

# Análise Computacional de Alfa-Glucosidases de Insetos Hematófagos

**EDGAR GUTIERREZ INFANTE**

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE

DARCY RIBEIRO - UENF

CAMPOS DOS GOYTACAZES/RJ

OUTUBRO DE 2013

# Análise Computacional de Alfa-Glucosidases de Insetos Hematófagos

**EDGAR GUTIERREZ INFANTE**

Dissertação apresentada ao Centro e Biociências e Biotecnologia da Universidade Estadual do Norte Fluminense como parte das exigências para obtenção do Título de Mestre em Biociências.

Orientadora: Prof. Dr<sup>a</sup> Marílvia Dansa de Alencar

Co-orientador: Prof. Dr. Thiago Motta Venancio

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE

DARCY RIBEIRO - UENF

CAMPOS DOS GOYTACAZES/RJ

OUTUBRO DE 2013

# Análise Computacional de Alfa-Glucosidases de Insetos Hematófagos

Dissertação apresentada ao Centro de Biociências e Biotecnologia da Universidade Estadual do Norte Fluminense como parte das exigências para obtenção do Título de Mestre em Biociências.

Aprovada em 30 de Outubro de 2013.

Comissão examinadora:

---

Prof. Fernando Ariel Genta. Lab. Bioquímica de Insetos. FIOCRUZ

---

Prof. Jorge Hernandez Fernandez LQFPP - UENF

---

Prof. Beatriz dos Santos Ferreira. LBT – UENF.

---

Prof. Marilvia Dansa de Alencar. LQFPP – UENF (Orientador)

---

Prof. Thiago Motta Venancio LQFPP – UENF (Coorientador)

## **AGRADECIMENTOS**

A minha família, pelo incentivo e apoio que em tudo momento foi importante para continuar no caminho da formação pessoal e profissional.

Aos Professores Marílvia Dansa e Thiago Venancio pela oportunidade de crescimento, disponibilidade e pelas ensinanzas, que foram essenciais para a realização de este trabalho.

Aos professores do laboratório LQFPP, Olga Lima Tavares, Jorge Hernandez, Katya Valesky e Flavia Burgues, pelos ensinamentos enriquecedores.

A meus amigos Maycon Bruno Almeyda, July Narvaez e Alexander Caytuero, pelos muitos momentos compartilhados e pela sua atenção incondicional.

Aos meus colegas do Programa de Pós-graduação de Biociências e Biotecnologia, nos quais me proporcionaram grandes momentos e eternas recordações.

A Marlene e Beatriz, secretarias do PPGBB, que muito atenta e competentemente que foram de enorme ajuda durante meu passo pela UENF.

Por todas as pessoas que, direta ou indiretamente, contribuíram para o desenvolvimento deste trabalho.

## SUMÁRIO

LISTA DE FIGURAS.....	vi
LISTA DE TABELAS.....	vii
ABREVIATURAS.....	viii
RESUMO.....	ix
ABSTRACT.....	x
1. INTRODUÇÃO.....	1
1.1. Ordem Hemiptera .....	2
1.2. Evolução da hematofagia em artrópodes.....	4
1.3. Hemozoína.....	6
1.4. Alfa-glucosidase.....	9
2. OBJETIVOS.....	13
2.1. Objetivos gerais.....	13
2.2. Objetivos Específicos.....	13
3. MATERIAL E MÉTODOS .....	14
3.1. Seqüências protéicas dos organismos selecionados.....	14
3.2. Busca de proteínas com alta similaridade com a alfa-glucosidase Rpro ACI96031 .....	14
3.3. Predição de regiões transmembranares. ....	15
3.4. Alinhamento múltiplo de seqüências.....	15
3.5. Construção da árvore filogenética.....	15
3.6. Alinhamento múltiplo de estruturas.....	15
4. RESULTADOS. ....	16
4.1. Resíduos marcadores das enzimas da família alfa-amilase e das alfa-glucosidases.....	16
4.2. Árvore filogenética das alfa-glucosidases.....	18
4.3. Alinhamento múltiplo das sequências das alfa-glucosidases.....	28
4.4. Alinhamento múltiplo das estruturas das alfa-glucosidases de R. prolixus.....	33
5. DISCUSSÃO.....	36
6. CONCLUSÕES.....	40
7. REFERÊNCIAS BIBLIOGRÁFICAS .....	41

## LISTA DE FIGURAS

Figura 1. Filogenia de insetos vivos e extintos.....	3
Figura 2. Estrutura do grupo heme ou ferro-protoporfirina IX.....	7
Figura 3. Estrutura do dímero de heme e da hemozoina.....	8
Figura 4. Estrutura da alfa-glucosidase de <i>Thermotoga maritima</i> .....	10
Figura 5. Alinhamento estrutural de seqüências de 24 estruturas PDB de alfa-glucosidases..	10
Figura 6. Organização dos domínios na árvore filogenética das alfa-glucosidases.....	19
Figura 7. Região A da árvore filogenética das alfa-glucosidases.....	20
Figura 8. Região B da árvore filogenética das alfa-glucosidases.....	21
Figura 9. Região C da árvore filogenética das alfa-glucosidases.....	22
Figura 10. Região D da árvore filogenética das alfa-glucosidases.....	23
Figura 11. Região E da árvore filogenética das alfa-glucosidases.....	24
Figura 12. Região F da árvore filogenética das alfa-glucosidases.....	25
Figura 13. Região G da árvore filogenética das alfa-glucosidases.....	26
Figura 14. Conservação dos resíduos Y, D, H na árvore filogenética.....	27
Figura 15. Preferências de aminoácido em relação a ligação de íons ferro.....	29
Figura 16. Distribuição dos resíduos H, C, D e E na região 410 – 530 do alinhamento.....	30
Figura 17. Distribuição dos resíduos H, C, D e E na região 370 – 570 do alinhamento.....	31
Figura 18. Estrutura da alfa-glucosidase de <i>R. prolixus</i> Rpro ACI90631.1.....	32
Figura 19. Alinhamento múltiplo das estruturas das alfa-glucosidases de <i>R. prolixus</i> .....	34
Figura 20. Alinhamento múltiplo das seqüências das alfa-glucosidases de <i>R. prolixus</i> .....	35

