presença dos inibidores pepstatina e E-64 cumpriu seu papel, pois a atividade das proteinases do tipo catepsina L foi quase banida neste experimento (cerca de 2% remanescentes), quando comparado à atividade do tratamento com MTS. O dado da inibição não é observável nos gráficos já que representam uma proporção da atividade total. A presença de inibidores de protease não afetou a atividade de alfa e beta glicosidase. Contudo, os padrões de distribuição no gráfico de De Duve sofreu influência, especialmente na atividade de beta-glicosidase, deixando em dúvida a integridade das frações subcelulares (figura 33). A razão mais

provável para esse arranjo é a presença do solvente dos inibidores (DMSO 2% (v/v)) que pode ter influenciado a estabilidade das membranas.

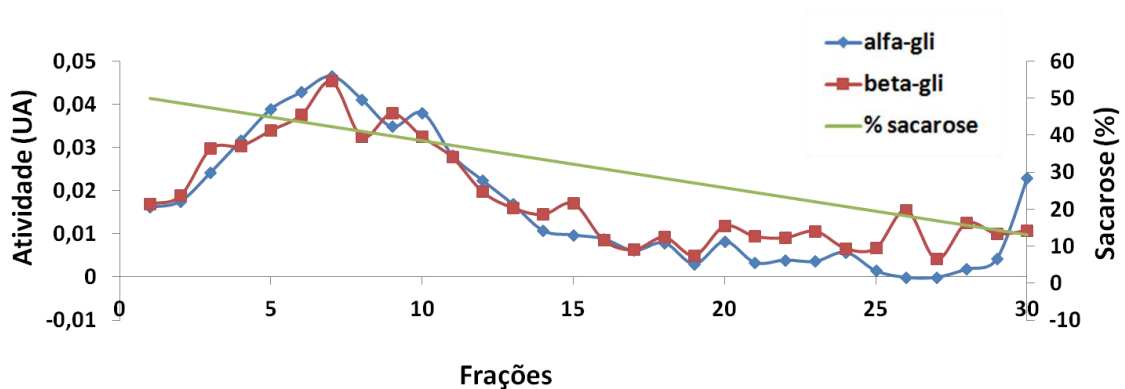


**Figura 34.** Centrifugação diferencial de homogeneizados do tecido de V2 em meio isotônico na presença de MTS (1 mM). Após a centrifugação diferencial, cada fração foi congelada e descongelada três vezes e novamente centrifugada gerando uma porção solúvel (SP) e uma insolúvel (PP). Solúvel (representado por barras brancas) e insolúvel (representado por barras preenchidas) de cada fração são ensaiados e foram aqui representados na ordem de obtenção, da esquerda para a direita (P1, P2, P3 e FS). A altura do histograma é proporcional à atividade específica relativa e sua área é proporcional à porcentagem da atividade enzimática total.

No experimento de centrifugação diferencial na presença de MTS (1mM) foi possível repetir o padrão de distribuição dos marcadores de membranas (figura 34). A atividade de cathepsina L é parcialmente solúvel e parcialmente ligada à fração insolúvel (ou fração de membranas) enriquecida especialmente em P2 e P3. Contudo, sua distribuição foi bastante parecida com o resultado obtido por Silva et al. (1996) quanto à atividade de fosfatase ácida (um marcador lisossômico). A desprezível atividade no solúvel final (FS) indica que houve pouca contaminação do tecido (durante a dissecação) por conteúdo luminal intestinal.

#### 4.3.2. Fracionamento celular (resolução em gradiente de sacarose)

As frações subcelulares resultantes da centrifugação diferencial foram selecionadas para serem resolvidas por centrifugação em gradiente de sacarose. Com o objetivo resolver vesículas de secreção e lisossomos, as frações escolhidas correspondem a P2 e P3, pois estas são enriquecidas na atividade de alfa-glicosidase, de beta-glicosidase e de catepsina L (os dois primeiros são marcados de membranas e o último é marcador de material vesicular solúvel – lisossomo e vesícula de secreção).



**Figura 35.** Atividade enzimática de frações subcelulares de V1 centrifugadas em gradiente de sacarose. alfa-gli, atividade sobre p-nitrofenil-alfa-D-glicosídeo. beta-gli, atividade sobre p-nitrofenil-beta-D-glicosídeo. UA, (unidades arbitrárias).

Esse método permite que frações subcelulares de densidades diferentes se separem. Considerando isso, as atividade de alfa e beta-glicosidase, por apresentarem o mesmo padrão de distribuição no gradiente de sacarose (figura 35), estão associadas, provavelmente compondo o mesmo compartimento subcelular. Esse resultado é o mesmo para frações subcelulares de V1 e V2.

Podemos dizer que frações subcelulares (vesículas) de dupla membrana estão sendo isoladas neste experimento. Contudo, não foi possível obter enzimas solúveis dessas amostras. Ensaio para atividade de catepsina L indicaram que o conteúdo dessas vesículas é liberado entre o processo de centrifugação diferencial e a aplicação no gradiente de sacarose, mesmo com a manutenção da pressão osmótica. Para acessar as enzimas solúveis que são secretadas, uma nova estratégia foi traçada: a identificação das proteínas secretadas com uso de técnicas proteômicas sobre o conteúdo solúvel de V1.

### 4.3.3. Estudo dos marcadores celulares selecionados

O conjunto de dados da literatura guiou a escolha dos marcadores celulares (em especial: Silva et al. (1996) e Ferreira et al. (1988)). Para isso, a atividade hidrolítica descrita nesses trabalhos é relacionada com as identidades das sequências do transcriptoma intestinal. Como em todos os casos houve mais de uma sequência com anotações idênticas, análises de expressão baseadas no número de reads, RT-PCR e qPCR foram fundamentais na escolha do marcador.

Foi selecionado um marcador de cada compartimento na rota secretória regulada (Griffiths e Simons, 1986), além das catepsinas L, das quais se acumulam evidências da sua secreção nos Heteroptera (Billingsley e Downe, 1988; Lopez-Ordoñez, *et al.*, 2001).

#### 4.3.3.1. Alfa-glicosidases gerais de *D. peruvianus*

As alfa-glicosidases gerais aqui consideradas envolvem as alfa-glicosidases e alfa-manosidases. As alfa-glicosidases são as enzimas marcadoras de membrana perimicrovilar e já foram purificadas a partir do tecido intestinal de *D. peruvianus* (Silva e Terra, 1995; Silva et al., 1996). A atividade de alfa-manosidase acompanha membranas de maior densidade em *Rhodnius prolixus*, sendo considerada assim como enzima marcadora de membrana microvilar (Ferreira *et al.*, 1988). A associação da alfa-manosidase com a membrana microvilar pode ser inferida também por apresentar o mesmo padrão de distribuição subcelular que a atividade de beta-glicosidase (também marcadora de membrana microvilar) (Silva *et al.*, 1996).

A análise de sequências obtidas no pirosequenciamento permitiu a identificação de 11 sequências de alfa-glicosidases gerais, sendo que destas, três apresentavam a maior representatividade quanto à quantidade de *reads*. Elas foram denominadas DPAIpMan, DPAlphaGlu1 e DPAlphaGlu2 e são sequências formadas, respectivamente, por 2363, 1319 e 677 *reads*, contrastando com os 169 *reads* da quarta sequência de RNAm mais abundante do transcriptoma intestinal.

A DPAIpMan é uma enzima da família 31 das glicosídeo hidrolases (pfam01055) e tem a maior quantidade de *reads*. Já as sequências da DPAlphaGlu1

e DPAlphaGlu2 são enzimas da família 13 das glicosídeo hidrolases com domínio catalítico compartilhado com as alfa-amilases (pfam00128).

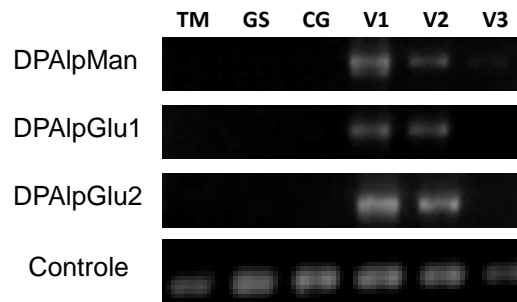
A família 31 inclui tanto enzimas com atividade de alfa-manosidase (EC 3.2.1.24) como também enzimas com atividade de alfa-glicosidase (EC 3.2.1.20) (<http://www.cazy.org/GH31.html>). Já a família 13 não inclui enzimas com atividade de alfa-manosidase. Sendo assim, enzimas da família 13 serão consideradas para análise como enzimas responsáveis pela atividade de alfa-glicosidase típica.

Mesmo pertencendo à mesma família, a DPAlphaGlu1 difere da DPAlphaGlu2 por apresentar um domínio transmembranar na extremidade carboxiterminal (figura 36). A presença do domínio transmembranar torna DPAlphaGlu1 uma forte candidata como a enzima marcadora de membrana perimicrovilar purificada por Silva e Terra (1996).

```
DP_AlphaGlu1 494 IPAAANVGFKFKNEKIDTTSKITTFRPSCMVLTASSNDNSSASSVFASIFAILLACLAVTLFRSN 555
DP_AlphaGlu2 494 VPSPNTGFKAGQLFNWSDVILPLSPRSGVVFRY----- 525
```

**Figura 36.** Alinhamento da porção C-terminal da DPAlphaGlu1 e DPAlphaGlu2. Domínio transmembranar sublinhado (predito por HMMTOP v. 2.0). Aminoácidos marcados com o fundo preto e cinza são idênticos e similares, respectivamente.

A sequência de DPAlphaGlu1 foi usada em um Blast (Blastp) e as duas sequências de maior *score* foram do também Heteroptera *Riptortus pedestris* e ambas sequências, quando submetidas à previsão pelo programa HMMTOP, também apresentaram alça transmembranar na região C-terminal (número de acesso BAN20159 e BAN20192.1). Já quando a DPAlphaGlu2 foi usada no Blastp o maior *score* foi um proteína de *R. pedestris* que, quando submetida a previsão, não apresentava a região transmembranar (número de acesso BAN20149.1). Essa análise indica que a alça transmembranar pode ser uma característica conservada e não um simples artefato de montagem.

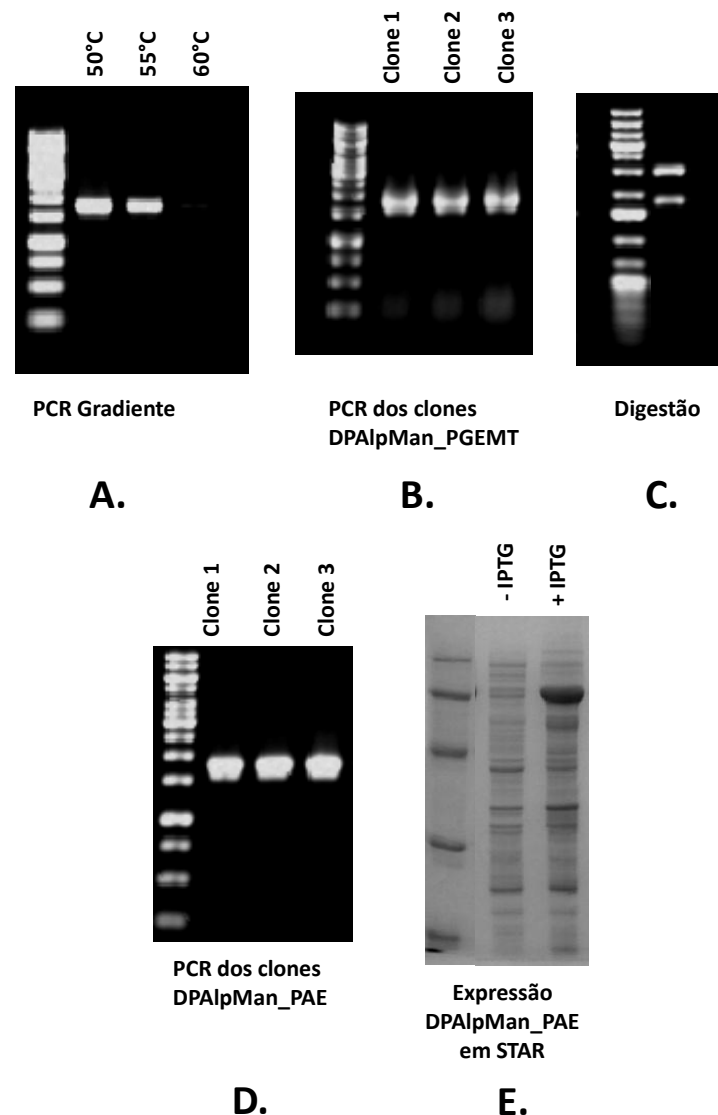


**Figura 37.** Expressão dos RNAm que codificam diferentes alfa-glicosidases avaliada por RT-PCR. TM, túbulos de Malpighi; GS, glândula salivar; CG, corpo gorduroso; V1, primeiro ventrículo do intestino médio; V2, segundo ventrículo do intestino médio; V3, terceiro ventrículo do intestino médio.

Os genes das três alfa-glicosidases gerais são expressos apenas nos tecidos do intestino médio (figura 37). A expressão em V3 foi detectada, mas com menor intensidade.

#### 4.3.3.2. DPAIpMan

O processo de clonagem da DPAIpMan foi feito usando como DNA molde o cDNA produzido a partir do RNAm extraído do intestino de *D. peruvianus*. Dessa forma, foi obtida uma banda de tamanho correspondente, conforme mostrado na figura 38 A, que foi purificada e o DNA foi ligado ao vetor PGEMT gerando os clones (figura 38 B) que, por sua vez, foram digeridos com enzimas de restrição correspondentes (figura 38 C).



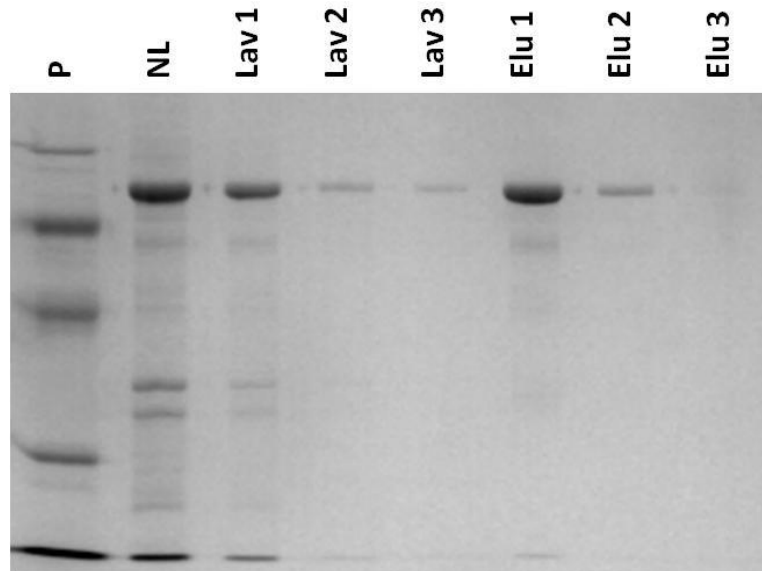
**Figura 38.** Eletroforeses obtidas durante na clonagem e expressão heteróloga de DPAIpMan. **A-D** são eletroforeses em gel de agarose. **A**, **B** e **D** são eletroforeses do produto das reações de amplificação. **E** eletroforese de proteínas (SDS-PAGE) corada por Coomassie blue. **A**. PCR realizado sobre cDNA do intestino, foram utilizadas três temperaturas de pareamento (50°, 55° e 60°). **B**. PCR de três clones de DPAIpMan\_PGEMT. **C**. Digestão da construção DPAIpMan\_PGEMT Clone1 com as enzimas de restrição BamHI e PstI. **D**. PCR de três clones de DPAIpMan\_PAE. **E**. Proteínas expressas em *E. coli* antes (-IPTG) e horas após a adição de IPTG (+IPTG).

A banda de tamanho correspondente foi purificada do gel e o DNA com ponta coesivas foi ligado a um vetor de expressão (PAE), também com pontas coesivas correspondentes, gerando assim a construção DPAIpMan\_PAE com seus respectivos clones (figura 38 D). O plasmídeo de expressão foi usado na transformação de células de *E. coli* da cepa STAR que, após a adição de IPTG, produziu a enzima recombinante (figura 38 E).

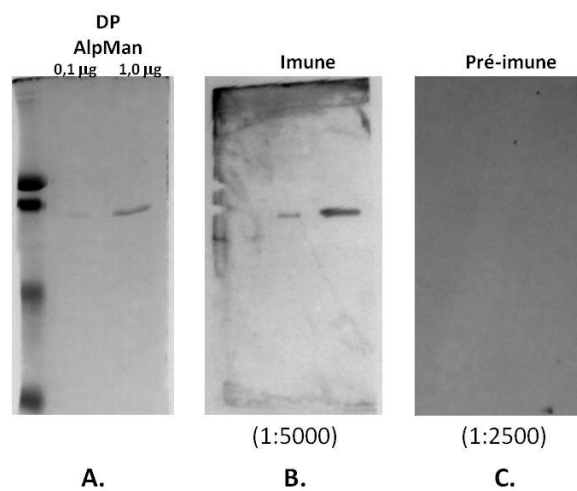
A alfa-manosidasidase de *D. peruvianus* (DPAIpMan) recombinante foi parcialmente purificada para a sua inoculação em coelho (figura 39). O anticorpo



gerado (anti-DPAIpMan) é capaz de reconhecer a proteína recombinante e o soro pré-imune não reconhece a proteína recombinante nas condições testadas (figura 40).



**Figura 39.** Eletroforese (SDS-PAGE) das proteínas insolúveis expressas por *E. coli* (Star) durante o processo de purificação de DPAIpMan recombinante. **P.** Padrão de massa molecular. **NL.** Proteínas não ligadas à resina de purificação. **Lav1-3.** Proteínas liberadas em lavagens consecutivas. **Elu1-3.** Proteínas liberadas com a adição de 500mM de imidazol ao tampão de lavagem.

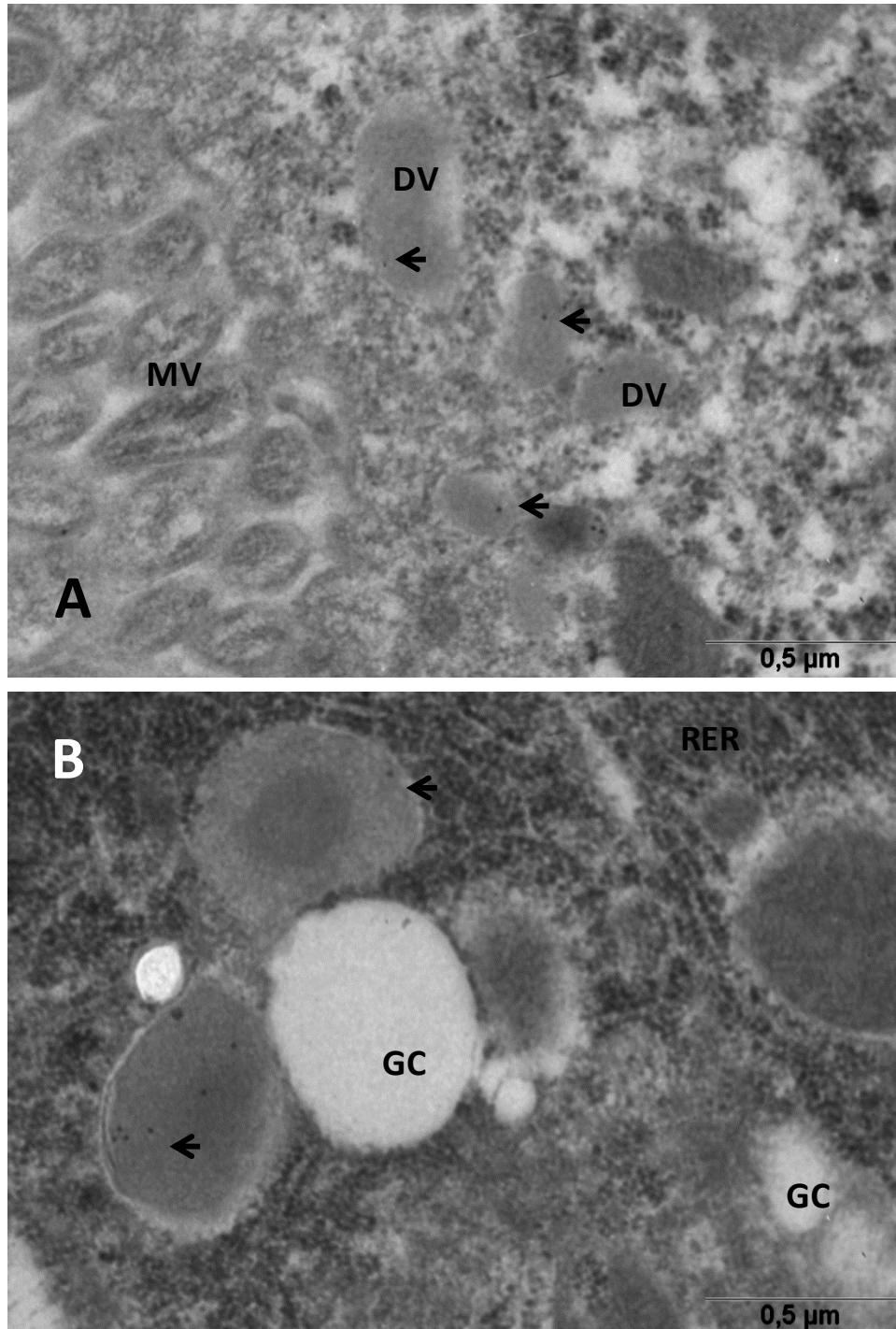


**Figura 40.** Eletroforese (SDS-PAGE) e *western blot* da DPAIpMan recombinante. **A.** Eletroforese da DPAIpMan recombinante semi-purificada, corada por Coomassie Blue. **B.** *Western blot* com soro imune anti-DPAIpMan com título de 1:5000. **C.** *Western blot* com soro pré-imune com título de 1:2500. Superexposto para assegurar a ausência de sinal.

A produção do anti-DPAIpMan e o reconhecimento da proteína recombinante indica a possibilidade da marcação da membrana microvilar em experimentos de imunocitocalização.

A presença de atividade de alfa manosidase associada as membranas celulares apicais foi primeiramente demonstrada em *Rhodnius prolixus*, nesse caso ligada à membrana rica em proteínas, ou seja à membrana microvilar (Terra et al., 1988; Ferreira et al., 1988). A presença desta atividade ligada à membrana microvilar em *D. peruvianus* foi posteriormente verificada e com distribuição celular similar à beta glicosidase. A atividade de alfa manosidase é parcialmente ligada à membrana e parcialmente solúvel (Silva e Terra, 1994; Silva et al., 1996).

O anticorpo policlonal anti-DPMan foi capaz de gerar marcações esparsas em vesículas de secreção de dupla membrana, bem como em vesículas associadas ao complexo de Golgi (figura 41 A e B). O mesmo experimento realizado com soro pré-imune de mesma diluição não gerou marcações significativas.



**Figura 41.** Micrografia de enterócito de V1 de *Dysdercus peruvianus* imunomarcado por anti-DPAIpMan. **A.** Região do ápice celular. **B.** Região que inclui o retículo endoplasmático rugoso e complexo de Golgi. Em ambos os casos a diluição do anticorpo foi 1:500. As setas indicam as regiões marcadas. MV, microvilosidade. DV, vesícula de secreção (de dupla membrana). GC, complexo de Golgi. RER, retículo endoplasmático rugoso.

É importante notar a fixação das amostras para experimentos de imunomarcção nem sempre gera micrografias com definição suficiente para a distinção da dupla membrana. A inferência da natureza da organela nesses casos se

dá não só pela marcação em si mas também pela topologia, eletrondensidade e comparação com micrografias da literatura (em especial Silva et al., 1995).

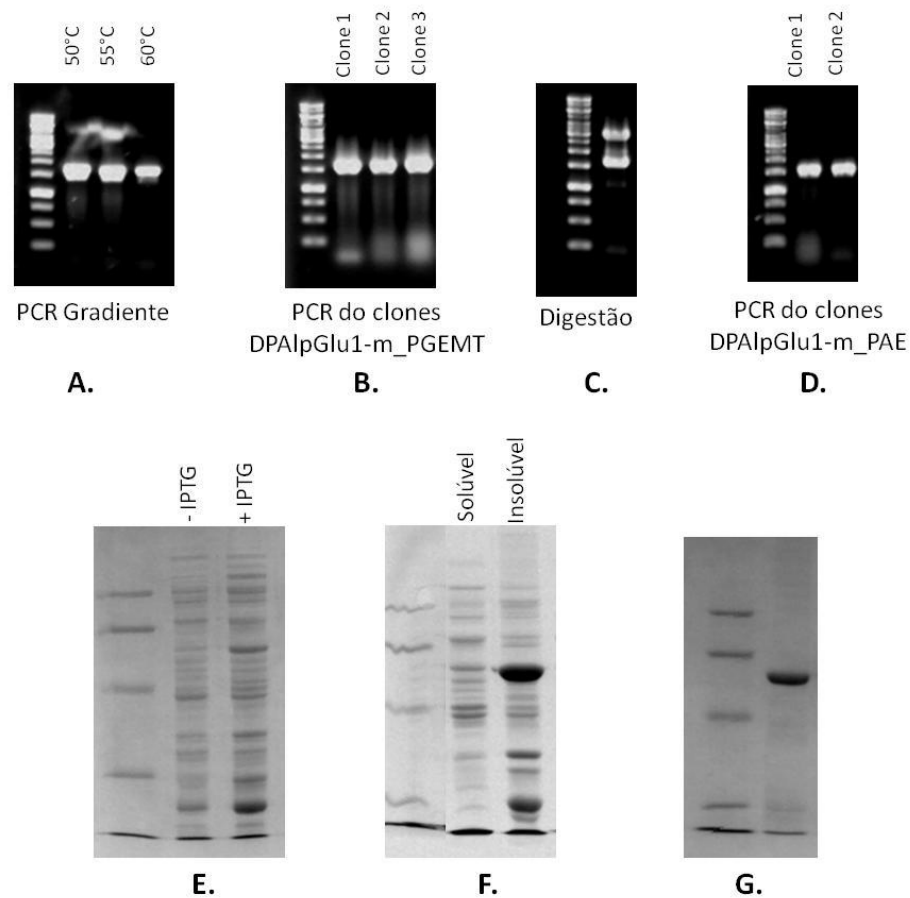
Nas condições testadas, anti-DPAIpMan não foi capaz de marcar a membrana microvilar, possivelmente devido ao título baixo do anticorpo. Contudo, a marcação das vesículas trans-Golgi e das vesículas de secreção indica sua presença na rota secretória. Outras organelas, como aquelas relacionadas aos lisossômos, não apresentaram marcação nas condições do experimento.

#### **4.3.3.3. DPAlphaGlu1**

O processo de clonagem da DPAlphaGlu1 foi feito inicialmente com a região que codifica a porção transmembranar da enzima, necessitando um passo adicional no qual foi isolado o trecho que codifica a região solúvel e, então, foi inserido em um vetor de expressão. O clone obtido inicialmente da DPAlphaGlu1 completa (DPAlphaGlu1-PGEMT) foi utilizado como amostra em uma nova reação de amplificação. Nesta reação foi retirada a porção dos nucleotídeos que codificam a região transmembranar da proteína, passo importante antes da expressão heteróloga em bactérias (figura 42 A).

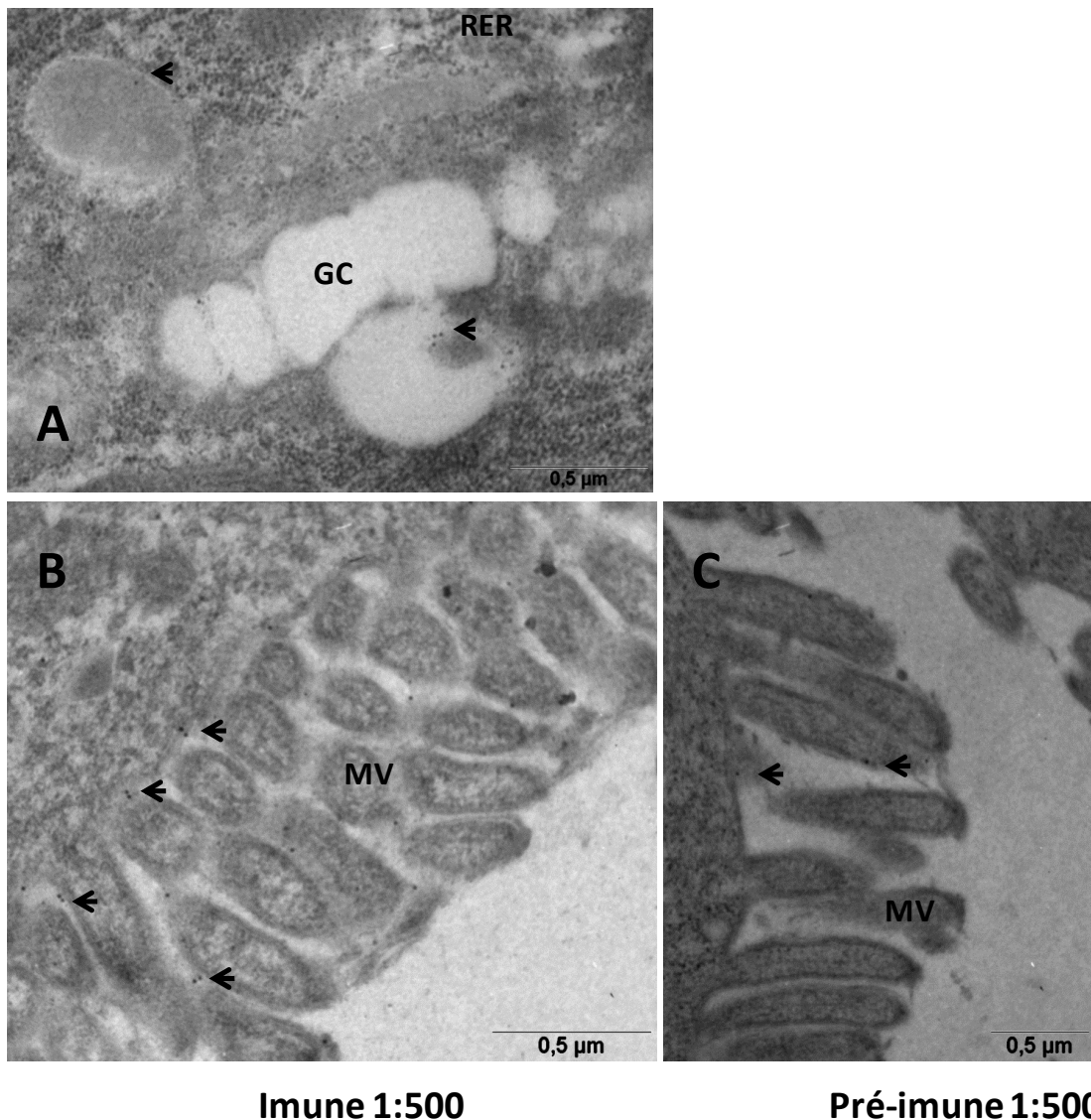
O novo clone gerado (DPAlphaGlu1-m\_PGEMT) foi digerido com enzimas de restrição e o trecho liberado foi ligado no vetor de expressão PAE gerando a construção DPAlphaGlu1-m\_PAE (figura 42 B, C e D). Bactérias *E. coli* com essa construção são capazes de produzir a proteína recombinante após a adição de indutor no meio (figura 42 E.).

A proteína recombinante produzida na forma insolúvel foi parcialmente purificada (figura 42 F. G.). Essa proteína parcialmente purificada foi inoculada em coelho que, por sua vez, gerou o anticorpo policlonal anti-DPAlphaGlu1.



**Figura 42.** Eletroforeses obtidas durante processo de clonagem e expressão heteróloga de DPAlphaGlu1 sem hélice transmembrana (-m). **A, B** e **D** são eletroforeses dos polinucleotídeos resultantes das reações de amplificação. **E-F** são eletroforeses de proteínas (SDS-PAGE) coradas por Coomassie Blue. **A.** PCR realizado sobre plasmídeo DPAlphaGlu1\_PGEMT, foram utilizadas três temperaturas de pareamento (50°, 55° e 60°C). **B.** PCR de três clones de DPAlphaGlu1-m\_PGEMT. **C.** Digestão da construção DPAlphaGlu1-m\_PGEMT Clone 1 com as enzimas de restrição. **D.** PCR de dois clones de DPAlphaGlu1-m\_PAE. **E.** Proteínas expressas em *E. coli* antes (- IPTG) e horas após a adição de IPTG (+ IPTG). **F.** Produto solúvel e insolúvel da lise de bactérias após a indução da produção da proteína recombinante. **G.** Proteína recombinante parcialmente purificada (segunda raia).

O anticorpo policlonal anti-DPAlphaGlu1 foi capaz de gerar marcações em vesículas associadas ao complexo de Golgi, como pode ser visto na figura 43 A. A marcação do ápice celular restringiu-se ao exterior das microvilosidades com partículas livres ou aderidas à membrana perimicrovilar (figura 43 B). O tratamento com soro pré-imune nas mesmas condições experimentais geraram marcações de padrão próximo ao conseguido com o soro imune, mas em uma densidade menor (figura 43 C).

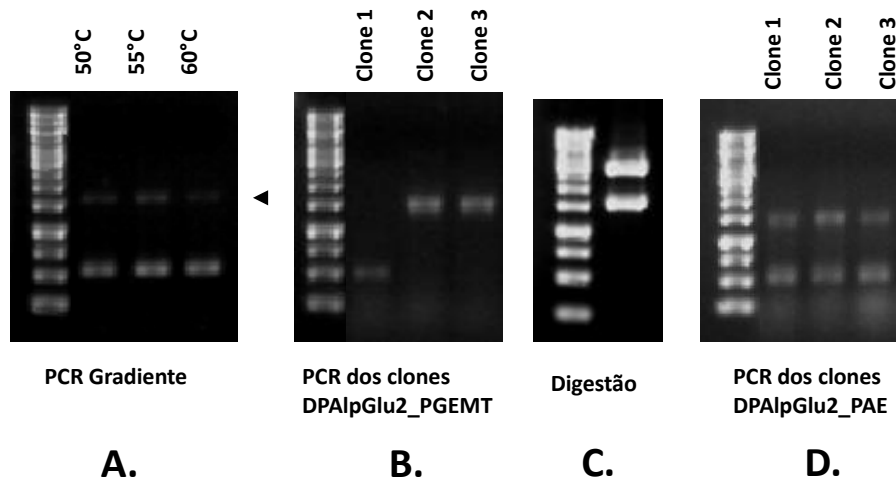


**Figura 43.** Micrografia de enterócito de V1 de *Dysdercus peruvianus* imunomarcado por anti-DPAlphaGlu1. **A.** Região que inclui o retículo endoplasmático rugoso e complexo de Golgi marcado com soro imune. **B.** Região do ápice celular marcado com soro imune. **C.** Região do ápice celular marcado com soro pré-imune (controle negativo). Em todos os casos a diluição do anticorpo foi 1:500. As setas indicam as regiões marcadas. MV, microvilosidade. DV, vesícula de secreção (de dupla membrana). GC, complexo de Golgi. RER, retículo endoplasmático rugoso.

#### 4.3.3.4. DPAlphaGlu2

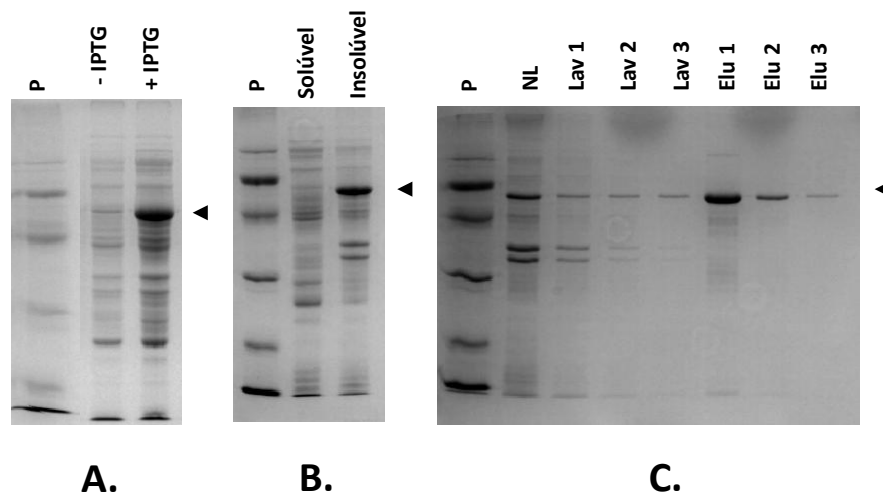
No processo de clonagem da DPAlphaGlu2 foram obtidas duas bandas na primeira reação de amplificação, uma delas com o tamanho esperado e outra com o tamanho bem menor (figura 44 A). A banda indicada pela cabeça de seta foi selecionada para os processos subsequentes de clonagem. Após a ligação do DNA extraído no vetor PGEMT, foram gerados clones da construção DPAlphaGlu2\_PGEMT (figura 44 B). O Clone 2 DPAlphaGlu2\_PGEMT foi utilizado

no processo de digestão (figura 44 C) e o DNA depois de purificado foi inserido no vetor PAE também digerido que, após a transformação de bactérias, deram origem aos clones DPAlphaGlu2\_PAE (figura 44 D).



**Figura 44.** Eletroforeses obtidas no processo de clonagem da DPAlphaGlu2. **A.** PCR realizado sobre cDNA do intestino, foram utilizadas três temperaturas de pareamento (50°, 55° e 60°). A banda indicada pela cabeça de seta foi selecionada para as etapas de digestão (ver **C.**) clonagem no vetor PAE (ver **D.**) **B.** PCR de três clones de DPAlphaGlu2\_PGEMT (clonagem da banda indicada em **A.**). **C.** Digestão da construção DPAlpMan\_PGEMT Clone 2 com as enzimas de restrição XhoI e PstI. **D.** PCR de três clones de DPAlphaGlu2\_PAE.