## LISTA DE TABELAS

Tabela 1. Aminoácidos conservados das enzimas da família alfa-amilase:.....	16
Tabela 2. Resíduos característicos das alfa-amilases com diferente especificidade..	17
Tabela 3. Características gerais da árvore filogenética das alfa-glucosidases.....	28
Tabela 4. Propriedades bioquímicas das alfa-glucosidases de <i>R. prolixus</i> .....	33
Tabela 5. Template para geração das estruturas das alfa-glucosidases .....	33

## **ABREVIATURAS**

**Aglu:** Alfa-glucosidase

**Hz:** Hemozoina

**MPMV:** Membrana Perimicrovilar

**RNAi:** RNA de interferência

**GH:** Glicosil Hidrolases

**HMMER:** Hidden Markov Models Software

**PDB:** Protein Data Base

**Rpro:** Rhodnius prolixus

**PFAM:** Protein Family Database



## RESUMO

Insetos são as causas de alta mortalidade e morbidade em populações humanas, sobretudo como resultado da transmissão de patógenos por espécies hematófagas. O combate a doenças transmitidas por estas espécies é, portanto, um desafio que requer novas e eficientes estratégias. Abordagens lógicas para essas iniciativas devem ser baseadas em estudos de fisiologia, bioquímica e biologia molecular de insetos vetores de doenças. *Rhodnius prolixus* é um hemíptero que pertence à família Reduviidae. Este inseto tem atividade de síntese de hemozoína (Hz) na fração protéica de epitélio intestinal. O presente trabalho tem por objetivo comparar alfa-glucosidases de *R. prolixus* com alfa-glucosidases de outros organismos utilizando ferramentas de bioinformática, buscando compreender a evolução destas enzimas e que características estruturais podem ser importantes na atividade formadora de Hz. A partir da utilização do programa HMMER e com base no domínio conservado de alfa-glucosidases, foram detectadas e analisadas estas enzimas em diversas espécies. As seqüências destas proteínas foram utilizadas na construção filogenética da família das alfa-glucosidases. A tríade catalítica das alfa-glucosidases (Asp107, Glu175, Asp242 na proteína de *R. prolixus* Rpro|ACI96031.1) são resíduos invariáveis entre os membros do clã GH-H. A ausência de algum destes resíduos indica que a proteína pode ser considerada inativa. Duas das sete regiões da árvore filogenética caracterizam-se por apresentarem a maior quantidade de alfa-glucosidases com a tríade catalítica completa, além disso, estas regiões têm totalmente conservados os resíduos YDH, ainda mais conservados que os resíduos catalíticos. Também foi encontrada no alinhamento múltiplo de seqüências uma alta concentração de resíduos de ácido aspártico e ácido glutâmico, resíduos conhecidos por interagir com o átomo de ferro.

**Palavras-chave:** hemozoína, alfa-glucosidase, *R. prolixus*, evolução, bioinformática.

## ABSTRACT

Insects are causes of high mortality and morbidity in human populations, primarily as a result of pathogen transmission by sucking species. The fight against diseases transmitted by these species is therefore a challenge that requires new and efficient strategies. Logical approaches for these initiatives should be based on studies of physiology, biochemistry and molecular biology of insect disease vectors. *Rhodnius prolixus* is a hemiptera belonging to the family Reduviidae, this insect has synthesis activity of hemozoin (Hz) in the protein fraction of the intestinal epithelium. The main object of this work is to compare alpha-glucosidase from *R. prolixus* with alpha-glucosidases from other organisms using bioinformatics tools, in order to understand the evolution of these enzymes searching for sequences and structural features that may be important to understand hemozoin activity. HMMER program was used for detect and analyze alfa-glucosidases in several species, based on the conserved domain of alpha-glucosidase. The sequences of these proteins were used in the construction of the phylogenetic tree family of alpha-glucosidases. The catalytic triad of alpha-glucosidase (Asp107, Glu175, Asp242 in *R. prolixus* protein Rpro | ACI96031.1) are invariant among members of the clan GH-H. The absence of some of these residues indicates that the protein can be considered inactive. Two of the seven regions of the phylogenetic tree characterized by presenting the greatest amount of alpha-glucosidases with the complete catalytic triad. These regions have the conserved residues YDH, and also the catalytic residues are fully conserved too. It was also found in the multiple sequence alignment a high prevalence of aspartic and glutamic acid residues. These residues are known to interact with the iron atom in many proteins.