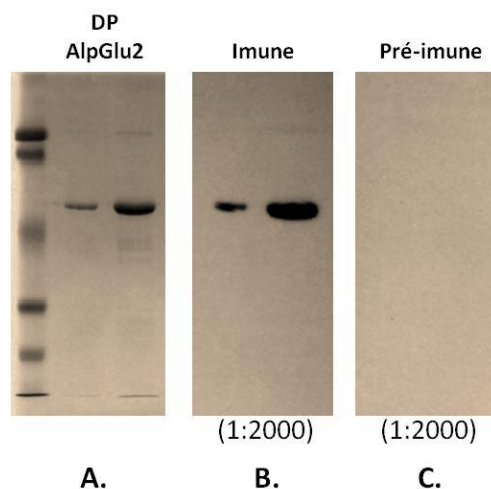
O clone de expressão DPAlphaGlu2\_PAE foi usado na transformação de células competente de *E. coli* da cepa STAR e a produção da proteína recombinante foi induzida pela adição de IPTG (figura 45 A) sendo produzida totalmente na forma insolúvel (figura 45 B). Após a solubilização da proteína recombinante em ureia 8M foi realizado um passo de purificação em coluna de afinidade, conforme pode ser visto na figura 45 C.



**Figura 45.** Eletroforese (SDS-PAGE) das proteínas expressas em *E. coli* (STAR-DE3) durante o processo de expressão e purificação da DPAlphaGlu2 recombinante. **A.** Proteínas expressas em *E. coli* antes (-IPTG) e horas após a adição de IPTG (+IPTG). **B.** Produto solúvel e insolúvel da lise de bactérias após a indução da produção da proteína recombinante. **C.** Eletroforese do processo de purificação da proteína recombinante sendo: NL. proteínas não ligadas à resina de purificação. Lav1-3. proteínas liberadas em lavagens consecutivas. Elu1-3. proteínas liberadas com a adição de 500mM de imidazol ao tampão de lavagem. Em todos os casos a cabeça de seta indica a proteína recombinante.

As alfa-glicosidases são descritas na digestão em vários insetos, sendo que sua atividade pode ser solúvel ou ligada à membrana. No caso da DPAlphaGlu2 estaríamos lidando com a atividade solúvel, já que não foi possível prever alças transmembranares para a enzima.

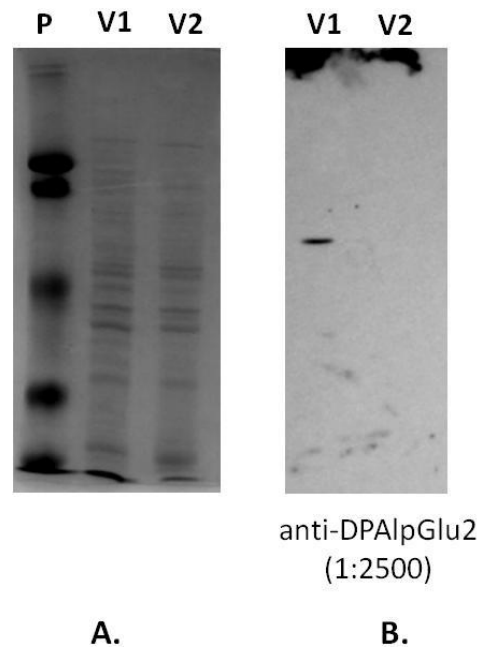
O anticorpo policlonal anti-DPAlphaGlu2 produzido em coelho foi capaz de reconhecer a proteína recombinante sendo que o soro pré-imune não reconhece a proteína recombinante nas condições experimentais apresentadas (figura 46).



**Figura 46.** Eletroforese (SDS-PAGE) e *western blot* da DPAlphaGlu2 recombinante. **A.** Eletroforese da DPAlphaGlu2 recombinante semi-purificada (segunda e terceira raia), corada por Coomassie Blue. **B.** *Western blot* com o soro imune anti-DPAlphaGlu2 com o título de 1:2000. **C.** *Western blot* com soro pré-imune com título de 1:2000.

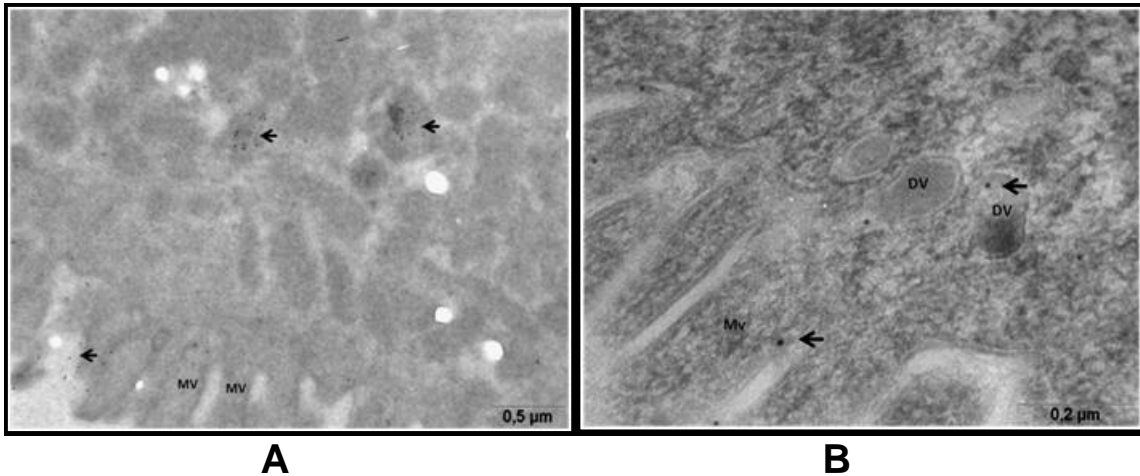


DPAlphaGlu2 não apresenta domínio transmembranar como a DPAlphaGlu1, logo, podemos prever ser uma enzima solúvel. Foi possível confirmar a DPAlphaGlu2 como proteína solúvel através da marcação de *western blot* usando apenas proteínas solúveis extraídas tanto de V1 quanto V2 (figura 47). A marcação de uma banda no tamanho esperado é visível apenas em V1, mas vale a pena ressaltar um fenômeno recorrente em *western blots* feitos para essa enzima. Uma forte marcação aparece na parte superior da membrana, sugerindo que grande parte da proteína não entra completamente no gel durante a eletroforese. Esse intrigante fenômeno ocorre também com a enzima recombinante, mas em menor grau.

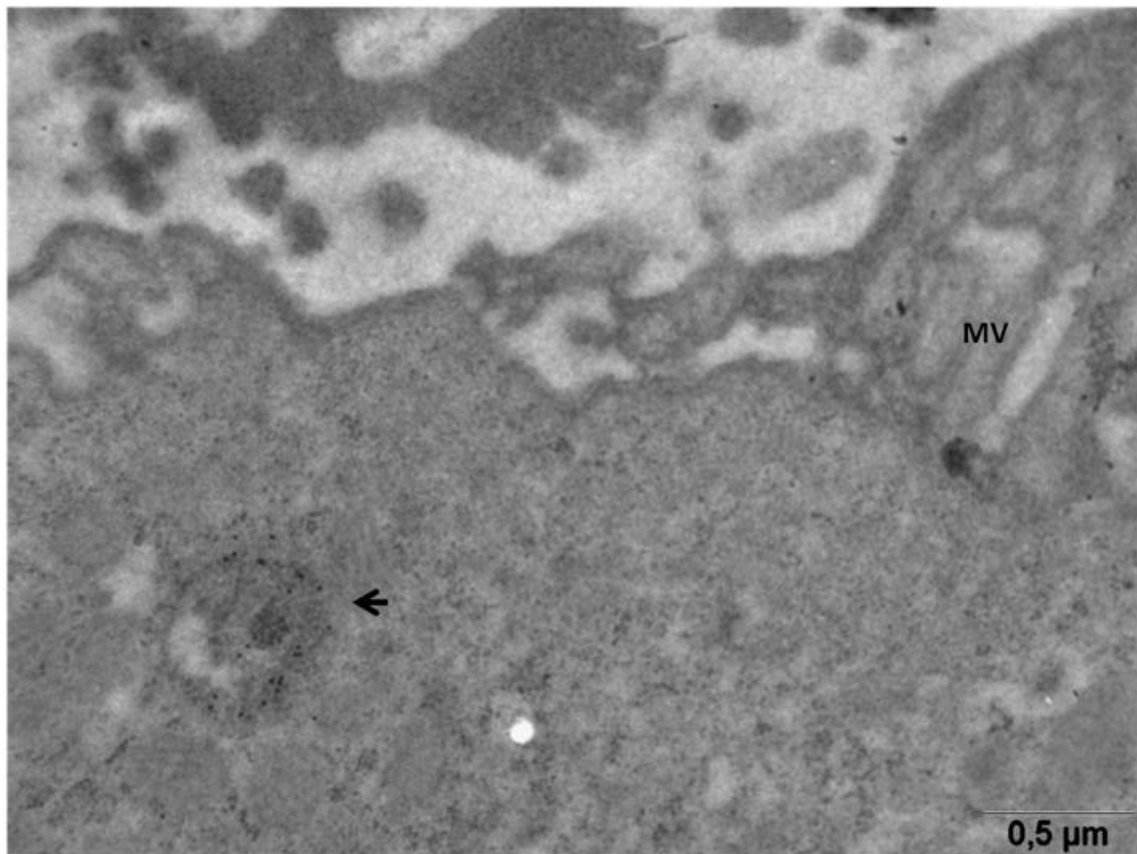


**Figura 47.** Reconhecimento da DPAlphaGlu2 de *Dysdercus peruvianus* por *western blotting*. **A.** SDS-PAGE das proteínas solúveis de V1 e V2 coradas por Coomassie Blue. **B.** *Western blot* com anticorpos policlonais anti-DPAlphaGlu2 com título de 1:2500.

O reconhecimento de uma única banda dentro da massa molecular esperada permite a continuidade dos experimentos passando para experimentos de imunocitocalização.



**Figura 48.** Micrografias de enterócitos de V1 de *D. peruvianus* imunomarcados por anti-DPAlphaGlu2. Setas indicam as regiões marcadas. MV, microvilosidade. A e B são experimentos idênticos variando apenas na qualidade do contraste.



**Figura 49.** Micrografia de enterócitos de V2 de *D. peruvianus* imunomarcados por anti-DPAlphaGlu2. Setas indicam as regiões marcadas. MV, microvilosidade.

Considerando os aspectos gerais da marcação, podemos dizer que compartimentos intracelulares e a luz intestinal foram marcados.

O anti-DPAlphaGlu2 marca compartimento intracelulares com conteúdo heterogêneo, composto por regiões mais e menos eletrodensas (figuras 48 e 49).

Tais compartimentos se assemelham às vesículas de secreção descritas para espécie.

#### 4.3.3.5. DPAPN

A aminopeptidase é uma enzima relacionada à digestão intermediária de proteínas e sua atividade se dá no espaço entre a membrana microvilar e a membrana perimicrovilar.

Foram encontradas diversas sequencias que codificam aminopeptidases de diferentes famílias gênicas (ver tabela 9). Para basear a escolha da sequencia que codifica a aminopeptidase intestinal quanto sua classificação, foi levado em consideração o trabalho realizado com a aminopeptidase de membrana do Hemiptero *A. pisum* (Cristofolletti *et al*, 2006), no qual a enzima purificada e sequenciada do intestino do inseto foi uma aminopeptidase N (família M1). O passo seguinte foi estimar o nível de expressão da enzima através do número de *reads* que formaram as sequencias analisadas.

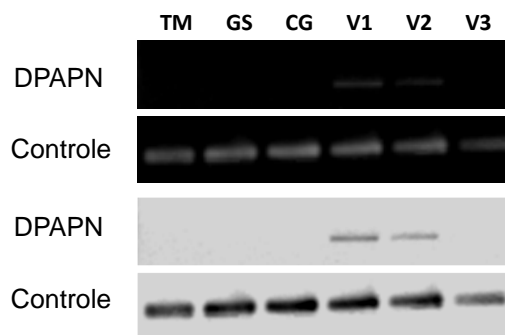
**Tabela 9.** Aminopetidases obtidas no pirosequenciamento, anotação e número de *reads*.

Selecionada	Anotação	Família	Nº de reads	Completa
	leucil aminopeptidase	peptidase_M17	823	sim
>	aminopeptidase N	peptidase_M1	5145	sim
	aminopeptidase N	peptidase_M1	72	sim
	leucil aminopeptidase	peptidase_M17	59	sim
	gluzincin	peptidase_M1	16	-
	leucil aminopeptidase	peptidase_M17	34	sim
	metionina aminopeptidase	APP_MetAP	19	sim
	metionina aminopeptidase	APP_MetAP	39	sim
	aminopeptidase N	peptidase_M1	11	-
	gluzincin	peptidase_M1	18	-
	metionina aminopeptidase	APP_MetAP	6	-
	gluzincin	peptidase_M1	6	-
	metionina aminopeptidase	APP_MetAP	10	-

Na tabela 9 uma aminopeptidase N completa apresenta, aproximadamente, 100 vezes mais *reads* que as outras aminopeptidases N. Essa grande quantidade de *reads* deve refletir sua expressão, reforçando sua importância na digestão. A sequencia foi selecionada para estudos mais profundos das suas características e

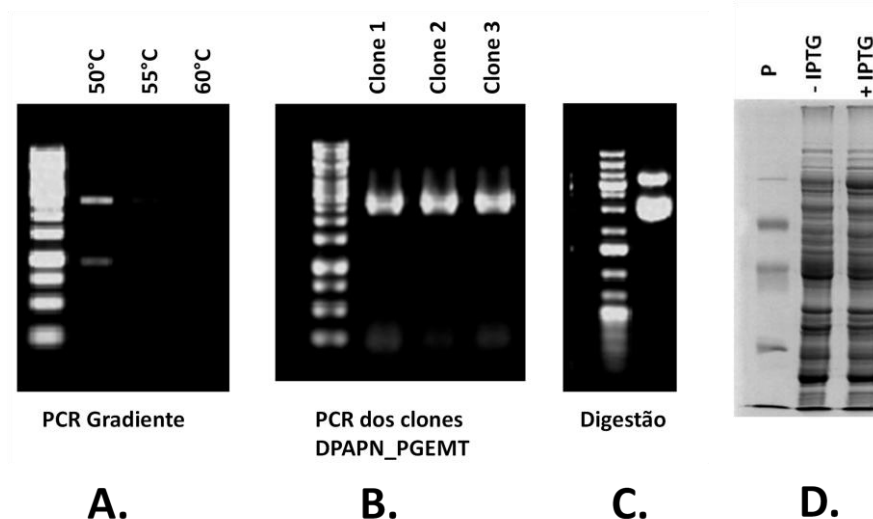
seu padrão de expressão em outros tecidos e será aqui denominada apenas de DPAPN.

A expressão da aminopetidase (DPAPN) se dá nos ventrículos V1 e V2 (figura 50). Sua expressão restrita a tecidos do intestino médio indica ser uma enzima relacionada aos processos digestivos. Ter seu mRNA expresso apenas em tecidos digestivos reforça sua escolha como enzima responsável pela atividade de aminopeptidase do intestino de *D. peruvianus*. DPAPN não apresentaria sítio de modificação de âncora de GPI segundo predição (big-PI Predictor).



**Figura 50.** Expressão do RNAm que codifica a DPAPN foi avaliada por RT-PCR. TM, túbulos de Malpighi; GS, glândula salivar; CG, corpo gorduroso; V1, primeiro ventrículo do intestino médio; V2, segundo ventrículo do intestino médio; V3, terceiro ventrículo do intestino médio. A mesma figura é apresentada de duas formas para melhor visualização.

A clonagem da DPAPN usou como DNA molde o cDNA obtido a partir do RNA extraído do intestino médio de *D. peruvianus*. Após a reação de amplificação foi possível notar duas bandas (figura 51 A), uma do tamanho esperado e outra bem menor. A DNA da banda de tamanho esperado, após purificação, foi inserida no vetor de clonagem PGENT (figura 51 B) e um dos clones obtidos foi submetido à digestão com enzimas de restrição (figura 51 C).



**Figura 51.** Eletroforeses obtidas no processo de clonagem da DPAPN. **A.** PCR realizado sobre cDNA do intestino foram utilizadas três temperaturas de pareamento (50°, 55° e 60°). **B.** PCR de três clones de DPAPN\_PGEMT. **C.** Digestão da construção DPAPN\_PGEMT Clone 1 com as enzimas de restrição BamHI e EcoRI. **D.** SDS-PAGE de uma tentativa de expressão recombinante da DPAPN. Proteínas expressas em *E. coli* antes (- IPTG) e horas após a adição de IPTG (+ IPTG)

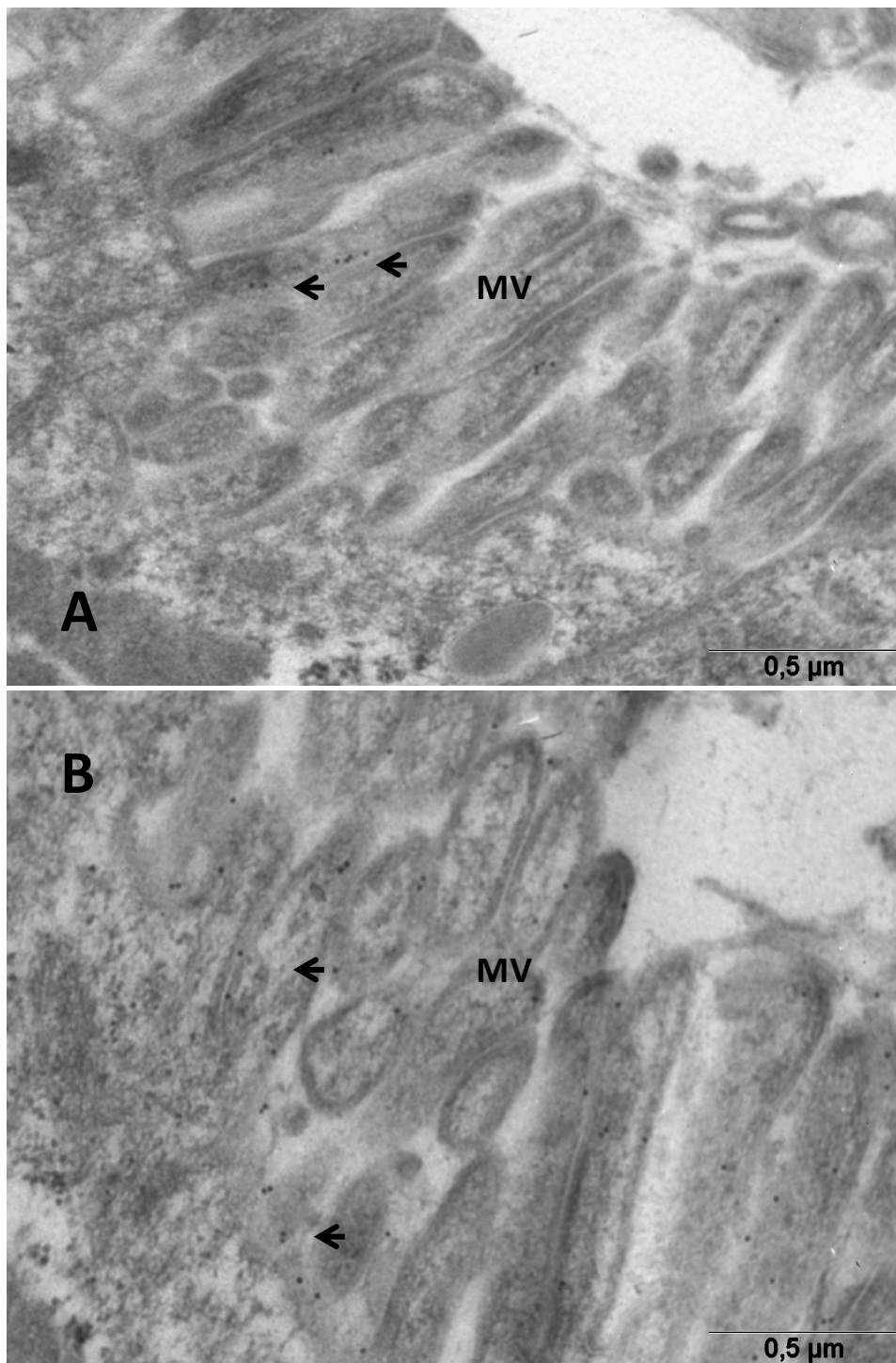
Foi possível clonar a sequência da DPAPN em um plasmídeo de expressão gerando a construção DPAPN\_PAE. A nova construção foi sequenciada e não apresentou qualquer mutação que inviabilizasse seu uso. Contudo, diversas tentativas de expressão em *E. coli* foram feitas sem sucesso na expressão heteróloga da enzima. As tentativas incluíram variações na temperatura de crescimento das bactérias em meio líquido, uso de diferentes cepas e variação na concentração do indutor (IPTG). A figura 51 D é um resultado típico para esse experimento.

Uma abordagem experimental foi adotada como forma de superar a dificuldade na produção desta proteína recombinante para a produção de anticorpos. A produção de peptídeos sintéticos foi usada como fonte de epítopos na geração de anticorpos policlonais em coelho.

A estratégia para a produção do anticorpo policlonal anti-DPAPN foi através da imunização de coelho com o peptídeo sintético baseado na sequência proteica encontrada. O peptídeo KRDFINVPSDITNDC foi sintetizado e conjugado com carreador KLH (hemocianina de molusco) para o processo de imunização do coelho.

O soro imune gerado foi purificado com coluna de afinidade com o peptídeo sintético imobilizado. O anti-DPAPN, como previsto, é capaz de marcar as laterais das microvilosidades, no espaço entre membranas (figura 52 A). Há, contudo, marcação com padrão batante difuso obtida com o soro pré-imune que inclui o

padrão experimental (figura 52 B). Com esse resultado não é possível afirmar que a imunização com o peptídeo sintético foi o único fator responsável pela especificidade da marcação do anti-DPAPN, podendo haver anticorpos prévios no coelho capazes de reconhecer a DPAPN.



**Figura 52.** Micrografia de enterócito de V1 de *Dysdercus peruvianus* imunomarcado por anti-DPAPN. **A.** Região do ápice celular marcado com soro imune. **B.** Região do ápice celular marcado com soro pré-imune (controle negativo). Em ambos os casos a diluição do anticorpo foi 1:1000. As setas indicam as regiões marcadas. MV, microvilosidade.

#### 4.3.3.6. Compilação dos anticorpos produzidos

A produção da proteína recombinante, sua purificação e validação do anticorpo produzido foram detalhadas para cada marcador, assim como foi feito para as catepsinas L. Como já mencionado, dificuldades experimentais não permitiram a validação da especificidade dos anticorpos em todas as condições testadas. Evidências claras de marcação inespecíficas foram fundamentais para a não utilização do anti-DPCAL1 e, por sua vez, as evidências de especificidade levaram a diante estudos com as demais. A tabela 10 resume esses esforços.

**Tabela 10.** Resumo da produção e experimentos realizados com diferentes anticorpos.

Anticorpo	Antígeno	Tratamento do soro	Western blot (recombinante)*	Western blot (homog.)**	Micrografia (reconhecimento diferencial)***
Anti-DPCAL1	proteína recombinante	precipit. por (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	sim	não	não testado
Anti-DPCAL10	proteína recombinante	precipit. por (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	sim	?	sim
Anti-DPAlphaGlu1	proteína recombinante	precipit. por (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	não testado	?	sim
Anti-DPAlphaGlu2	proteína recombinante	precipit. por (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	sim	sim	sim
Anti-DPAlphaMan	proteína recombinante	precipit. por (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	sim	?	sim
Anti-DPAPN	peptídeo sintético	precipit. por (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	NA	?	sim

\* indica a capacidade de o anticorpo reconhecer a proteína recombinante usada como antígeno. \*\* indica a capacidade de o anticorpo reconhecer uma proteína única em uma amostra complexa (como o homogeneizado intestinal). \*\*\* indica a capacidade de o anticorpo gerar marcações em cortes histológicos quando comparado com o seu pré-imune (nas mesmas concentrações). NA, não se aplica. ?, resultados inconclusivos.

#### 4.3.4. Proteômica do conteúdo solúvel de V1

Foi isolado o conteúdo solúvel da primeira fração intestinal de *D. peruvianus* (V1) e com esse material rico em proteínas secretadas foi realizado experimentos de proteômica. Saber quais proteínas estão presentes na luz intestinal poderia confirmar as previsões feitas para a secreção de enzimas tipicamente lisossômicas na participação da digestão do alimento.

Foi possível identificar no conteúdo de V1 seis catepsinas D, DpCatD1, 4, 5, 6, 7 e 8. A análise de expressão destes genes indicava estarem presentes preferencialmente em V1, mostrando que essa fração intestinal teria alguma função na digestão de proteínas além da porção V2, que tem grande atividade de cisteíno proteinase (Silva e Terra, 1994). Devemos considerar também que DpCatD1, 4, 5, 6

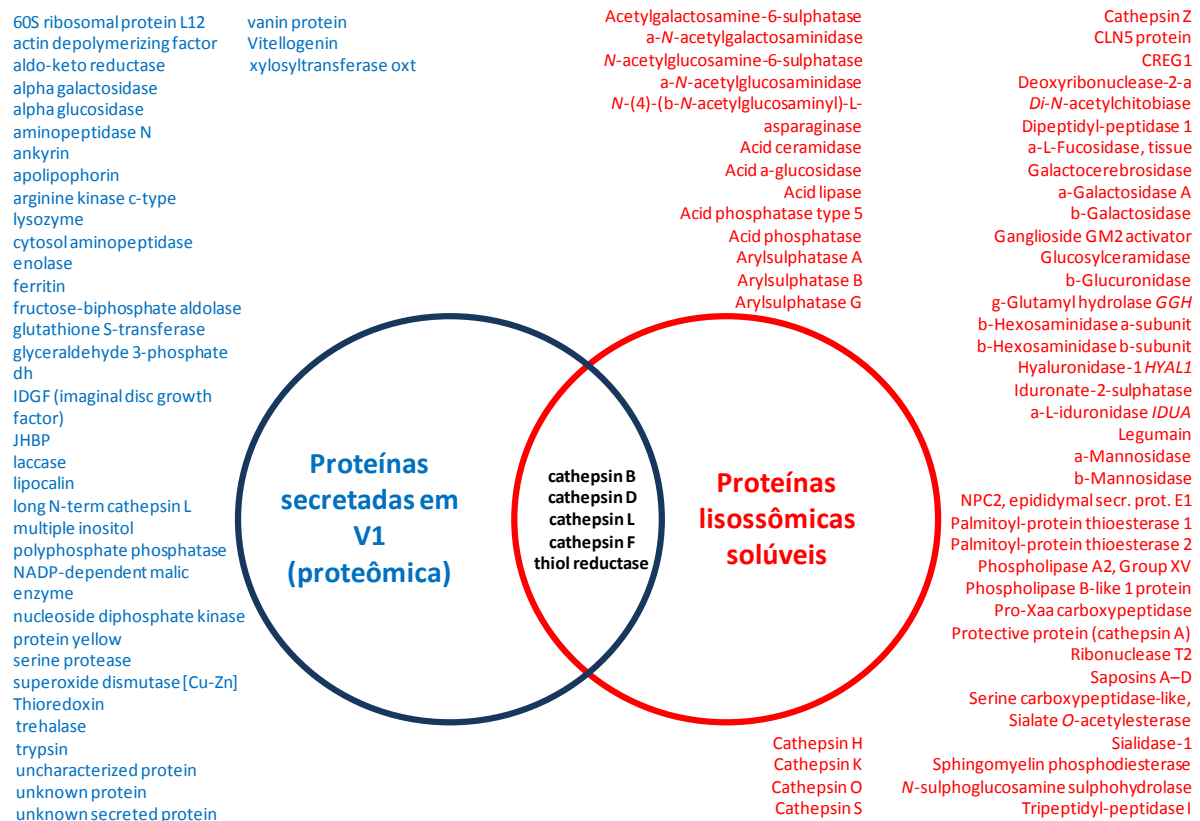
e 7 apresentam os resíduos catalíticos conservados bem como os sinais de N-glicosilação. Segundo Balczun e colaboradores (2012) a catepsina D com a região consenso de enzima lisossômica poderia ser isolada do lúmen intestinal. Ainda não foi possível identificar a DpCatD10, enzima provavelmente lisossômica, no lúmen intestinal, mas sabemos que sua expressão é extremamente baixa no tecido intestinal e, portanto, sua detecção no material luminal total é menos provável.

Foi possível identificar cinco das dez catepsinas L em estudo. São elas DpCaL2, 3, 4, 5 e 8, sendo que das catepsinas L clássicas (DpCaL3, 4, 5 e 8) a DpCaL5 apresentou a maior contagem de espectros, confirmando os dados de expressão que indicavam a expressão gênica preferencial em V1 (ver figura 13). Com apenas um peptídeo detectado, o DpCaL10 foi também identificada. Essa detecção é importante pois, mesmo com baixa expressão gênica, essa proteína foi encontrada no lúmen de V1, indicando sua secreção.

Três glicosidases foram encontradas com grande contagem de espectros (que é um indício de sua presença na amostra). As duas alfa-glicosidases estudadas (DpAlphaGlu1 e DpAlphaGlu2) estão presentes na amostra, o que indica que a DpAlphaGlu1 não se mantém restrita a membrana perimicrovilar, sendo parcialmente solubilizada. A terceira alfa glicosidase, não presente nos estudos prévios, pertence a família 31 das glicosídeo hidrolases (pfam01055). Esta família inclui enzimas que são alfa galactosidases (ou similar), sugerindo que seria esta a enzima responsável pela hidrólise da rafinose presente na semente de algodão.

A evolução dos Heteroptera está associada à volta à alimentação de polímeros de insetos que ancestralmente se adaptaram a sugar seiva. Devido à perda das enzimas de digestão inicial nesses ancestrais, enzimas derivadas de enzimas lisossômicas e que são agora encontradas na luz intestinal foram usadas na digestão (Terra e Ferreira, 2012). A presença de uma grande quantidade de enzimas lisossômicas na luz do intestino gera uma dúvida quanto à sua origem – seriam as enzimas secretadas isoladamente ou o lisossômo descarregaria todo seu conteúdo na luz intestinal? Para ter mais indícios da participação de outras enzimas lisossômicas na digestão intestinal luminal foi feita uma comparação das proteínas solúveis sabidamente lisossômicas (Schröder et al., 2010) com as proteínas encontradas no proteoma da material solúvel do lúmen de V1 (figura 53 e tabela suplementar 2).





**Figura 53.** Proteínas encontradas no lúmen de V1 em relação às proteínas descritas para o lisossomo. Anotações em azul (à esquerda) representam as proteínas do proteoma que não apresentam relação com proteínas lisossômicas. Anotações em vermelho (à direita) correspondem às proteínas solúveis descritas para o lisossomo de acordo com a revisão de Schröder et al. (2010). Proteínas encontradas em V1 previstas para ocorrerem em lisossomos foram escritas em preto na interseção dos círculos.

Essa abordagem nos permite acessar diretamente quais enzimas passaram pelo processo de secreção e são agora encontradas no lúmen intestinal. Caso os lisossomos tivessem seu conteúdo secretado, como os osteoclastos de vertebrados, esperaríamos uma grande correspondência das proteínas identificadas em V1 com as proteínas lisossômicas. Dentre as proteínas identificadas no proteoma da fração solúvel de V1, apenas as catepsinas B, D, L e F e uma tiolredutase correspondem a proteínas tipicamente lisossômicas, indicando haver apenas secreção parcial do conteúdo lisossômico.

## 5. Discussão

### 5.1. Fisiologia molecular da digestão em *Dysdercus peruvianus*

#### 5.1.1. Digestão e absorção no intestino médio

Foi detectada a presença de absorção de glicose em V1 e absorção de aminoácidos em V2 (Silva e Terra, 1994; Bifano et al., 2010). Com esse dado é possível prever que os transportadores correspondentes seriam expressos preferencialmente nos respectivos compartimentos. Com a grande quantidade de transportadores de glicose obtidos no transcriptoma intestinal não é possível sugerir o transportador responsáveis pela absorção de glicose vista por Bifano *et al.* (2010). Como forma de delimitar melhor os possíveis transportadores de glicose, foi realizado um experimento que compara transcriptomas das porções V1 e V2, buscando transportadores diferencialmente expressos nos tecidos. As análises levam a identificação de potenciais transportadores de glicose com identidade diferente do já relatado na literatura (Bifano et al., 2010). A expressão diferencial em V2 de um gene que transcreve um transportador de glicose nos leva a pensar que haveria absorção de glicose (ou outro monossacarídeo) também nessa porção intestinal. Agora, porém, temos mais evidências de que o principal transportador de glicose em V1 seria realizado pelo gene representado aqui pelo contig\_834, e não pelo contig\_1905 (código de acesso ACY03364.1) como relatado por Bifano et al. (2010). Isso, além de gerar um melhor panorama dos genes expressos, indica um bom candidato à experimentos de inibição de expressão gênica (RNAi) para confirmar sua atuação na absorção de glicose em abordagens futuras.

Segundo Silva e Terra (1994), a cisteino proteinase é a principal proteinase de *D. peruvianus*, atuando como enzima solúvel no lúmen intestinal, mais especificamente no lúmen de V2 (uma fração do intestino médio). Pesquisas recentes revelam que, dentre as cisteino proteinases, as catepsinas L são as únicas quantitativamente importantes no intestino médio de insetos (Terra e Ferreira, 2012), havendo diversos relatos de expressão de catepsina L no intestino médio de Heteroptera hematófagos (Lopez-Ordoñez, *et al.*, 2001; Kollien, *et al.*, 2004; Waniek, *et al.*, 2012). Como forma de compreender melhor a presença desse tipo de enzima

no lúmen intestinal, foram selecionadas dentre as vinte e cinco sequencias de cisteino peptidases dez que correspondem às catepsinas L.

É importante salientar que as sequencias DPCAL1-4 já haviam sido descritas em trabalhos prévios do laboratório, mas ainda não publicados. Nestes trabalhos prévios era atribuído a DPCAL2 função lisossômica e a DPCAL1, 3 e 4, função como enzima digestiva (Bifano *et al.*, não publicado).

Os dados de expressão indicam que DPCAL2 participaria de uma função comum aos tecidos digestivos, os túbulos de Malpighi e ao corpo gorduroso, não sendo expresso nas glândulas salivares. A ausência de expressão da DPCAL2 nas glândula salivares e a descoberta da DPCAL10, que é expressa em todos os tecidos estudados, tornou mais claro que DPCAL2 não seria uma proteína de tão ampla distribuição e, portanto, talvez não fosse um marcador de lisossomo em todos os tecidos. De qualquer forma, a grande porção N-terminal dessa proteína, bem como seu padrão de expressão bastante peculiar, torna essa proteína uma candidata para estudos mais gerais para fazer inferências sobre sua função. Dentro dos insetos, há cisteino proteinases com o mesmo domínio N-terminal da DPCAL2 com a anotação “26/29-kDa proteinase”. A comparação de sequencias aponta na mesma direção já que, somado ao fato de ter o N-terminal longo, a DPCAL2 pertence ao grupo das “26/29-kDa proteinase” e não das outras catepsinas L estudados.

Inicialmente, foi descoberta uma atividade proteolítica que distinguia proteínas próprias das exógenas quando injetadas na cavidade abdominal da larva da mosca *Sarcophaga peregrina* (Suzuki e Natori, 1985). A proteína responsável pela atividade foi purificada e mostrou ser um heterodímero com subunidades de 26 e 29 kDa e, posteriormente, foi verificado que ambas as subunidades eram transcritas pelo mesmo gene (Saito et al., 1992; Fujimoto et al., 1999). Esse gene está presente em outros insetos e, mesmo com a identidade muito próxima, não é sensível aos inibidores específicos de catepsina L (Fujimoto et al., 1999). Os autores não relatam a presença do domínio conservado, que está presente na região da subunidade 26 kDa. Originalmente a atividade da 26/29-kDa proteinase foi descrita em hemócitos da larva de *Sarcophaga peregrina* e não foi encontrada atividade no corpo gorduroso (Suzuki e Natori, 1985). O estudo de expressão da DPCAL2 não foi feito nos hemócitos, mas sua expressão foi detectada em corpo gorduroso o que não seria esperado pela comparação com *S. peregrina*. Há ainda a possibilidade de

o tecido estar contaminado por hemócitos e mostrar um falso-positivo. De qualquer forma, sua expressão é marcante também nos túbulos de Malpighi e intestino médio. Sua atuação deve incluir, como descrito para as outras 26/29-kDa proteinases, a defesa imunológica.

Esses dados sugerem que DPCAL1 e DPCALs de 3 a 9 são enzimas que participam do processo digestivo, pois apenas o intestino médio expressa os genes que codificam essas enzimas. A expressão preferencial das DPCALs que parecem ser responsáveis pela digestão no lúmen intestinal reforça a proposta de Silva e Terra (1994) de que as cisteino proteinases seriam secretadas principalmente em V2. Contudo, o padrão de expressão da DPCAL5 se torna bastante intrigante, visto ter uma grande expressão e ser expressa preferencialmente em V1, porção com baixa atividade de cisteino proteínase em *D. peruvianus* verificada no trabalho de Silva e Terra (1994). A menor expressão da DPCAL10 no intestino difere dos dados obtidos no transcriptoma de *Rhodnius prolixus*. Nele, a possível catepsina L lisossômica (RP-428) apresenta superexpressão no intestino médio, mas muito próximo do corte estatístico (Ribeiro et al., 2014). A ativação da DPCAL1 recombinante demonstra que as enzimas classificadas como digestivas são capazes, sim, de atuar na hidrólise do alimento. A DPCAL1 se autoativa em pH baixo como pode ser vista a ativação da pCAL3 de *Tenebrio molitor* (Beton et al., 2012).