**Keywords:** hemozoin, alpha-glucosidase, *Rhodnius prolixus*, bioinformatics.

## 1. INTRODUÇÃO

A classe dos insetos compreende mais de um milhão de espécies. Destes, aproximadamente 14.000 espécies, distribuídas em 400 diferentes gêneros, têm o hábito de se alimentar de sangue de vertebrados. A hematofagia é uma característica polifilética tendo evoluído várias vezes ao longo da evolução destes artrópodes. Embora não se conheça o momento exato em que os primeiros insetos hematófagos surgiram, há uma tendência em ajustar a origem da maioria dos grupos no período Cretáceo (Lukashevich e Mostovski, 2003; Mans e Neitz, 2004).

Em insetos a evolução do hábito de sugar sangue parece ter ocorrido através de dois caminhos principais: pela presença de partes bucais pré-adaptadas para perfurar e através da associação prolongada entre insetos e vertebrados (Lehane, 2005). Nesta associação que deve ter ocorrido em ninhos, os insetos inicialmente se alimentavam de fezes, fungos e restos de vertebrados provenientes das mudas, penas ou pelos. A ingestão regular ou acidental desses materiais provavelmente favoreceu a seleção de indivíduos que poderiam utilizar eficientemente este material (Lehane, 2005).

Uma vez que os insetos passaram a adquirir sangue de forma regular, é provável que seu alto valor nutricional tenha direcionado o desenvolvimento de um grupo de insetos que regularmente exploravam sangue como fonte de alimento. Isto teria se desenvolvido progressivamente, através de adaptações fisiológicas, comportamentais e morfológicas, inicialmente para hematofagia facultativa e, eventualmente, em alguns insetos, para hematofagia obrigatória (Marshall, 1981).

Uma segunda hipótese explica o surgimento de insetos hematófagos a partir de espécies de insetos que se alimentavam de seiva de plantas. Os sugadores de seiva podem ter perdido suas proteases digestivas, pois estas representariam uma dificuldade alimentar para estes insetos. Dessa forma, seletivamente, alguns destes genes foram perdidos em algumas linhagens de fitófagos. Posteriormente, para mudarem seu hábito alimentar para a hematofagia, estes insetos precisaram readquirir sua atividade proteolítica. Sugere-se que, tendo perdido suas tripsinas, estes insetos tivessem que fazer uso de catepsinas intracelulares, redirecionando-as para digestão extracelular (Lehane, 2005). Hemípteros fazem uso de proteases do

tipo catepsinas no seu processo digestivo, com grande eficiência, isto é consistente com a hipótese de que tenham surgido a partir de sugadores de seiva vegetal (Houseman *et al.*, 1985; Billingsley e Downe, 1989; Terra, 1988). Entre os hematófagos encontram-se ainda insetos que apresentam serino-proteases ativas. Isto pode ter ocorrido devido à manutenção do gene e posterior ativação deste, com a evolução da hematofagia.

### 1.1. Ordem Hemíptera

Na classe Insecta, o grupo Paraneoptera inclui duas superordens: Psocodea, que compreende as ordens Psocoptera (traças e psocóides) e Phthiraptera (piolhos); e a superordem Condylgnatha, que inclui a ordem Hemiptera. Na ordem Hemiptera se encontram barbeiros, percevejos, pulgões e cigarras e é a ordem de Paraneoptera, a mais bem sucedida e mais numerosa (Silva *et al.*, 2004).

A ordem Hemiptera possivelmente surgiu no Carbonífero (**Figura 1**), tendo sua maior diversificação no Mesozóico, aliada à emergência das angiospermas, já que as peças bucais destes insetos são altamente eficientes na extração de fluidos vegetais, visto que teriam as partes bucais pré-adaptadas ao hábito picador-sugador. A evolução das partes bucais foi de crucial importância para possibilitar que se desenvolvessem no grupo, o hábito de se alimentar de sangue (Lehane, 1991).

A ordem Hemiptera é dividida em duas subordens: Homoptera e Heteroptera, que diferem, sobretudo, quanto a seus hábitos alimentares (Silva *et al.*, 1995). Homoptera compreende insetos que se alimentam principalmente de seiva de plantas, enquanto insetos da subordem Heteroptera possuem uma dieta mais variada. A ordem Hemiptera possui muitas espécies de grande importância médica e veterinária. No entanto, esta ordem também possui membros que são considerados pragas ou vetores de patógenos de plantas, como é o caso do *Dysdercus peruvianus*.