A catepsina L do pulgão *Acyrtosiphon pisum*, também um Hemiptera, é uma enzima presa à membrana perimicrovilar, o que previne a excreção das enzimas (Cristofolletti, et al., 2003). Como pelo menos 20% da atividade de cisteino proteínase do conteúdo intestinal está associada ao material insolúvel (membranas) em *D. peruvianus* (Silva e Terra, 1994), foi pesquisada a presença de hélices transmembranas e sítios de modificação por âncora de GPI, mas não houve predição em nenhum dos casos, sugerindo que caso haja alguma forma de ligação à membrana, a ligação se dê por outros mecanismos.

A marcante presença das catepsinas L em V1 parece contrapor os dados de que a V2 seria a porção intestinal responsável pela digestão de proteínas. Contudo, um olhar cauteloso sobre o trabalho de Silva e Terra (1994) indica que cerca de 25% da atividade de cisteino proteínase está restrita à V1 (soma da atividade presente nas células com a atividade presente no conteúdo). Logo, a presença destas proteínases no conteúdo luminal de V1 complementa o dado da literatura, sugerindo

que V1 teria, também, participação significativa na digestão de proteínas e que V1 contribuiria com enzimas na digestão de proteínas em V2, já que o alimento segue de V1 para V2.

A presença de catepsina D atuando na digestão em outros Heteroptera já foi verificada (Terra et al., 1988; Defferrari et al., 2011; Balczun *et al.*, 2012), mas parece pequena a sua participação na digestão em *D. peruvianus* (Silva e Terra, 1994). Contudo, a presença marcante de transcritos dos genes de catepsina D nos transcriptomas, sua regulação pela alimentação e a presença das enzimas no lúmen intestinal parece indicar alguma participação no processo digestivo.

Considerando a variação espacial na expressão dado pelas análises de expressão diferencial entre V1 e V2 e pelas expressões gênicas das proteinases, podemos dizer que há especialização na produção de proteinases de acordo com a porção do intestino médio. V1 seria responsável pela produção preferencial de catepsinas D e V2 pela secreção de catepsinas L. Como sabemos, essas duas classes de proteinases têm especificidades diferentes, podendo atuar complementarmente na digestão de proteínas. Somamos a essa discussão a participação das lipases que são preferencialmente expressas em V2. A presença de lipases é fundamental para esse inseto, já que lipídeos são o principal componente da semente de algodão (Silva e Terra, 1994). Como o conteúdo de V1 é descarregado em V2 e a maior residência do alimento é em V2 (72 h para V2 contra 31 h para V1, (Silva e Terra, 1994)), podemos considerar que as enzimas secretadas em V1 atuam na digestão em V2.

A DPAPN descrita aqui não tem a mesma sequência do gene relatado por Costa et al. (2011). Além da diferença entre as sequências, o gene do citado trabalho apresenta expressão em diversos tecidos, incluindo tecidos como corpo gorduroso e túbulos de Malpighi. Não foi possível encontrar a sequência do trabalho citado dentre as apresentadas aqui e nem mesmo no banco de dados brutos (*output* do montador Newbler), que implica em sua baixa representatividade no intestino médio. A expressão da DPAPN foi verificada em V1 e V2, mas não em outros tecidos estudados, indicando a participação da DPAPN na digestão. Esse resultado indica que o gene estudado deve ser o responsável pela transcrição da enzima cuja atividade aparece distribuída por todos os ventrículos, com V1 respondendo por 60% da atividade total (Silva e Terra, 1994). Trabalhos com *A. pisum* mostraram que toda atividade de aminopeptidase está presente na fração de membranas (insolúvel) e

que a ligação dela na membrana se daria por âncora de GPI (Cristofolletti *et al.*, 2003; Cristofolletti *et al.*, 2006). Em *D. peruvianus* e *R. prolixus*, porém, a atividade de aminopeptidase estaria na forma solúvel no espaço entre a membrana microvilar e a perimicrovilar (Silva *et al.*, 1996; Ferreira *et al.*, 1988) e, como esperado, não há predição de âncora de GPI. A marcação da aminopeptidase no espaço entre membranas reforça que a proteína estaria, sim, presente no espaço perimicrovilar e, analisando os dados já obtidos conjuntamente com os estudos citológicos, a DPAPN deve corresponder à enzima isolada e caracterizada por Costa *et al.* (2010), embora sua sequência difira do relatado por eles.

DPAphaGlu1, DPAphaGlu2 e DPMan são expressas apenas no intestino médio, especialmente em V1 e V2, indicando participarem da digestão. O resultado está de acordo com os dados enzimáticos publicados em que a atividade de carboidrases seria importante em V1 e presente também em V2 (Silva e Terra, 1994; Silva *et al.*, 1996). Há em *D. peruvianus* uma importante atividade de alfa glicosidade presa a membrana perimicrovilar em V1 que foi isolada e caracterizada (Silva e Terra, 1994; Silva e Terra, 1995). O conjunto de dados indicam ser a DPAphaGlu1 a enzima responsável pela atividade de membrana. Já a DPAphaGlu2 seria a principal responsável pela atividade de alfa-glicosidase das frações solúveis presente no conteúdo do intestino médio de *D. peruvianus*. Com base em Silva e Terra (1994), essa atividade solúvel corresponde a 76% da atividade de alfa-glicosidase do lúmen intestinal de V1. Os dados de proteômica indicam a presença da DPAphaGlu1 no material solúvel. O mecanismo de solubilização da enzima DpAlphaGlu1 ainda não foi definido, mas pode estar relacionado à proteólise parcial da enzima, com a liberação da fração solúvel da enzima de sua fração membranar. Como previsto, DpAlphaGlu2 foi a alfa glicosidase com maior contagem de espectros e confirmou os dados de imunomarcção (nos quais foi demonstrada sua marcação na luz intestinal).

A atividade hidrolítica sobre a rafinose presente na semente de algodão se dá por duas fontes de enzimas: uma é a alfa galactosidase presente na própria semente e a segunda é a alfa galactosidase produzida pelo epitélio intestinal do *D. peruvianus* (Silva e Terra, 1997). A alfa galactosidase encontrada no conteúdo solúvel de V1 pertence a família 31 das glicosidases. Esse dado causou surpresa, já que outro gene candidato (pertencente a família 27, pfam02065) era o foco de nossa atuação para explicar a atividade de alfa galactosidase em V1. A presença de

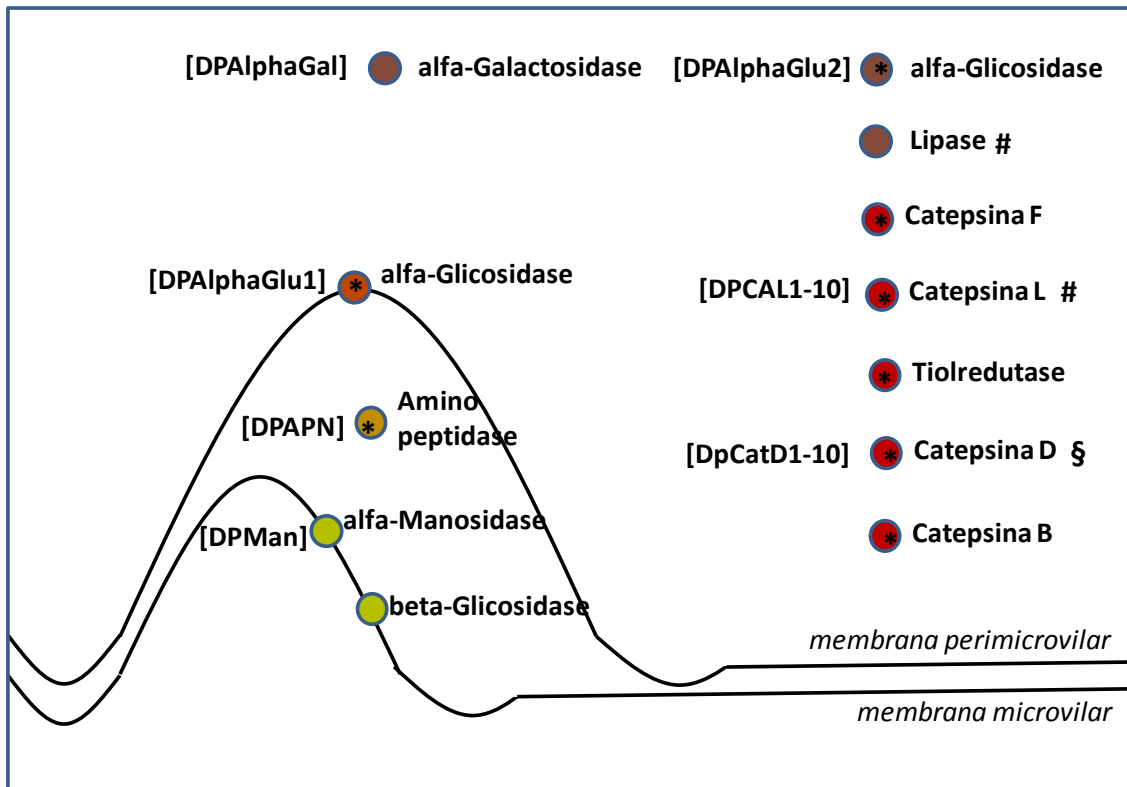
uma alfa galactosidase é fundamental para a digestão do principal carboidrato da semente de algodão. A atuação da alfa galactosidase sobre rafinose libera galactose e sacarose que, por sua vez, serve de substrato para as alfa glicosidases solúvel e de membrana, liberando glicose e frutose no lúmen de V1 e V2 (Silva e Terra, 1997).

### 5.1.2. Compreensão global da digestão

O alimento ingerido por *Dysdercus peruvianus* é uma suspensão do conteúdo da semente de algodão (e pode incluir a semente de outras Malvaceae). Essa suspensão com algum material já solubilizado é encaminhado para a primeira porção do intestino médio, V1 (Silva e Terra, 1994). Em V1 há a ação de três glicosidases principais: uma alfa-glicosidase solúvel (DPA $\alpha$ Glu2), uma alfa-glicosidase presa a membrana perimicrovilar que pode ser solubilizada (DPA $\alpha$ Glu1) e uma alfa-galactosidase solúvel (DPA $\alpha$ Gal) (Silva e Terra, 1994; Silva et al., 1996). Há atuação também de uma beta-glicosidase e um alfa-manosidase (DPMan) presas à membrana microvilar. A atuação dessas enzimas sobre os substratos da semente como a rafinose gera os monossacarídeos galactose, glicose e frutose, que são absorvidos em V1 por transportadores de glicose do tipo GLUT, com o transporte conjunto com água (Silva e Terra, 1994; Bifano et al., 2010). Há secreção de catepsina D em V1 (DpCatD1, 4-8) e catepsina L (DPCAL2-5, 8 e 10), iniciando a digestão de proteínas. Dentre as catepsinas L secretadas em V1 há uma que também participa na defesa imunológica (DPCAL2) e outra que atua também nos lisossomos em outros tecidos (DPCAL10). A secreção das catepsinas L em V1 correspondem a 25% da atividade do total do intestino (Silva e Terra, 1994). A secreção de proteinases inclui, também, catepsinas B e F além de uma tiolredutase. Em V1, ainda, há atuação de uma aminopeptidase (DPAPN) do espaço entre membranas, que atua também no material luminal (Silva e Terra, 1994; Silva et al., 1996; Costa et al., 2011).

As enzimas secretadas em V1 seguem para V2 junto com o bolo alimentar, onde ocorrerá a principal parte da digestão de proteínas com a atuação principal das catepsinas L. Nesta porção do intestino ocorre a principal absorção de aminoácidos (Silva e Terra, 1994). Há em V2 a secreção de glicosidases que são responsáveis pelo final da digestão dos carboidratos, há também a atuação de lipases que são fundamentais na digestão do principal constituinte da semente. O conteúdo de V2 é

passado para V3, onde há atuação de bactérias simbiotes e, depois de uma longa permanência (70-90 h) nesses dois compartimentos, o conteúdo luminal passa rapidamente por V4 (Silva e Terra, 1994). O modelo de digestão em *D. peruvianus* e a contribuição do presente trabalho na sua compreensão foi resumido na figura 54.



**Figura 54.** Modelo da distribuição das enzimas digestivas em *Dysdercus peruvianus*. Conjunto de enzimas que participam da digestão baseado na literatura (ver texto) e as contribuições deste trabalho. Os círculos coloridos representam a proteína e sua localização. À direita do círculo está a atividade (verificada ou predita) para esse proteína. Em colchetes, à esquerda do círculo colorido, o nome da proteína estudada (que apresenta um ou mais genes que transcrevem tal enzima). \*, enzima encontrada no conteúdo solúvel de V1 (ou seja, é secretada). #, enzima cujo gene é preferencialmente expresso em V2. §, enzima cujo gene é preferencialmente expresso em V1

## 5.2. Secreção de enzimas digestivas

### 5.2.1. Secreção de proteinases e a rota endossomo-lisossomo

A presença das peptidases catepsina D e catepsina L é recorrente na digestão em Heteroptera (Houseman e Downe, 1983a). Os trabalhos indicam a participação majoritária das catepsinas L (Terra et al., 1988; Silva e Terra, 1994; Defferrari et al., 2011). Contudo, não há consenso na participação relativa das duas classes de proteinases na digestão nesses insetos. Essas proteinases são membros



de grupos de enzimas descritas como lisossômicas, tornando sua ação extracelular um fato marcante (Houseman e Downe, 1982). Assim, abordaremos a participação das duas classes de proteinases na digestão enfatizando características que permitiram sua atuação na digestão extracelular.

#### 5.2.1.1. Catepsina D

Já há fortes evidências da participação das catepsinas D na digestão nos Heteroptera. Podemos citar a série de trabalhos que definiram a presença desta proteinase ácida no intestino médio do barbeiro *Rhodnius prolixus* (Houseman e Downe, 1982, 1983a, 1983b; Terra et al., 1988), no também barbeiro *Triatoma infestans* (Balczun et al., 2012) e no percevejo *Oncopeltus fasciatus* (Defferrari et al., 2011). Contudo, o estudo realizado com *Dysdercus peruvianus* não evidenciou a participação da catepsina D na hidrólise de proteína (hemoglobina como substrato), ficando demonstrada apenas a grande importância da atividade de cisteíno proteinase (Silva e Terra, 1994). Contudo, poderia ter sido negligenciada por ser minoritária a participação da catepsina D na digestão em *D. peruvianus*. Nesse contexto foi realizada a identificação dos genes responsáveis pela transcrição de catepsinas D e sua caracterização em aspectos que envolvem função fisiológica e secreção para luz intestinal.

Considerando que foi possível prever peptídeo sinal em todas as sequências analisadas, é possível dizer que todas as sequências correspondem às proteínas que são endereçadas para o sistema de membranas. Dentre as catepsinas D estudadas, as que apresentaram os resíduos catalíticos intactos apresentaram também de um a quatro sinais de N-glicosilação. Isso é um indício de que essas proteínas podem, hipoteticamente, ser encaminhadas para a via lisossômica. As proteínas com a falta de algum resíduo catalítico apresentam nenhum ou seis sinais de N-glicosilação. O sinal de N-glicosilação pode ser considerado um requisito para o encaminhamento lisossômico, sendo que sua ausência já foi relacionada com o encaminhamento de enzimas para a rota secretória (Tikkanen et al., 1997).

Diferentemente do que ocorre com as catepsinas L (veja abaixo), há no caso das catepsinas D uma forma alternativa de se distinguir a função lisossômica da enzima. Segundo Padilha et al. (2009) é possível distinguir as catepsinas D

lisossômicas daquelas que assumiram outra função no decorrer da evolução (como o caso da catepsina D lisossômica em relação à pepsina gástrica nos vertebrados). A forma de discriminar é analisar a presença de uma região consenso (DxPxPx(G/A)P), chamada de alça de prolina, que é conservada apenas nas catepsinas D que mantiveram sua ação intracelular (lisossômica). Foram isoladas no lúmen intestinal do barbeiro *T. infestans* duas catepsinas D, sendo que uma delas apresenta a alça de prolina (Balczun *et al.*, 2012). Esse resultado, entretanto, contradiz os resultados gerais obtidos para organismos como vertebrados, moscas e crustáceos (Padilha *et al.*, 2009; Rojo *et al.*, 2010).

O primeiro indício de que a DpCatD10 seria a catepsina D lisossômica é sua expressão em tecidos não digestivos. Os genes de catepsina D envolvidos em digestão não são expressos nos outros tecidos, ficando, portanto, com expressão restrita ao intestino médio. Quando consideramos a presença da alça de prolina entre as catepsinas D estudadas, apenas DpCatD10 apresenta todos os resíduos dessa alça, confirmando que apenas essa enzima teria função tipicamente lisossômica (ou de digestão intracelular). É claro que a dinâmica da secreção de enzimas no intestino ainda não é completamente compreendida para descartarmos sua ação extracelular neste órgão, mas há maior segurança em atribuir a essa enzima a participação na digestão intracelular (ou lisossômica) nos outros órgãos (que não o intestino médio).

A função lisossômica, como pressão seletiva, manteria características conservadas (como alça de prolina). Em uma filogenia, o resultado seria a formação de um agrupamento de proteínas com a mesma função. Já os produtos de duplicação gênica que divergiram para a função de digestão extracelular, acumulam mudanças até se deslocarem para fora do grupo lisossômico. Isso está de acordo com os resultados obtidos, já que a DpCatD10 se agrupa com outras catepsinas D sabidamente lisossômicas. Mais uma vez, isso não inviabilizaria sua secreção em um tecido específico, como parece ocorrer no intestino médio dos Hemiptera em geral.

A presença de várias das catepsinas D estudadas no lúmen de V1 demonstra que as enzimas citadas seguem a rota secretória e atingem o lúmen intestinal, confirmando as predições *in silico*, baseadas em características como a presença do sinal de N-glicosilação. Como podemos perceber nos dados de expressão gênica, há pequena expressão da DpCatD10 no intestino médio de *D. peruvianus*. Assim,

não identificar sua presença no conteúdo luminal de V1 não assegura sua ausência, já que foi isolada uma catepsina D do tipo lisossômica no lúmen intestinal de *T. infestans* (Balczun *et al.*, 2012).

Genes de enzimas relacionadas à digestão de alimentos podem ser regulados pela ingestão (Marinotti *et al.*, 2006). Este mecanismo permite a produção da enzima quando o seu substrato está disponível. Segundo essa ideia, haveria outra forma de diferenciar as enzimas digestivas das enzimas lisossômicas baseada na indutibilidade pela alimentação. Essa característica foi documentada para a atividade proteolítica em *R. prolixus* e para nível de expressão gênica em *T. infestans* (Houseman e Downe, 1983b e Balczun *et al.*, 2012, respectivamente). Genes de catepsina D têm regulação dupla, podendo permanecer com expressão basal típica dos *housekeeping* ou ser ativado para determinadas funções fisiológicas (Zaidi *et al.*, 2008). Isso ocorre também em insetos, por exemplo, no caso do ciclo vitelogênico (Dittmer e Raikhel, 1997). Logo, o acoplamento do aumento da expressão gênica com a ingestão de alimentos (e o não acoplamento no caso da DpCatD10) parece indicar sua função.