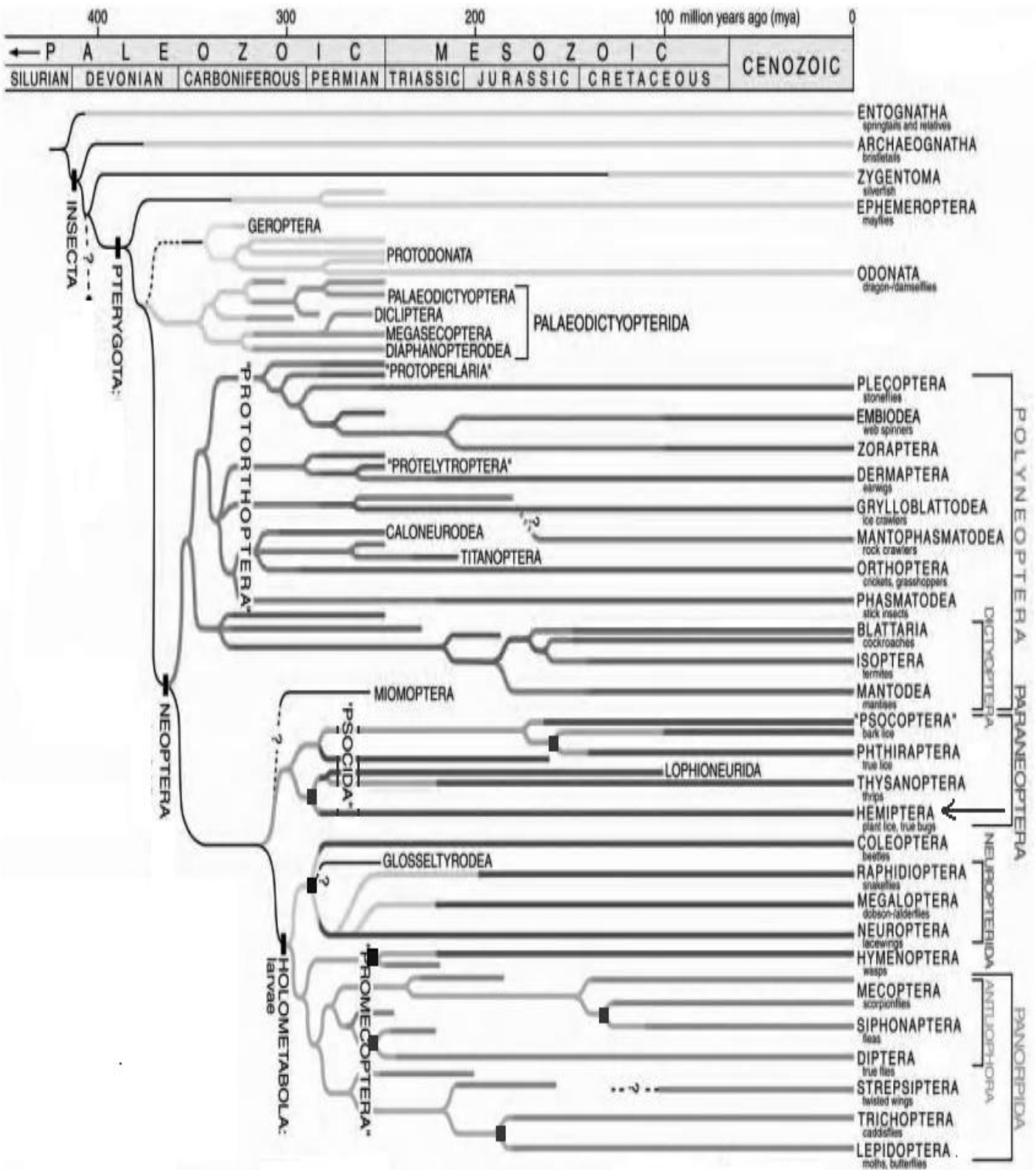


Figura 1. Filogenia de insetos vivos e extintos (Kristensen, 1999; Boudreaux, 1979).

A matriz peritrófica consiste em uma estrutura anatômica quitino-protéica que envolve o bolo alimentar da maioria dos insetos, levando à compartimentalização do processo digestivo (Terra *et al.*, 1979; Terra e Ferreira, 1994). Hemípteros diferem

dos outros insetos por possuírem membrana perimicrovilar (MPMV), que consiste em uma membrana extracelular de natureza fosfolipídica que envolve as microvilosidades do intestino (Terra, 1988). As membranas perimicrovulares foram descritas tanto em Heteroptera, quanto em Fulgoroidea, superfamília de Homoptera onde se encontram cigarras, pulgões, cigarrinhas e colchonilhas (Reger, 1971; Lane e Harison, 1979; Andries e Torpier, 1982; Baerwald e Decarpio, 1983).

Entre vários grupos de insetos paraneopteros, somente os insetos das ordens Hemiptera e Thysanoptera possuem membranas perimicrovulares, contrastando com Psocoptera e Phthiraptera que não possuem. Isto sugere que as membranas perimicrovulares apareceram inicialmente em um ancestral Condylgnatha (Silva *et al.*, 2004). Acredita-se que o desenvolvimento das membranas perimicrovulares ocorreu pela necessidade de uma maior eficiência no processo de absorção de aminoácidos essenciais, já que estes se encontravam em baixíssimas concentrações no floema (Terra, 1990). MPMV estão em contato direto com o bolo alimentar, separando este do epitélio intestinal. MPMV se estendem da base das microvilosidades em direção ao lúmen e delimitam um compartimento fechado, o espaço perimicrovilar. Este espaço é localizado entre a membrana microvilar e a membrana perimicrovilar. A função dessa membrana é a compartimentalização do processo digestivo, aumento da eficiência da absorção de aminoácidos, imobilização de algumas enzimas e proteção química e mecânica do epitélio intestinal (Terra e Jordão, 1991).

## 1.2. Evolução da hematofagia em artrópodes

Em termos de número de espécies conhecidas, Arthropoda é o filo mais bem sucedido dos metazoários. Somente insetos apresentam mais de um milhão de espécies. Dessa enorme diversidade genética aproximadamente, 14.000 espécies distribuídas em 400 diferentes gêneros, têm o hábito de ingerir sangue de vertebrados (Ribeiro, 1995).

A hematofagia surgiu várias vezes ao longo da evolução dos artrópodes (Ribeiro, 1995). É possível que a hematofagia tenha surgido no início da divisão de mamíferos e aves, acompanhadas pela diversificação e expansão dos animais terrestres. Antecedentes de hematófagos modernos, frequentemente têm várias

características que são pré-condições para hematofagia. Algumas dessas pré-adaptações são de natureza ecológica e a proximidade com os vertebrados foi principalmente o que facilitou o desenvolvimento dessas características (Lehane, 2005).

A associação persistente entre insetos e vertebrados parece ter se iniciado através da atração de insetos para os ninhos e tocas dos vertebrados onde encontrariam calor e umidade. Em áreas semi-áridas ou áridas, o habitat oferecido pelos ninhos pode ter sido também de fundamental importância para a sobrevivência dos insetos, já que estes locais apresentavam um abundante suprimento de nutrientes (Pearman, 1960). Nesta associação com os ninhos, os insetos inicialmente se alimentavam de fezes, fungos e restos de vertebrados provenientes das mudas, penas ou pelos. A ingestão regular ou acidental desses materiais provavelmente favoreceu a seleção de indivíduos que poderiam utilizar eficientemente este material (Lehane, 2005). Adaptações morfológicas foram claramente necessárias, como a modificação das peças bucais para o desenvolvimento da hematofagia (Emmerson *et al.*, 1973). As partes bucais destes insetos se desenvolveram para alimentação baseada inicialmente em matéria orgânica, sendo quase em sua maioria mastigadoras.

Uma vez que os insetos passaram a adquirir sangue de forma regular, é provável que seu alto valor nutricional tenha favorecido o desenvolvimento de espécies que regularmente exploravam sangue como fonte de alimento. Isto teria se desenvolvido progressivamente, através de novas adaptações fisiológicas, comportamentais e morfológicas, inicialmente para hematofagia facultativa e, eventualmente, em alguns insetos, para hematofagia obrigatória (Marshall, 1981).

A associação de insetos e vertebrados pode ter sido estabelecida ainda devido à atração dos insetos pelas secreções do hospedeiro vertebrado, ou pelo uso do vertebrado como sítio de aquecimento (Lehane, 2005). Mas, provavelmente o fator mais importante, foi o uso dos excrementos do hospedeiro como hábitat larval. *Haematobia irritans* é um bom exemplo de colonizador de fezes para a criação de larvas. Estas só deixam o hospedeiro vertebrado para realizar a oviposição nas fezes do mesmo (Mohr, 1943). Este tipo de adaptação permite que a fêmea adulta, durante a dispersão de suas larvas, não se afaste do seu hospedeiro. Assim, o aparecimento do hábito de se alimentar de sangue em fêmeas pode ser explicado

pelas vantagens oferecidas, não só pelo fato do sangue ser altamente nutritivo, mas também devido ao uso do ambiente do hospedeiro para oviposição.

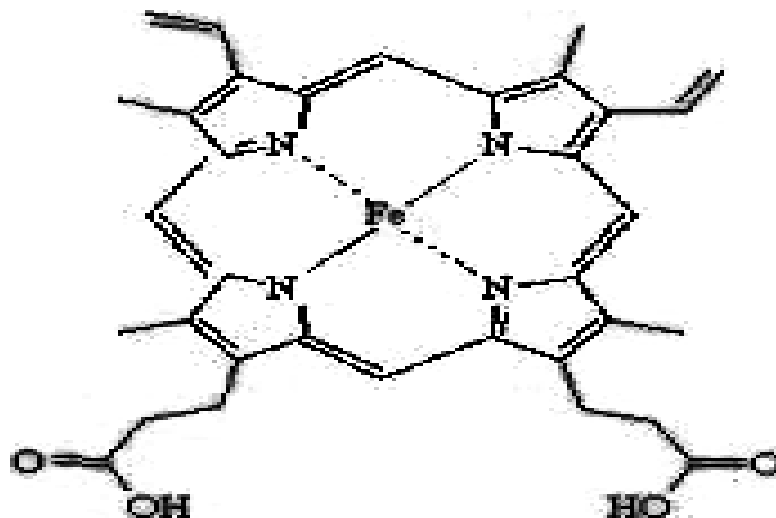
Nos machos, mesmo aqueles que não se alimentam de sangue, a questão pode ser explicada pelo fato de que a associação com o vertebrado minimiza o tempo e a energia investidos na procura da fêmea. Assim, estando o macho em associação com o vertebrado, mesmo sem se alimentar de sangue, ele garante que a fêmea que se alimenta de sangue chegará ao local de alimentação, onde ele se encontra.

A segunda teoria acerca da rota que possibilitou a evolução da hematofagia se baseia no desenvolvimento a partir de alguns insetos ancestrais que possuíam a morfologia das partes bucais pré-adaptadas para perfurar. Insetos entomófagos são fortes candidatos a esta conversão. Rhagionidae é uma família onde se encontram insetos predadores de outros insetos. Nesta família algumas poucas espécies se tornaram hematófagas. Estes entomófagos seriam atraídos aos ninhos pelos insetos presentes, e dessa forma, começariam a também ter contato com vertebrados. Estes insetos predadores teriam adaptações morfológicas e fisiológicas, tais como enzimas digestivas eficientes e partes bucais adaptadas, que facilitariam uma mudança para o hábito de se alimentar de sangue. Este hábito inicialmente estaria relacionado a um evento ocasional, podendo ser modificado para alimentação exclusivamente sanguínea, pela continuada associação com vertebrado (Lehane, 2005).