O conjunto de dados mostra que diferentes características definem dois grupos distintos de catepsinas D nos Heteroptera, as catepsinas D ligadas à digestão de alimentos ingeridos e as catepsinas D tipicamente lisossômicas. O segundo grupo parece atuar em todos os órgãos na digestão intracelular e seu gene parece não ser regulado pela alimentação. Já as catepsinas D ligadas à digestão parecem ser genes que divergiram mais recentemente no processo evolutivo e parecem refletir a reaquisição da capacidade de digerir polímeros no lúmen intestinal. Essa reaquisição teve impacto na regulação desses genes, cuja regulação é definida pela alimentação em alguns casos. A secreção das catepsinas D no intestino médio ocorre a despeito da função tipicamente lisossômica em outros organismos, evidenciando peculiaridades da evolução do grupo (Heteroptera).

#### **5.2.1.2. Catepsina L**

Os indícios que nos permitem definir quais enzimas fazem parte de digestão dos alimentos no lúmen intestinal (ou digestivas) e quais enzimas participam da digestão intracelular (ou lisossômica) são a expressão nos diferentes tecidos e a

comparação de sequências. Contudo, tais informações não nos permite dizer como se dá a entrada de enzimas tipicamente lisossômicas na rota secretora.

Considerando que enzimas encaminhadas para o lisossomo devem ser N-glicosiladas, a ampla presença do sinal de N-glicosilação nas DPCALs não nos permite, sob esse critério, diferenciar as enzimas digestivas das lisossômicas. Isso talvez seja uma indicação de que a sinalização para o encaminhamento lisossômicos seja compartilhada com as enzimas encaminhadas para a rota secretória. O receptor de manose-6-fosfato é responsável pelo encaminhamento da catepsina L e outras hidrolases ácidas para o lisossomo. Logo, essas hidrolases devem conter um sítio de modificação de manose-6-fosfato. A transferência da GlcNac-1-fosfato para uma manose específica ligada na catepsina L é um passo crítico no encaminhamento para o lisossomo (Bao et al., 1996; Brulke and Bonifacino, 2009).

Na busca por outra característica que nos permita distinguir o encaminhamento celular, podemos considerar a presença do *signal patch*. A ação da fosfotransferase é dependente do reconhecimento do *signal patch*. Esse sinal é caracterizado por 2-3 lisinas da superfície da molécula com 34 Å (em alguns casos a 40 Å) de distância entre elas (Tikkanen et al. 1997; Brulke e Bonifacino, 2009). Como os modelos estruturais de todas as DPCALs apresentaram pelo menos uma medida definida pelo *signal patch* não é possível distinguir as DpCaLs baseado nos sinais de encaminhamento. Como todas as DpCaLs apresentam essas características, elas seriam reconhecidas pela fosfotransferase e encaminhadas ao lisossomo das células intestinais.

Dentro do sistema endossomo-lisossomo há a possibilidade da catepsina L ser encaminhada para a rota secretória se superexpressa. A simples variação na expressão poderia indicar um mecanismo de secreção destas proteinases, assim como foi visto em experimentos de variação da expressão em catepsina L humana (Kane et al., 1988). Para isso, contudo, é necessária a presença de resíduos conservados na região carboxi-terminal (Chauhan et al., 1998). Essas duas características são respeitadas pelas DPCALs, tendo sido registrado tanto a superexpressão das catepsinas L no intestino (quando comparado ao nível de expressão da DPCAL10 em outros tecidos) quanto à presença da conservação carboxi-terminal. Em resumo, provavelmente todas as DpCaLs são capazes de seguir a rota lisossômica e serem secretadas quando superexpressas, assim como a

catepsina L humana. Nesse contexto, não se espera retenção ou secreção preferencial de nenhuma das catepsinas L, já que todas compartilham as características para o encaminhamento lisossômico e a secreção devido à superexpressão no intestino médio.

Os dados até então parecem indicar uma rota de secreção das catepsinas L que inclui a rota lisossômica. Isso torna ainda mais intrigante como seria essa rota de encaminhamento intracelular, visto que aparentemente as enzimas aqui estudadas compartilhariam os sinais de encaminhamento ao lisossomo. Em *Rhodnius prolixus* a atividade de catepsina L (naquele momento chamada de catepsina B) foi verificada em lisossomos e especialmente associada a vesículas de secreção, sendo sugerido a presença de duas populações de lisossomos (Billingsley e Downe, 1988).

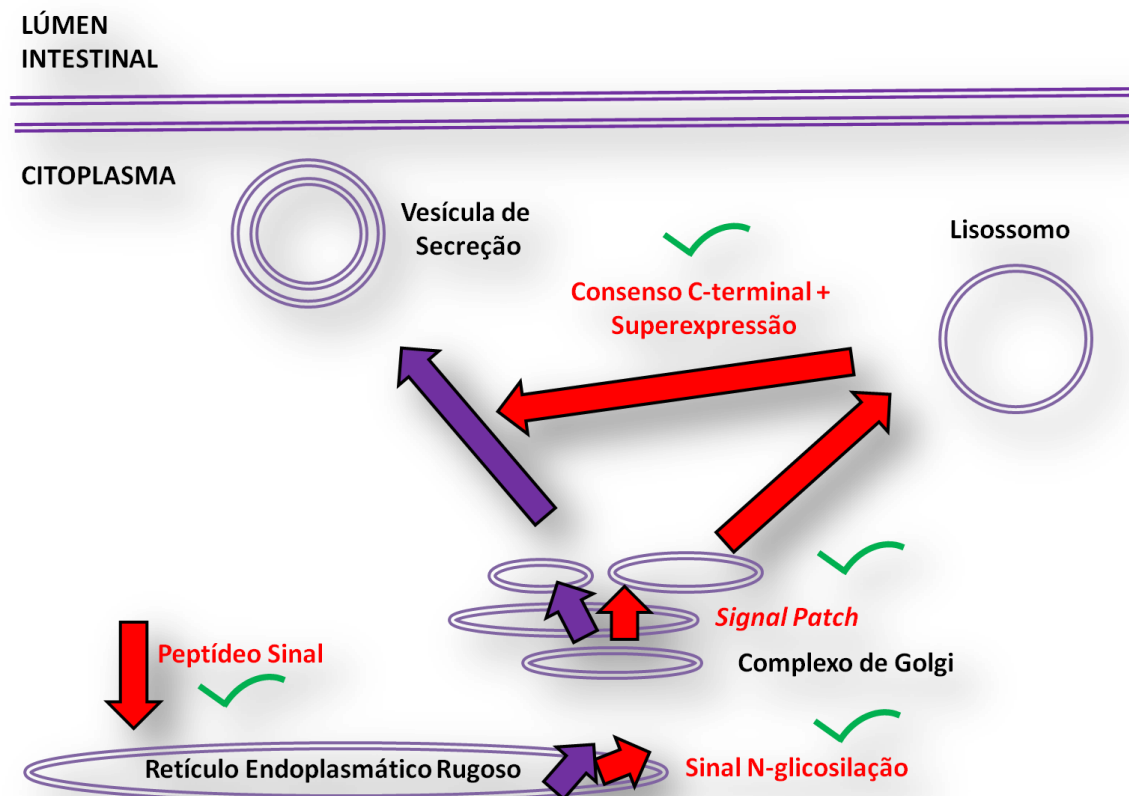
O pH de ativação (usando os dados obtidos com a DPCAL1) tem importância na compreensão da fisiologia deste inseto, já que a enzima não seria ativada no pH da luz intestinal que é de 5,6-6,2 (Silva e Terra, 1994). Isso indica a necessidade da enzima passar por compartimentos mais ácidos, como o lisossomo. A consequência direta desta lógica é a liberação das enzimas previamente ativadas no interior da célula (nas vesículas de secreção), exatamente como é visto nos experimentos de histoquímica em *R. prolixus* (Billingsley e Downe, 1988). Esse indício da passagem das catepsinas L pelo sistema endossomo-lisossomo está de acordo com o conjunto de predições do encaminhamento intracelular.

A hipótese de que as catepsinas L passariam pelo sistema endossomo-lisossomo e atingiriam a rota secretória é reforçada por dois experimentos complementares: a citolocalização da DPCAL10 e o proteoma de conteúdo de V1. A semelhança das organelas chamadas de lisossomos por Billingsley e Downe (1988) com as organelas marcadas pelo anti-DPCAL10 é clara. Mas, diferente da concepção clássica de lisossomo de membrana simples, esta organela apresenta conjuntos de membranas no seu interior. Considerando que o trabalho citado indica ação proteolítica do tipo catepsina L, essa organela deve, sim, ser relacionada com o lisossomo. Já quanto à presença de várias DPCALs no lúmen de V1, podemos dizer que atingem o meio extracelular (são secretadas). A identificação da DPCAL10 no material luminal indica, também, que a secreção ocorre para as DPCALs em geral no intestino de *Dysdercus peruvianus*.

### 5.2.1.3. Secreção de proteinases

A presença das proteinases catepsina D e catepsina L como agentes na hidrólise do alimento ingerido foi melhor caracterizada nesse estudo. Os dados sugerem que o encaminhamento das proteinases lisossômicas para a rota secretora se dê por supreexpressão, como foi visto no caso de células de mamíferos (Kane et al., 1988).

O modelo proposto para a secreção das proteinases lisossômicas se dá conforme a figura 55. A presença de peptídeo sinal permitiria a entrada no sistema de membranas. O sinal de N-glicosilação e a presença do *signal patch* permitiria o endereçamento para o sistema endossomo-lisossomo. A superexpressão, associado à presença do consenso C-terminal, permitiria o acesso das catepsinas à rota secretora.



**Figura 55.** Ligação entre a rota endossomo-lisossomo e a rota secretória. A rota endossomo-lisossomo e a rota secretória estão indicadas com setas vermelhas e roxas, respectivamente. Escrito em vermelho os atributos necessários para seguir a rota em vermelho. Sinais em verde indicam a presença dos atributos nas catepsinas L analisadas.

Para testar esse modelo, experimentos complementares foram feitos. A imunomarcção da DPCAL10 em elementos tanto da rota secretora quanto do sistema endossomo-lisossomo reforça o modelo. Outro dado que reforça essa proposta é a detecção do peptídeo da DPCAL10 no material solúvel do conteúdo de V1, assim como a detecção das outras catepsinas D e L aqui tratadas. Ou seja, os estudos de expressão gênica e análises *in silico* predizem os dados experimentais de imunomarcção e proteômica.

De maneira geral, é possível compreender o mecanismo celular que permite o acesso das proteinases tipicamente lisossômicas ao meio extracelular, aquisição importante no processo evolutivo do grupo que permite a digestão de proteínas.

### 5.2.2. Secreção de enzimas digestivas no sistema de duas membranas

Nos experimentos de centrifugação diferencial foi possível obter resultados similares aos obtidos por Silva et al. (1996), exceto pela proporção das barras das frações P1, P2 e P3 na abscissa (% de proteína). A diferença se deve ao processo de homogeneização, aqui realizado de maneira mais branda, liberando menor quantidade de pequenas frações subcelulares e aumentando proporcionalmente as barras de P1 em relação às de P2 e P3.

O resultado da atividade de catepsina L, por ser inédito, não foi possível comparar diretamente com a literatura. Contudo, sua distribuição foi bastante parecida com o resultado obtido por Silva et al. (1996) quanto à atividade de fosfatase ácida (um marcador lisossômico), ficando parcialmente solúvel e parcialmente ligada à fração insolúvel (ou fração de membranas). Vale a pena, contudo, salientar que a atividade de fosfatase ácida foi descrita em *Rhodnius prolixus* com distribuição subcelular similar a atividade de aminopeptidase, uma enzima do espaço perimicrovilar (Ferreira et al., 1988). Diferentemente de *R. prolixus* e do também Hemiptera *Eurygaster integriceps*, a atividade de fosfatase ácida em *D. peruvianus* é enriquecida (>3) em P2, P3 e P4, mostrando ser um marcador não só do espaço perimicrovilar como também das frações lisossômicas em *D. peruvianus* (inferido a partir dos trabalhos: Ferreira et al., 1988; Silva et al., 1996; Allahyari et al., 2010).

Como análise geral desses resultados, podemos dizer que o processo de resolução subcelular é fiel aos dados da literatura quanto à distribuição da atividade

enzimática pelas frações. Analisando isoladamente a distribuição subcelular da atividade de catepsina L, podemos dizer que segue o mesmo padrão de distribuição de outro marcador lisossômico (Silva *et al.* 1996), indicando sua presença ali. Considerando o trânsito intracelular dessa enzima, sua localização subcelular deve incluir também vesículas de secreção e essas, por sua vez, se comportariam no processo de fracionamento subcelular como os lisossomos, como visto na seção anterior. É importante notar que, mesmo como uma enzima sabidamente solúvel, a atividade de catepsina L é verificada em frações insolúveis mesmo depois do processo de congelamento e descongelamento. Isso indica que frações subcelulares membranosas são capazes de reter a atividade de catepsina L, o que deve refletir em estruturas subcelulares capazes de gerar tal efeito.

De acordo com Silva *et al.* (1996) homogeneizados teciduais não permitem a separação da atividade de alfa e beta-glicosidase e, portanto, não permitem a separação das membranas micro e perimicrovilares. Os autores só conseguiram a separação das atividades em preparações de conteúdo intestinal, situação em que a membrana perimicrovilar apresenta extensões luminais. O mesmo dado foi obtido nesta tese e, por não haver um pico exclusivo para uma das atividades, temos um indício que vesículas contendo apenas uma das membranas não ocorrem no tecido. Com isso, podemos inferir que o encaminhamento celular das duas classes de enzimas é feita por uma rota comum que envolve a presença de dupla membrana. Este resultado concorda com o modelo da origem da membrana perimicrovilar proposto por Silva *et al.* (1995). Nele, a formação de vesícula de duas membranas se dá ainda nas cisternas trans-Golgi, o que levaria à presença de apenas um tipo de vesícula de secreção que carregaria os dois marcadores celulares.

Os anticorpos que reconhecem proteínas sabidamente digestivas têm um padrão de marcação bastante constante das organelas. Anti-DPMan, anti-DPAlphaGlu1, anti-DPAlphaGlu2 e anti-DPAPN marcam grandes vesículas originadas no sistema trans-Golgi e marcam também as vesículas de secreção (de dupla membrana). Exceto por fatores como falta de resolução das membranas e baixo título do anticorpo, as marcações seguem o modelo proposto para a formação da membrana perimicrovilar e a localização dos marcadores (Silva *et al.* 1995; Silva *et al.*, 1996). Os compartimentos intracelulares marcados são de dupla membrana, assim como as marcações obtidas em Silva *et al.* (1995) para a alfa-glicosidase de membrana. Somando a esse modelo de secreção proposto e com base na



marcação da DPAlphaGlu2 e sua presença no lúmen de V1, podemos dizer que a secreção de enzimas digestivas solúveis tipicamente digestivas ocorrem pelo sistema de vesículas de dupla membrana, assim como previsto para as enzimas tipicamente lisossômicas que participam da digestão. Vale a pena ressaltar que anti-DPMan, anti-DPAlphaGlu1, anti-DPAlphaGlu2 e anti-DPAPN não foram capazes de marcar organelas do sistema endossomo-lisossômo.

Considerando agora as marcações específicas para a catepsina L lisossômica (DPCAL10), há indícios de uma rota celular diferente para as proteinases lisossômicas. A marcação de organelas ainda relacionadas aos lisossomos aconteceu apenas com o anti-DPCAL10 e não para os outros anticorpos, sugerindo haver rotas distintas (mas convergentes) para a secreção de enzimas digestivas no inseto estudado. A superexpressão de catepsinas (L no caso) gera um efeito global com a secreção de todas as catepsinas L, mesmo as menos expressas.

Essa rota distinta para secreção de enzimas tipicamente lisossômicas poderia incluir uma quantidade maior de enzimas ou mesmo todas as enzimas lisossômicas. O proteoma do material solúvel luminal de V1 foi útil para definir quais outras enzimas tipicamente lisossômicas são secretadas no intestido médio de *D. peruvianus*. Usando uma lista completa de enzimas solúveis tipicamente lisossômicas (Schröder et al., 2010) foi possível compara-las com as proteínas identificadas no proteoma. As proteínas compartilhadas são apenas as catepsinas B, D, L e F e, de maneira inédita, foi encontrada uma tiol redutase que está relacionada a redução de pontes dissulfeto, facilitando o desdobramento e proteólise – provavelmente funcionando em sinergia com as catepsinas secretadas.

Quanto às catepsinas, já haviam dados fortes para sua secreção, incluindo dados de genoma e transcriptoma indicando expansões gênicas, provavelmente relacionadas à nova função de digestão extracelular. Análise recente do genoma de *R. prolixus* indicou uma expansão dos genes de catepsina D, que foi relacionado à substituição das enzimas digestivas do tipo serino-proteinases por proteinases lisossomais (Mesquita et al., 2015). Essas expansões já haviam sido previstas para outras proteinases lisossomais que provavelmente assumiram função digestiva em *R. prolixus*. Seu transcriptoma já indicava a expansão bem como expressão em tecidos intestinais de catepsinas D, L e F (Ribeiro et al., 2014).

## 6. Conclusões

A digestão de proteínas baseada em enzimas originalmente lisossômicas é uma característica marcante nos Hemiptera, em especial os Heteroptera que voltaram a uma alimentação de polímeros (Terra e Ferreira, 2012). Assim como visto para *Rhodnius prolixus*, essa característica parece ter surgido com a duplicação desses genes, gerando o que é chamado de expansão gênica (Ribeiro et al., 2014; Mesquita et al., 2015). Dessa série de duplicações gênicas, parece haver a manutenção de um gene para função puramente lisossômica e a divergência funcional dos demais genes para a função de digestão extracelular.

O gene que mantém a função lisossômica parece não responder aos estímulos da alimentação, além de apresentar expressão nos mais diversos tecidos. Já os genes que se especializaram na digestão extracelular têm a expressão aumentada com a ingestão de alimento, indicando sua função.

Parece não haver diferença nas características relacionadas ao encaminhamento celular das proteínas produzidas por esses genes. O direcionamento para a rota secretória provavelmente está relacionado a superexpressão dos genes que garantem o acesso das proteinases lisossômicas à essa rota.

O modelo proposto por esse trabalho está resumido na figura 56. Nele podemos ver a localização celular relativa de cada marcador bem como as duas rotas de secreção aqui tratadas. É uma compilação que leva em conta os trabalhos prévios com *Dysdercus peruvianus* e as contribuições da pesquisa aqui desenvolvida.



descrevem diferentes atividades enzimáticas no processo digestivo, associando transcritos e as proteínas responsáveis por cada atividade hidrolítica. Foi possível indicar a identidade das glicosidases e peptidases bem como a localização da atuação das lipases. Nesse processo se verificou prováveis equívocos na literatura quanto à identidade da enzima aminopeptidase e do transportador de glicose do tipo GLUT. Foi possível ampliar o modelo de digestão incluindo a participação das catepsinas D, de uma alfa-glicosidase solúvel e da possível participação de uma tiolredutase.

Por fim, podemos fazer não só inferências sobre o encaminhamento e secreção de enzimas no sistema de membranas de *D. peruvianus*, mas também compreender o processo digestivo em si, confirmando e ampliando o modelo da digestão nesse organismo tão peculiar.