### 1.3 Hemozoína

A proteína mais abundante no sangue de mamíferos é a hemoglobina (Hb), constituindo uma importante fonte de aminoácidos para insetos hematófagos. A degradação da hemoglobina pelo sistema digestório desses insetos resulta na liberação de altas concentrações de heme, o grupo prostético da hemoglobina. O grupo heme, ou ferro-protoporfirina IX (**Figura 2**), pertence à família das porfirinas que são moléculas biologicamente importantes relacionadas aos processos de captura e transdução de energia. Essas moléculas são formadas por um macrociclo composto de quatro anéis pirrólicos unidos por ligações metilênicas (Moreira *et al.*, 2004).





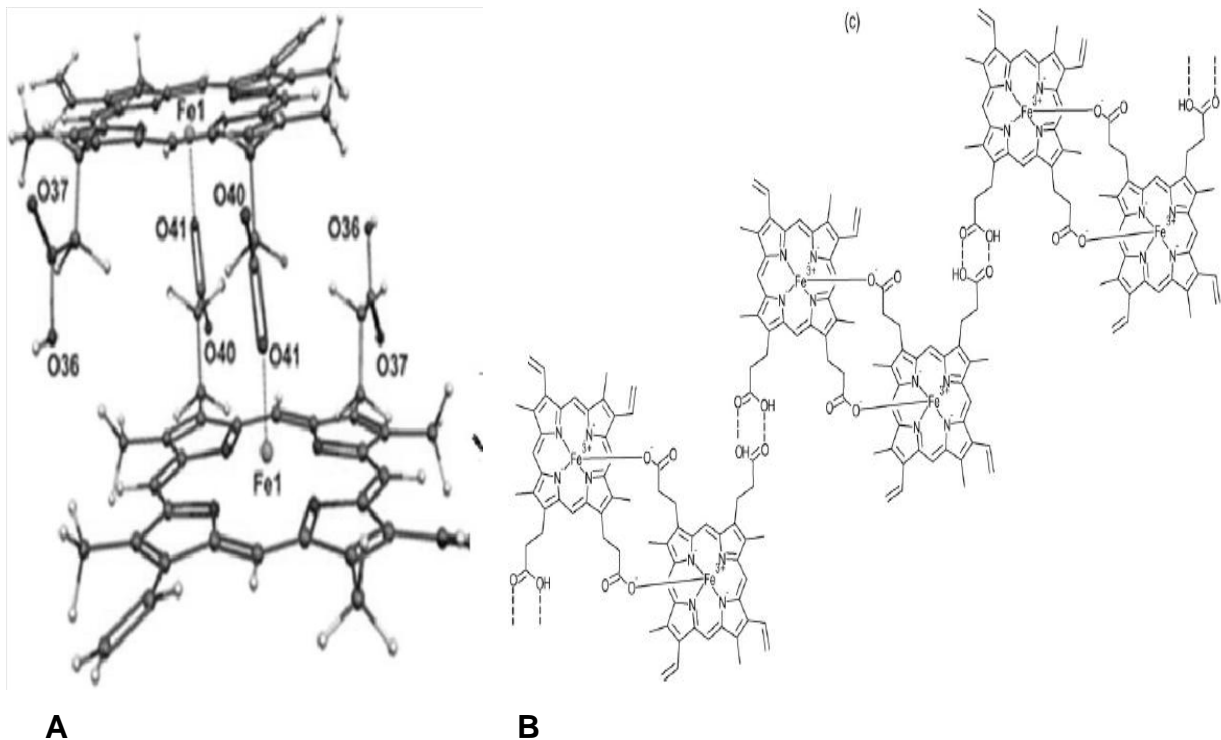
**Figura 2:** Estrutura do grupo heme ou ferro-protoporfirina IX.

A liberação de grandes concentrações de heme na sua forma monomérica, durante a digestão da hemoglobina no intestino médio de espécies hematófagas poderia ser danosa, já que a permeabilidade e a seletividade de membranas celulares são alteradas na presença desta molécula (Smith *et al.*, 1993). A toxicidade do heme está relacionada também a sua capacidade de promover a geração de espécies reativas de oxigênio (Liu *et al.*, 1985; Gutteridge e Smith, 1988) que causam danos oxidativos a lipídios, proteínas e DNA (Aft e Muller, 1984; Vicent *et al.*, 1988).

Devido ao caráter tóxico das moléculas de heme, vários mecanismos de detoxificação dessa molécula foram desenvolvidos ao longo da evolução, por diferentes organismos hematófagos. Estes mecanismos incluem a detoxificação de heme por degradação enzimática (Paiva-Silva *et al.*, 2002, 2006), a ligação de heme a proteínas antioxidantes (Dansa Petretski *et al.*, 1995), a ação de antioxidantes que previnem a formação de espécies reativas de oxigênio (Paes e Oliveira, 1999; Graça-Souza *et al.*, 1999) e a formação de hemozoína (Hz) (Oliveira *et al.*, 1999).

A hemozoína possui em sua estrutura uma molécula de heme associada à outra através de ligações iônicas entre o átomo de Fe de uma molécula de heme e o grupo carboxila dissociado de um grupamento propionato do anel porfirínico de outra molécula de heme, como descrito por espectrometria infravermelha e de absorção de raios X (Slater *et al.*, 1991). A estrutura química da hemozoína corresponde a um

dímero inicial (**Figura 3A**) que vai gerar um cristal composto de vários dímeros associados unidos por ligações de hidrogênio (Egan, 2008) (**Figura 3B**).