## 7. Referências Bibliográficas

- Adang, M.J. Insect aminopeptidase N. In: Barrett, A.J., Rawlings, N.D., Woessner, J.F. Handbook of Proteolytic Enzymes 2nd Ed. Elsevier Academic Press, 1, 296-299, 2004.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P. Molecular biology of the cell. Garland Science, 4<sup>ed</sup>, 2007.
- Allahyari, M., Bandani, A.R., Habibi-Rezaei, M. Subcellular fractionation of midguts cells of the sunn pest *Eurygaster integriceps* (Hemiptera: Scutelleridae): Enzyme markers of microvillar and perimicrovillar membranes. Journal of Insect Physiology, 56, 710-717, 2010.
- Baerwald, R.J., Delcarpio, J.B. Double membrane-bounded intestinal microvilli in *Oncopeltus fasciatus*. Cell and Tissue Research, 232, 593-600, 1983.
- Bagshaw, R.D., Mahuran D.J., Callahan, J.W. A proteomic analysis of lysosomal integral membrane proteins reveals the diverse composition of the organelle. Molecular & Cellular Proteomics, 4.2. 133-143, 2005.
- Balczun, C., Siemanowski, J., Pausch, J.K., Helling, S., Marcus, K., Stephan, C., Meyer, H.E., Schneider, T., Cizmowski, C., Oldenburg, M., Höhn, S., Meiser, C.K., Schuhmann, W., Shaub, G.A. Intestinal aspartate proteases TiCatD and TiCatD2 of the haematophagous bug *Triatoma infestans* (Reduviidae): Sequence characterisation, expression pattern and characterisation of proteolytic activity. Insect Biochemistry and Molecular Biology, 42, 240-250, 2012.
- Bao, M., Booth, J.L., Elmendorf, B.J., Canfield, W.M. Bovine UDP-N-acetylglucosamine: lysosomal enzyme N-acetylglucosamine-1-phosphothansferase. The Journal of Biological Chemistry, 271, 31437-31445, 1996.
- Beton, D., Guzzo, C., Ribeiro, A.F., Farah, C., Terra, W.R. The 3D structure and function of digestive cathepsin L-like proteinases of *Tenebrio molitor* larval midgut. Inset Biochemistry and Molecular Biology, 42, 655-664, 2012.
- Bifano, T. D., Alegria, T. G.P., Terra, W. R. Transporters involved in glucose and water absorption in the *Dysdercus peruvianus* (Hemiptera: Pyrrhocoridae) anterior midgut. Comparative Biochemistry and Physiology. B, Biochemistry & Molecular Biology, 157, 1-9, 2010.

- Billingsley, P.F., Downe, A.E.R. Ultrastructural localization of cathepsin B in the midgut of *Rhodnius prolixus* Stal (Hemiptera: Reduviidae) during blood digestion. *Int. J. Insect Morphol. & Embryol.*, 17, 295-302, 1988.
- Braulke, T., Bonifacino, J.S. Sorting of lysosomal proteins. *Biochimica et Biophysica Acta*, 1793, 605-614, 2009.
- Chauhan, S.S., Ray, D., Kane, S.E., Willingham, M.C., Gottesman, M.M. Involvement of carboxy-terminal amino acids in secretion of human lysosomal protease cathepsin L. *Biochemistry*, 37, 8584-8594, 1998.
- Conner, G.E. Cathepsin D. In: Barrett, A.J., Rawlings, N.D., Woessner, J.F. *Handbook of Proteolytic Enzymes 2nd Ed.* Elsevier Academic Press, 1, 43-52, 2004.
- Costa, I.A., Samuels, R.I., Bifano, T.D., Terra, W.R., Silva, C.P. Purification and partial characterization of an aminopeptidase from the midgut tissue of *Dysdercus peruvianus*. *Comparative Biochemistry and Physiology, part. B*, 158, 235-241, 2011.
- Cristofolletti, P.T., Terra, W.R. Specificity, anchoring, and subsites in the active center of a microvillar aminopeptidase purified from *Tenebrio molitor* (Coleoptera) midgut cells. *Insect Biochemistry and Molecular Biology*, 29, 807-819, 1999.
- Cristofolletti, P. T., Ribeiro, A. F., Terra, W. R. The cathepsin L-like proteinases from the midgut of *Tenebrio molitor* larvae: sequence, properties, immunocytochemical localization and function. *Insect Biochemistry and Molecular Biology*, 35, 883-901, 2005.
- Cristofolletti, P. T., Ribeiro, A. F., Deraison, C., Rahbe, Y., Terra, W. R. Midgut adaptation and digestive enzyme distribution in a phloem feeding insect, the pea aphid *Acyrtosiphon pisum*. *Journal of Insect Physiology*, 49, 11-24, 2003.
- Cristofolletti, P.T., Sousa, F.A.M., Rahbé, Y., Terra, W.R. Characterization of a membrane-bound aminopeptidase purified from *Acyrtosiphon pisum* midgut cells. *FEBS Journal*, 273, 5574-5588, 2006.
- De Duve, C. The lysosome turns fifty. *Nature Cell Biology*, 7, 847-849, 2005.
- Defferrari, M.S., Demartini, D.R., Marcelino, T.B., Pinto, P.M., Carlini, C.R. Insecticidal effect of *Canavalia ensiformes* major urease on nymphs of the milkweed bug *Oncopeltus fasciatus* and characterization of digestive peptidases. *Insect Biochemistry and Molecular Biology*, 41, 388-399, 2011.

- Dittmer, N.T., Raikhel, A.S. Analysis of the mosquito lysosomal aspartic protease gene: an insect housekeeping gene with fat body-enhanced expression. *Insect Biochem. Molec. Biol.*, 27, 323-335, 1997
- Espinoza-Fuentes, F.P., Terra, W.R. Physiological adaptations for digesting bacteria. Water fluxes and distribution of digestive enzymes in *Musca domestica* larval midgut. *Insect Biochem.*, 17, 809-817, 1987.
- Ferreira, C., Ribeiro, A. F., Garcia, E. S., Terra, W. R. Digestive enzymes trapped between and associated with the double plasma membranes of *Rhodnius prolixus* posterior midgut cells. *Insect Biochemistry*, 18, 521-530, 1988.
- Fortenberry, S.C., Schorey, J.S. Chirgwin, J.M. Role of glycosylation in the expression of human procathepsin D. *Journal of Cell Science.*, 108, 2001-2006, 1995.
- Fujimoto, Y., Kobayashi, A., Kurata, S., Natori, S. Two subunits of the insect 26/29-kDa proteinase are probably derived from a common precursor protein. *J. Biochem.*, 125, 566-573, 1999.
- Griffiths, G., Simons, K. The *trans* Golgi network: sorting at the exit of the Golgi complex. *Science*, 234, 438-443, 1986.
- Grimaldi, D., Engel, M.S. *Evolution of the insects*. Cambridge University Press, 2005
- Hoseman J.G., Downe, A.E.R. Endoproteinase activity in the posterior midgut of *Rhodnius prolixus* Stal (Hemiptera: Reduviidae). *Insect Biochem.*, 10, 363-366, 1980.
- Hoseman J.G., Downe, A.E.R. Characterization of an acidic proteinase from the posterior midgut of *Rhodnius prolixus* Stal (Hemiptera: Reduviidae). *Insect Biochem.*, 12, 651-655, 1982.
- Hoseman J.G., Downe, A.E.R. Cathepsin D-like activity in the posterior midgut of hemipteran insects. *Comp. Biochem. Physiol.*, 75B, 509-512, 1983.
- Hoseman J.G., Downe, A.E.R. Activity cycles and the control of four proteinases in the posterior midgut of *Rhodnius prolixus* Stal (Hemiptera: Reduviidae). *J. Insect Physiol.*, 29, 141-148, 1983.
- Kane, S.E., Troen, B.R., Gal, S., Ueda, K., Pastan, I., Gottesman, M. Use of a cloned multidrug resistance gene for coamplification and overproduction of major excreted protein, a transformation-regulated secreted acid protease. *Mol. Cell. Biol.*, 8, 3316-3321, 1988.

- Kirschke, H. Cathepsin L. In: Barrett, A.J., Rawlings, N.D., Woessner, J.F. Handbook of Proteolytic Enzymes 2nd Ed. Elsevier Academic Press, 2, 1097-1102, 2004.
- Kollien, A.H., Waniek, P.J., Nisbet, A.J., Billingsley, Schaub, G.A. Activity and sequence characterization of two cysteine proteases in the digestive tract of the reduviid bug *Triatoma infestans*. *Insect Molecular Biology*, 13, 569-579.
- Lane, N.J., Harrison, J.B. An unusual cell surface modification: a double membrane. *J. Cell Sci.*, 39, 355-372, 1979.
- Lopez-Ordoñez, T., Rodriguez, M.H., Hernández-Hernández, F.C., Characterization of a cDNA encoding a cathepsin L-like protein of *Rhodnius prolixus*. *Insect Molecular Biology*, 10, 505-511, 2001.
- Marinotti, O., Calvo, E., Nguyen, Q.K., Dissanayake, S., Ribeiro, J.M.C., James, A.A. Genome-wide analysis of gene expression in adult *Anopheles gambiae*. *Insect Molecular Biology*, 15, 1-12, 2006
- Mesquita, R.D., Vionette-Amaral, R.J., Lowenberger, C., Rivera-Pomar, R, ..., Oliveira, L.P. Genome of *Rhodnius prolixus*, an insect vector of Chagas disease, reveals unique adaptation to hematophagy and parasite infection. *PNAS Early Edition*, 1-6, 2015.
- Padilha, M.H.P., Pimentel, A.C., Ribeiro, A.F., Terra, W.R. Sequence and function of lysosomal and digestive cathepsin D-like proteinases of *Musca domestica*. *Insect Biochemistry and Molecular Biology*, 39, 782-791, 2009.
- Pereira, M.J.B., Albuquerque, F.A., Bastos, C.S. Pragas do algodoeiro: identificação, biologia e sintomas de ataque. *Rev. Bras. Ol. Fibros. Campina Grande*, 10, 1073-1117, 2006.
- Press, E.M., Porter, R.R., Cebra, J. The isolation and properties of a proteolytic enzyme, cathepsin D, from bovine spleen. *Biochemical Journal*, 74, 501-514, 1960.
- Ramos C. C. R., Abreu P.A.E., Nascimento, A.L.T.O., Ho P.L. A high-copy T7 *Escherichia coli* expression vector for the production of recombinant proteins with a minimal N-terminal His-tagged fusion peptide. *Brazilian Journal Medical and Biological Research*, 37, 1103-1109, 2004.
- Ribeiro, J.M., Genta, F.A., Sorgine, M.H.F., Logullo, R., Mesquita, R.D., ..., Oliveira, P.L. An insight into the transcriptome of the digestive tract of the bloodsucking bug, *Rhodnius prolixus*. *Plos Neglected Tropical Diseases*, 8, 1-31, 2014.



- Rojo, L., Sotelo-Mundo, R., García-Carreño, F., Gráf, L. Isolation, biochemical characterization, and molecular modeling of American lobster digestive cathepsin D1. *Comparative Biochemistry and Physiology*, 157, 394-400, 2010.
- Saito, H., Kurata, S., Natori, S. Purification and characterization of a hemocyte proteinase of *Sarcophaga*, possibly participating in elimination of foreign substances. *Eur. J. Biochem.*, 209, 939-944, 1992.
- Sangadala, S., Walters, F.S., English, L.H., Adang, M.J. A mixture of *Manduca sexta* aminopeptidase and phosphatase enhances *Bacillus thuringiensis* insecticidal CryIA(c) toxin binding and  $^{86}\text{Rb}^+\text{-K}^+$  efflux *in vitro*. *The Journal of Biological Chemistry*, 269, 10088-10092, 1994.
- Schröder, B.A., Wrocklage, C., Hasilik, A., Saftig, P. The proteome of lysosomes. *Proteomics*, 10, 4053-4076, 2010.
- Silva, C.P., Terra, W.R. Digestive and absorptive sites along the midgut of the cotton seed sucker bug *Dysdercus peruvianus* (Hemiptera: Pyrrhocoridae). *Insect Biochemistry and Molecular Biology*, 24, 493-505, 1994.
- Silva, C.P., Terra, W.R. An  $\alpha$ -glucosidase from perimicrovillar membranes of *Dysdercus peruvianus* (Hemiptera: Pyrrhocoridae) midgut cells. Purification and properties. *Insect Biochem. Molec. Biol.*, 25, 487-494, 1995.
- Silva, C.P., Terra, W.R. Alfa-galactosidase activity in ingested seeds and in the midgut of *Dysdercus peruvianus* (Hemiptera: Pyrrhocoridae). *Arquives of Insect Biochemistry Physiology*, 34, 443-460, 1997.
- Silva, C.P., Ribeiro, A.F., Gulbenkian, S., Terra, W.R. Organization, origin and function of the outer microvillar (perimicrovillar) membranes of *Dysdercus peruvianus* (Hemiptera) midgut cells. *J. Insect Physiol.*, 41, 1093-1103, 1995.
- Silva, C.P., Ribeiro, A.F., Terra, W.R. Enzymes markers and isolation of the microvillar and perimicrovillar membranes of *Dysdercus peruvianus* (Hemiptera: Pyrrhocoridae) midgut cells. *Insect Biochemistry and Molecular Biology*, 26, 1011-1018, 1996.
- Suzuki, T., Natori, S. Preferential degradation and elimination of foreign protein introduced into the abdominal cavity of *Sarcophaga peregrine* larvae. *Comp. Biochem. Physiol.*, 81A, 191-193, 1985.
- Waniek, P., Costa, J.E.P., Jansen, A.M., Costa, J., Araújo, C.A.C. Cathepsin L of *Triatoma brasiliensis* (Reduviidae, Triatominae): Sequence characterization,

- expression pattern and zymography. *Journal of Insect Physiology*, 58, 178-187, 2012.
- Tao, K., Stearns, N.A., Dong, J., Wu, Q., Sahagian, G.G. The proregion of cathesin L is required for proper folding stability, and ER exit. *Archives of Biochemistry and Biophysics*, 311, 19-27, 1994.
- Terra, W. R., Ferreira, C. Biochemistry and molecular biology of digestion. In: Gilbert, L. I. *Insect Molecular Biology and Biochemistry*. Academic Press, London, 365-418, 2012.
- Terra, W.R., Ferreira, C., Garcia, E.S. Origin, distribution, properties, and functions of the major *Rhodnius prolixus* midgut hydrolases. *Insect Biochemistry*, 18, 423-434, 1988.
- Tikkanen, R., Peltola, M., Oinonen, C., Rouvinen, J., Peltonen, L. Several cooperating binding sites mediate the interaction of a lysosomal enzyme with phosphotransferase. *The EMBO Journal*, 16, 6684-6693, 1997.
- Zaidi, N., Maurer, A., Nieke, S., Kalbacher, H. Cathepsin D: a cellular roadmap. *Biochemical and Biophysical Research Communications*, 376, 5-9, 2008.

## 8. Apêndices

### Material Suplementar I - Metodologia bioinformática para análise de dados Illumina:

**Pré-processamento e montagem.** Os dados obtidos pelo HiScanSq (Illumina, Inc., San Diego, CA) foram analisados em quatro etapas principais. Na etapa de obtenção dos dados brutos foi utilizado o *software* CASAVA 1.8.2 fornecido pela Illumina, que faz o *base call* dos dados brutos e os transforma em *reads* no formato fastq acompanhados dos *scores* de qualidade phred. Os *reads* foram visualizados utilizando o programa FastQC ([www.bioinformatics.bbsrc.ac.uk/projects/](http://www.bioinformatics.bbsrc.ac.uk/projects/)). Os que possuíam uma mediana de qualidade inferior a 28 na escala phred foram removidos utilizando os scripts de Smith et al (2011) a fim de evitar artefatos e reduzir o tempo das análises posteriores. Sequências dos índices e adaptadores utilizadas na confecção das bibliotecas de cDNA também serão removidas, utilizando a ferramenta de FASTX-Toolkit, versão 0.0.13 ([http://hannonlab.cshl.edu/fastx\\_toolkit/index.html](http://hannonlab.cshl.edu/fastx_toolkit/index.html)). A montagem de novo foi realizada pelos programas Velvet/Oases (Zerbino e Birney, 2008; Schulz et al 2012). Foram feitas quatro montagens para todas as amostras, com kmers de 51 a 81, com intervalos de 10pb, para as amostras de *Dysdercus peruvianus*. Essas montagens foram agrupadas utilizando a função merge. O arquivo fasta desses contigs foi então filtrado e um arquivo com os exemplares foi obtido. Esse filtro consiste na remoção de: transcritos redundantes, daqueles com tamanho inferior a 150pb, variantes de splicing e contigs de baixa confiança. As montagens passaram também por um blast contra a base de dados Univec (<http://www.ncbi.nlm.nih.gov/VecScreen/UniVec.html>), quando sequências de vetores e primers foram identificados como contaminantes e removidos. Em seguida foi feito um blast sobre a base de dados SwissProt ([http://web.expasy.org/docs/swiss-prot\\_guideline.html](http://web.expasy.org/docs/swiss-prot_guideline.html)) e os contigs reconhecidos foram anotados conforme o respectivo hit. **Anotação GO e Pfam.** Os contigs foram mapeados contra uma seleção da base de dados NCBI não-redundante (nr), utilizando o subgrupo de proteínas de Metazoa. Para isso foi utilizado o programa blastx com o cut-off de  $1e-10$  (Altschul et al, 1990; Camacho et al 2009). A anotação funcional dos termos de ontologia gênica (GO) e seus derivados foi realizada utilizando o programa Blast2GO (Conesa et al 2005), a partir de 20 hits do Blast para cada contig. Utilizando a ferramenta Interproscan v. 5RC1 (Zdobnov & Apweiler 2001), os modelos de *hidden Markov* (HMMs) presentes nos domínios da base PFAM e Profile foram recuperados para os contigs. **Análise de expressão diferencial.** O programa BWA (Li & Durbin 2010) foi utilizado para o mapeamento contra os transcriptomas de referência. Os arquivos .bam obtidos foram então

reorganizados, utilizando os programas Samtools (ferramenta sort, Li et al 2009) e convertidos em arquivos .sam. Os reads mapeados em mais de um sítio foram removidos usando a ferramenta MarkDuplicates do programa Picard ([http:// picard.sourceforge.net/](http://picard.sourceforge.net/)). Em seguida o número de *reads* alinhados contra cada contig do transcriptoma de referência foi obtido através de um script e tabelados no formato adequado para uso no programa DESeq (Anders & Huber, 2010), um pacote do R/Bioconductor (Gentleman et al 2004). Dentro do programa DESeq os dados foram transformados utilizando a função `estimateVarianceFunctions`, ignorando os grupos e calculando a variância levando em conta todas as amostras (procedimento recomendado pelos autores na ausência de réplicas). Em seguida foi realizado o teste binomial de expressão e os contigs cujas amostras tinham o `baseMean` inferior ao valor 5 foram removidos. Sobre o teste foi aplicada a correção de Benjamini-Hochberg (Benjamini & Hochberg 1995) para múltiplos testes, a fim de evitar os erros tipo I, ou seja, os positivos falsos.