**Figura 3: Estrutura do dímero de heme (Frosch *et al.*, 2009) A e da hemozoina (Egan, 2008) B.**

Nesta estrutura ocorre ligação entre o ferro de uma molécula de heme com o radical carboxilato da molécula de heme adjacente. Também ocorrem ligações de pontes de hidrogênio que auxiliam na ramificação da cadeia.

Em *Aedes aegypti* a associação da heme com a matriz peritrófica parece diminuir a toxicidade desta molécula no intestino do inseto mas esta espécie não parece ser capaz de formar Hz *in vivo* (Pascoa *et al.*, 2002). No *R. prolixus* trabalhos anteriores do nosso grupo tem demonstrado a capacidade de associação do heme, gerado pela degradação da hemoglobina, com as membranas perimicrovilares (Oliveira *et al.*, 2000; Silva, 2001, Silva *et al.*, 2007). Homogenatos de *Plasmodium* (Slater *et al.*, 1991) e uma fração particulada do conteúdo intestinal de *R. prolixus* (Oliveira *et al.*, 2000) são capazes de promover a formação de Hz *in vitro*.

Silva *et al.* (2007) mostraram que a formação de hemozoina em *R. prolixus* é dependente das membranas perimicrovilares. Essas membranas possuem  $\alpha$ -glucosidases como proteínas majoritárias e, por isso, estas enzimas têm sido utilizadas como marcadores moleculares de membranas perimicrovilares em

Hemiptera (Silva *et al.*, 1995). Estas enzimas parecem ser responsáveis pela formação de Hz no intestino médio de *R. prolixus* (Mury *et al.*, 2009).

Em alguns trabalhos com *Plasmodium falciparum* foi proposto que a atividade de formação de Hz estava relacionada à presença de uma enzima que foi denominada heme polimerase (Chou e Fitch *et al.*, 1992; Slater e Cerami, 1992). Posteriormente no *Plasmodium*, outra proteína, denominada HDP (heme detoxification protein) foi descrita como tendo um papel importante na conversão de heme em Hz (Jani *et al.*, 2008).

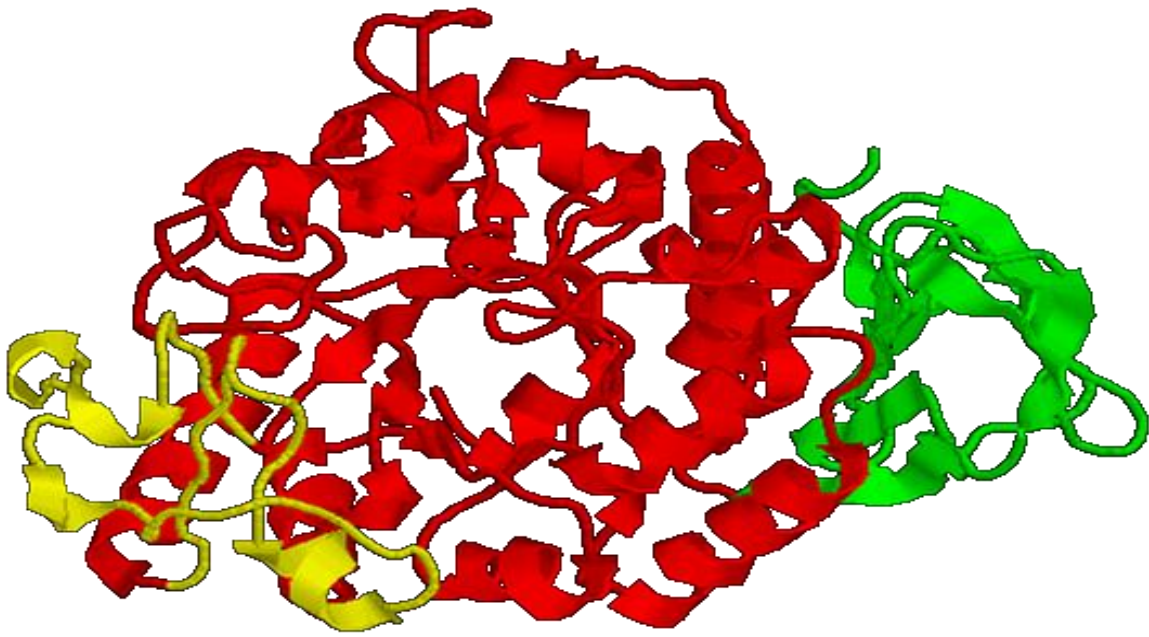
Tanto em *Plasmodium* como em outros organismos produtores de Hz, o mecanismo de biogênese de Hz é ainda pouco conhecido e sujeito a intensos debates (Dorn *et al.*, 1995; Fitch *et al.*, 1999; Tekwani e Walker, 2005). Alguns autores sugeriram que a formação de Hz era, na verdade, um processo dependente da presença de lipídeos específicos (Bendrat *et al.*, 1995 e Dorn *et al.*, 1995), outros de proteínas ricas em histidina (Sullivan *et al.*, 1996) e outros ainda que a própria Hz pré-formada poderia promover a formação de Hz, em uma reação autocatalítica sem o envolvimento de proteínas ou lipídios do parasita (Dorn *et al.*, 1995).