### **Referências bibliográficas (Material Suplementar I)**

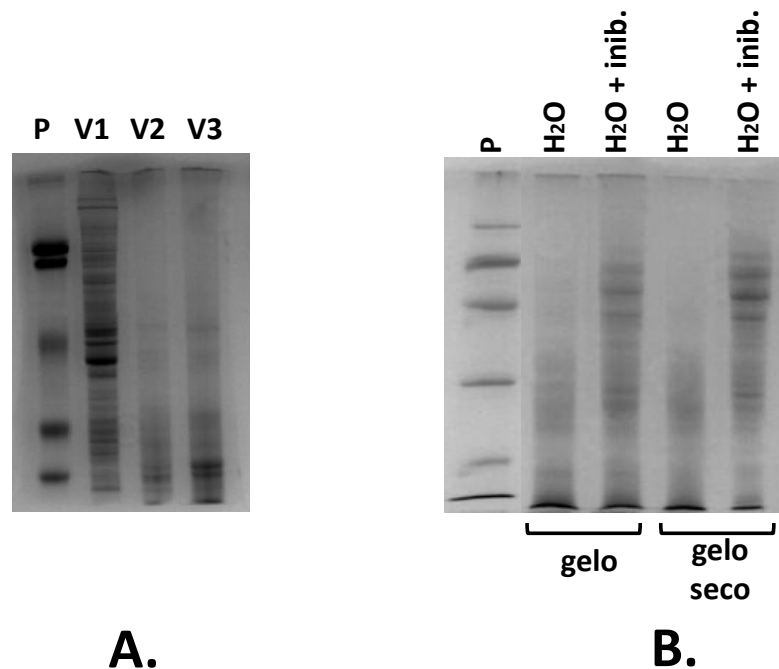
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic local alignment search tool. *J Mol Evol* 1990, 215: 403–410.
- Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biology*, 11, R106.
- Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing". *Journal of the Royal Statistical Society, Series B (Methodological)* 57:289–300.
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL: BLAST+: architecture and applications. *BMC Bioinformatics* 2009, 10:421.
- Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M: Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 2005, 21:3674–3676.
- Gentleman R, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y et al. 2004. Bioconductor: Open software development for computational biology and bioinformatics R. *Genome Biology*, Vol. 5, R80.
- Li H. and Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, 26, 589-595.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. and 1000 Genome Project Data Processing Subgroup. 2009. The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* 25, 2078-9.

- Smith S, Wilson NG, Goetz F, Feehery C, Andrade SCS, Rouse GW, Giribet G, Dunn CW. 2011. Resolving the evolutionary relationships of molluscs with phylogenomic tools. *Nature*, 480:364-367.
- Schulz M.H., Zerbino D.R., Vingron M., Birney, E. 2012. Oases: Robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics*, 28: 1086-1092.
- Zdobnov E.M. and Apweiler R. 2001. InterProScan - an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17(9): 847-8
- Zerbino DR, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18: 821–829.

### **Material Suplementar II - Evitando degradação das amostras**

Após dissecação dos tecidos, eles eram mantidos em gelo até o processo de homogeneização, no qual se adicionava inibidores de proteases, e após a homogeneização se realizava a quantificação de proteína do homogeneizado (ainda mantido em gelo) que era, por fim, fervido com tampão de amostra para SDS-PAGE na proporção correta para o experimento. Quarenta  $\mu\text{g}$  das amostras obtidas foram aplicadas em gel de poliacrilamida e podem ser vistas na figura 57 A. No gel é possível ver que a integridade das proteínas dos tecidos V2 e V3 não foi mantida, quando comparamos com as proteínas de V1.

V1, em *D. peruvianus* é o ventrículo responsável pela digestão de carboidratos e absorção de açúcares sendo pobre em atividade proteolítica. Já V2 e V3 são ventrículos associados à digestão de proteínas, apresentando uma forte atividade proteolítica. Isso faz com que as proteínas desses dois últimos tecidos sejam bastante suscetíveis à proteólise. O gel da figura 34 A indica que as proteínas de V2 e V3 perderam sua integridade, tipicamente acumulando material proteico de baixa massa molecular.



**Figura 57.** Eletroforese (SDS-PAGE) das proteínas de homogeneizados intestinais de *D. peruvianus*. **A.** Eletroforese das diferentes porções do intestino de *D. peruvianus* usada para transferência de proteínas para membrana de nitrocelulose. Homogeneizado com E-64 10 $\mu$ M e pepstatina 10 $\mu$ M. P. padrão de massa molecular; V1. Quarenta  $\mu$ g de proteína do homogeneizado de V1; V2. Quarenta mg de proteína do homogeneizado de V2; V3. Quarenta mg de proteína do homogeneizado de V3. **B.** Eletroforese de homogeneizados de V2 de *D. peruvianus*. Todos os homogeneizados foram obtidos de três fêmeas com V2 cheio. P. padrão de massa molecular; H<sub>2</sub>O; homogeneizado em água; H<sub>2</sub>O+inib.; homogeneizado na presença de inibidores; gelo: conservado em gelo até a homogeneização; gelo seco: conservado em gelo seco até a homogeneização.

Como não estava claro em que momento a proteólise ocorria, foi feito um experimento que busca comparar a forma de acondicionamento do tecido de V2 (tecido com menos proteínas íntegras na figura 57 A.) e o momento de adição dos inibidores de proteinases.

Das condições simuladas, a adição de solução contendo inibidores de proteases imediatamente após a dissecação foi a medida mais efetiva (figura 57 B). A manutenção apenas em gelo ou em gelo seco até a homogeneização claramente não garante a integridade das proteínas, mas manter em gelo-seco na presença de inibidores parece ser a medida mais eficiente, visto ter bandas mais bem definidas.

Foi testada também a copolimerização de inibidores de proteases no gel de poli(acrilamida) do SDS-PAGE, mas não afetou o resultado de nenhuma maneira.

**Tabela suplementar 1.** Relação de iniciadores utilizados em experimentos de clonagem, RT-PCR, qPCR. FW corresponde ao iniciador no sentido 5' para 3' e RV iniciador no sentido 3' para 5'.

<b>DPCAL clonagem</b>	
DPCAL1_BamHI	CGCGGATCCGGGTCGGTCTTTGAA
DPCAL1_PstI	TGGTTCTGCAGTTAAACCACAGGGTA
DPCAL2_BamHI	CGCGGATCCGCACGTGGTGATGCT
DPCAL2_EcoRI	CCGGAATTCTTACATTGTTACATA
DPCAL2_FW	GCACGTGGGGATGCTCCAGT
DPCAL2_RV	TTACATTGTTACATAAGTTG
DPCAL3_XhoI	CCGCTCGAGCTGTCAGTGTCTGAC
DPCAL3_HindIII	CCCAAGCTTTTAAACGTTAGGGTA
DPCAL5_BamHI	CGCGGATCCGTACCGTCGCTGACG
DPCAL5_EcoRI	CCGGAATTCCTACACTAGAGGGTA
DPCAL10_BamHI	CGCGGATCCGTATCCTTTTTTAAT
DPCAL10_PstI	TGGTTCTGCAGTTATACCAGAGGATA
DPCAL10_FW	GTATCCTTTTTTAATCTGGG
DPCAL10_RV	TTATACCAGAGGATAGCTGG
<b>DPCAL RT-PCR</b>	
DPCAL1_FW	GCCATGCCGGGTCGGTCTTT
DPCAL1_RV	CCAGCAGGAGCCGCACTGTC
DPCAL2_FW	AACTGCAGCGGGATTAGAGCCA
DPCAL2_RV	TGCCAAAGGACCAGCACGAACC
DPCAL3_FW	AGGGCCCAGGGATCCTGCAG
DPCAL3_RV	GGACCTACGTTGGCCACGGC
DPCAL4_FW	GGCGGCCTCATGACCTCAGC
DPCAL4_RV	CCCAGCTGGCTCCCCAGGAA
DPCAL5_FW	CCAAGGCCGTTGCGGAGCAT
DPCAL5_RV	ACGGTAGCCACCGCCTCCAT
DPCAL6_FW	TGCGGTTCGTGCTGGGCATT
DPCAL6_RV	ACTGGCGTCCATGCCAACGG
DPCAL7_FW	TTGGCCGCAGTGGGCTTTCC
DPCAL7_RV	GGGGTTGCATGACGCCTCGT
DPCAL8_FW	CAGTGCGGATCCTGCTGGGC
DPCAL8_RV	TACGGCCAGGACACCGTGGT
DPCAL9_FW	GCGGTTGGAGCACTCGAAGGG
DPCAL9_RV	CGCTGGCCGAACTAGAGCCA
DPCAL10_FW	GAGCCGCTGTGCAGTTCCGA
DPCAL10_RV	TGGCACTCGTGGCGATACCG
<b>DPCAL qPCR</b>	
DP_CAL1_FW	CCTCAAGGCGATGAAGAAGC
DP_CAL1_RV	TCGTAGTAGACACCTCCGGA
DP_CAL3_FW	GGTCCTGTATCTGTGGCCAT
DP_CAL3_RV	TGATCCAAGAACTCGGGCT
DP_CAL4_FW	CGAACTCCCGCAACTGTAAG

DP_CAL4_RV	GATACCGGTCCAACATTCGC
DP_CAL5_FW	TGTCTGGCTACGTGGACATT
DP_CAL5_RV	CACAACAAGAACACCGTGGT
DP_CAL10_FW	AAAACCTCTTGGGGAACGGGG
DP_CAL10_RV	CAGGGGATAGCTGGCACTC

---

#### Marcadores clonagem

DPMan_BamHI	CGCGGATCCAGTTTGGAGTCTGAC
DPMan_PstI	GTTCTGCAGTTAGTTGGTAATTTTG
DPAIpGlu1_BamHI	CGGATCCTCCGAGGACAAGTGG
DPAIpGlu1_HindIII	CCCAAGCTT TTAATTGCTCCTGAA
DPAIpGlu2_XhoI	CGCTCGAGAAGGAGGAATGGTGG
DPAIpGlu2_PstI	GTTCTGCAGTCAATAGCGGAACAC
DPAPN1_BamHI	CGCGGATCCGTTCTGCCGGATTAC
DPAPN1_EcoRI	CCGGAATTCTTATCTTGAAGACGA

---

#### Marcadores RT-PCR

DPMan_FW	AGTTTGGAGTCTGACGTTTA
DPMan_RV	TTAGTTGGTAATTTTGCCGACCC
DPAIpGlu1_FW	TCCGAGGACAAGTGGTGGCA
DPAIpGlu1_RV	TTAATTGCTCCTGAAGAGCG
DPAIpGlu2_FW	AAGGAGGAATGGTGGAGGGC
DPAIpGlu2_RV	TCAATAGCGGAACACGACGC
DPAPN1_FW	GTTCTGCCGGATTACTTCGT
DPAPN1_RV	TTATCTTGAAGACGACAAGG

---

#### DPCatD qPCR

DpCatD1_FW	GGCGAAAACCTGTCCCAAAA
DpCatD1_RV	CGACTCTCTGTTGTTGGTTTGT
DpCatD2_FW	AGGTGTTTGGTTTCGTTGTG
DpCatD2_RV	GCTACGGCCTCCTTCAGTAT
DpCatD3_FW	TGGAGACGGTGAAGAATTGTG
DpCatD3_RV	CTGCAGTCTACCGTAGCCAT
DpCatD4_FW	CACGAAAACGACAAAATGGGT
DpCatD4_RV	GTGTCCCTCTTGGCTTCTCT
DpCatD5_FW	GTGTTGGTCGTCGCTCTTTT
DpCatD5_RV	TGAAATCCCTGAGGCTCCTG
DpCatD6_FW	GGACCTGTAGCCAAGTGTCT
DpCatD6_RV	GCTTGAGGTTTGGTACGCC
DpCatD7_FW	CCTTGTTCCGGGTTCTCTT
DpCatD7_RV	TTGGGCGTTCATGTAGTTGC
DpCatD8_FW	CCAAGTGACCGTACAGCTCT
DpCatD8_RV	CCGTAACTCTTTCGGTCGT
DpCatD9_FW	CGAGAGTCCCGATGGCTATT
DpCatD9_RV	AACCCTATCATCCGGCTTCC
DpCatD10_FW	TCGTGTTTGGTGTGTTTGT
DpCatD10_RV	TCAGCGTAACATGCCTGACT

---

#### Controle qPCR

DP18S_I_FW	GATTCCTTCAGTGTAGCGCG
------------	----------------------



DP18S_I_RV	TCTTCTTAAGGGGACAGGCG
DP18S_II_FW	AAGGGCAGGGACGTAATCAA
DP18S_II_RV	CGCGCTACACTGAAGGAATC

**Tabela suplementar 2.** Proteínas solúveis identificadas no conteúdo de V1 de *Dysdercus peruvianus*. Estão representadas apenas proteínas com pelo menos dois espectros. O resultado corresponde a uma única amostra biológica.

Provável contaminante	Proteínas	Anotação	Contagem de espectros
	DPAPN	aminopeptidase N	33
*		JHBP	7
*		JHBP	6
		uncharacterized protein	2
		uncharacterized protein	8
		alpha galactosidase	10
		ankyrin	5
		cysteine protease (cathepsin F)	3
	DpAlpGlu2	alpha glucosidase	12
		unknown protein	7
	Tiolredutase	thiol reductase (gamma-interferon inducible)	4
		unknown secreted protein	6
		unknown secreted protein	5
		fructose-biphosphate aldolase	11
	DPCAL2	long N-term cathepsin L	6
	DPCAL3	cathepsin L	4
	DPCAL3	cathepsin L	4
	DPCAL4	cathepsin L	1
		lipocalin	4
		unknown protein	6
*		JHBP	5
	(AGO32660.1)	trehalase	5
	DpCatD5	cathepsin D	3
*		apolipophorin	3
		unknown protein	4
	DPCAL8	cathepsin L	2
	DPCAL3	cathepsin L	2
		nucleoside diphosphate kinase	6
		trypsin	4
		trypsin	3
		trypsin	2
	DpAlpGlu1	alpha glucosidase	6
		unknown secreted protein	3
		apolipophorin-III precursor	2
		unknown secreted protein	2

	unknown protein	2
	cathepsin B	2
	ankyrin	4
	unknown secreted protein	3
	superoxide dismutase [Cu-Zn]	2
	unknown secreted protein	3
	xylosyltransferase oxt	4
	glutathione S-transferase	3
DPCAL5	cathepsin L	4
	unknown secreted protein	2
	c-type lysozyme	2
	aldo-keto reductase	3
DpCatD7	cathepsin D	4
	serine protease	3
	arginine kinase	3
	IDGF (imaginal disc growth factor)	3
	apolipophorin	2
	JHBP	3
	protein yellow	2
*	vitellogenin	3
*	vitellogenin	2
*	vitellogenin	2
	60S ribosomal protein L12	2
	NADP-dependent malic enzyme	3
	ferritin	3
	cytosol aminopeptidase	3
	thioredoxin	3
	ankyrin	2
	vanin protein	2
	unknown secreted protein	2
	laccase	2
DPCAL3	cathepsin L	2
	unknown protein	2
	glyceraldehyde 3-phosphate dehydrogenase	2
*	enolase	2
	vitellogenin	2
DpCatD4	cathepsin D	2
	vitellogenin	2
DPCAL8	cathepsin L	2
	unknown protein	2
	serine protease	2
	multiple inositol polyphosphate phosphatase	2
	actin depolymerizing factor	2

JHBP (Haemolymph juvenile hormone binding protein)

## SÚMULA CURRICULAR

### DADOS PESSOAIS

André Coppe Pimentel

Brasileiro, nascido em São Paulo (SP) no dia 14 de Janeiro de 1984

### EDUCAÇÃO

*2011 Doutorado em andamento em Bioquímica.*

Instituto de Química - Universidade de São Paulo, IQ - USP, Brasil.

Bolsista do(a): Fundação de Amparo à Pesquisa do Estado de São Paulo, FAPESP, Brasil. Orientador: Walter Ribeiro Terra.

*2009 - 2011 Mestrado em Ciências Biológicas.*

Instituto de Química - Universidade de São Paulo, IQ - USP, Brasil.

Bolsista do(a): Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, CAPES, Brasil. Orientador: Walter Ribeiro Terra.

*2004 - 2009 Graduação em Licenciatura em Ciências Biológicas.*

Instituto de Biociências - Universidade de São Paulo, IB - USP, Brasil.

*2004 - 2008 Graduação em Bacharelado em Ciências Biológicas.*

Instituto de Biociências - Universidade de São Paulo, IB - USP, Brasil.

*1999 - 2002 Curso técnico/profissionalizante em Assessoria de Gerenciamento Empresarial.*

Fundação Instituto Tecnológico de Osasco, FITO, Brasil.

### PUBLICAÇÕES

Artigos completos publicados em periódicos

1. RIBEIRO, JOSÉ M. C. GENTA, FERNANDO A. SORGINE, MARCOS H. F. LOGULLO, RAQUEL MESQUITA, RAFAEL D. PAIVA-SILVA, GABRIELA O. MAJEROWICZ, DAVID MEDEIROS, MARCELO KOERICH, LEONARDO Terra, Walter R. FERREIRA, CLÉLIA **Pimentel**, **André C.** BISCH, PAULO M. LEITE, DANIEL C.

DINIZ, MICHELLE M. P. JUNIOR, JOÃO LÍDIO DA S. G. V. DA SILVA, MANUELA L. ARAUJO, RICARDO N. GANDARA, ANA CAROLINE P. BROSSON, SÉBASTIEN SALMON, DIDIER BOUSBATA, SABRINA GONZÁLEZ-CABALLERO, NATALIA SILBER, ARIEL MARIANO ALVES-BEZERRA, MICHELE , et al. ; An Insight into the Transcriptome of the Digestive Tract of the Bloodsucking Bug, *Rhodnius prolixus*. PLoS Neglected Tropical Diseases (Online), v. 8, p. e2594, 2014.

2. Tamaki, Fábio K. ; **Pimentel, André C.** ; DIAS, ALCIDES B. ; CARDOSO, CHRISTIANE ; Ribeiro, Alberto F. ; FERREIRA, CLÉLIA ; Terra, Walter R. . Physiology of digestion and the molecular characterization of the major digestive enzymes from *Periplaneta americana*. Journal of Insect Physiology, v. 70, p. 22, 2014.

3. Tamaki, Fábio K. ; Padilha, Marcelo H.P. ; **Pimentel, Andre C.** ; Ribeiro, Alberto F. ; Terra, Walter R. . Properties and secretory mechanism of *Musca domestica* digestive chymotrypsin and its relation with *Drosophila melanogaster* homologs. Insect Biochemistry and Molecular Biology, v. 42, p. 482-490, 2012.

4. Padilha, Marcelo H.P. ; **PIMENTEL, A. C.** ; Ribeiro, Alberto F. ; Terra, Walter R. . Sequence and function of lysosomal and digestive cathepsin D-like proteinases of *Musca domestica* midgut. Insect Biochemistry and Molecular Biology, v. 39, p. 782-791, 2009.

\*Para acessar o currículo completo e atualizado visite a página: <http://lattes.cnpq.br/7134879912374227>