#### 1.4 Alfa-glucosidases

Glucosidases (EC 3.2.1) são enzimas que catalisam reações de hidrólise de ligações O- ou S-glicosil, e estão envolvidas com uma grande variedade de eventos biológicos (Zechel e Whitters, 2000). Entre as glucosidases se encontram as  $\alpha$ -glucosidases ou  $\alpha$ -D-glucosideo-glucohidrolases (EC 3.2.1.20) Estas enzimas catalisam a hidrólise de ligações glicosídicas do tipo  $\alpha$ -1,4 das extremidades não redutoras de oligo ou polissacarídeos liberando  $\alpha$ -D-glucose com eficiência variada.

As alfa-glucosidases tem o domínio catalítico A com a estrutura de barril (alfa/beta)<sub>8</sub> que é uma das estruturas mais utilizadas na natureza pelas proteínas (**Figura 4**).

O domínio A das alfa-glucosidases tem a tríade catalítica composta por dois ácidos aspárticos e um ácido glutâmico (**Figura 5**), resíduos que são absolutamente conservados nas proteínas ativas. Os domínios B e C podem ser de diversos tamanhos ou inclusive, como no domínio C terminal, pode ser ausente em algumas alfa-glucosidases (Kumar, 2010).



**Figura 4: Estrutura da alfa-glucosidase de *Thermotoga maritima*. PDB 1OBB.** Os domínios individuais são marcados: domínio A (vermelho), consiste em uma estrutura de barril (alfa/beta)<sub>8</sub>, onde estão os resíduos catalíticos, o domínio B (amarelo), e domínio C terminal (verde). (Pujadas, 2001)

	52	63	92	106	223	232	256	263	320	330
1E3Z.pdb	NGGPYDLYDLG	NVQVIGD/VVNRKAG	LSLDGF	RI	DA	MFTVAKYK	SVTVYONHDIQ			
1A47.pdb	HGWARDFKRTN	NIRVILDFARNHTSP	MGIDGI	RL	DA	VFTFGGEWF	MWTFYONHDMQ			
1BG9.pdb	QGMPGRLYDLG	GVKAIADIVINRRTA	IGFDGW	RF	DF	SFAVAEIKW	AVTVYONHDSG			
1G94.pdb	TRQPVSYEL-Q	GVVIVDVLNRMAA	IGVKGF	RF	DA	PVVFGQEV	AVVTVYONHDSQ			
1GVI.pdb	HKOTADYFEID	GIRVILDAVENRCCY	FDIDGW	RL	DV	VYILGLIW	AFNLLSHTDTP			
1J0J.pdb	HKOTADYFEVD	GIRVILDAVENRCCY	FDIDGW	RL	DV	VYILGEIKW	AFNLLSHTDTS			
1JDC.pdb	EGFWRDFN-K-	GVVIVD/VVENRMR	YGAGGF	RF	DF	SFCVGLQW	AVTVYONHDSG			
1LMJ.pdb	HGQVDFYSFK	GIRVILDLPHRRGPF	MSVDGF	RF	DA	GIFLAEIKW	FVNFYONHDSMS			
1OSE.pdb	ERQPVSYKLC-C	GVVIVD/VVENRMCQ	IGVAGF	RI	DA	PFIFQAEVI	ALVTVYONHDSQ			
1QHO.pdb	HGWRDFPKQIE	GIRVIVD/VVENRSTP	HGADGL	RI	DA	IFLVGEIKW	LITVYONHDSMS			
1RP8.pdb	EGMPGRLYDID	GVAIADIVINRCA	LGFDAM	RL	DF	SLAVAEIKW	AVTVYONHDSG			
1UA7.pdb	WLQPTSYQIGN	GIRVIVDAVENRRTF	DGADGF	RF	DA	EFOYGEIKL	LVTVYONHDSY			
1UH3.pdb	HKOTQDYMAVD	KGIVILDAVENRAGD	YSVDGW	RL	DA	AAIIGERW	MGNLLSNDHIT			
1V3M.pdb	HGWARDFKRTN	NIRVILDFARNHTSP	LGVDGI	RV	DA	VFTFGGEWF	QVTVYONHDMQ			
1W9X.pdb	VGSAVDLYDLG	GVVIVD/VVNRKGG	LNLDGI	RI	DA	MFAVAEIKW	AVTVYONHDSQ			
1WPC.pdb	VGSAVDLYDLG	GIRVIGD/VVNRKGG	LGLDGF	RI	DA	MFAVAEIKW	AVTVYONHDSQ			
1XH0.pdb	ERQPVSYKLC-C	GVVIVD/VVENRMCQ	IGVAGF	RL	DA	PFIVQAEVI	ALVTVYONHDSQ			
1Z32.pdb	ERQPVSYKLC-C	GVVIVD/VVENRMCQ	IGVAGF	RI	DA	PFIVQAEVI	ALVTVYONHDSQ			
2D20.pdb	HKOTADYLAIQ	GIRVILDAVENRAGD	QGIDGW	RL	DV	ALIVGEIKW	LNLLSHTDTE			
2DIJ.pdb	HGWARDFKRTN	NIRVILDFARNHTSP	LGIDGI	RM	DA	VFTFGGEWF	QVTVYONHDMQ			
2FHF.pdb	WGDPFHYTVPE	GMVIVD/VVNRRTNA	YKIDGF	RF	DL	IYFFGEIKW	VVNVYONHDSQ			
2GVY.pdb	HGWRQDIYSLN	GMVIVD/VVNRMCY	YSIDGL	RI	DT	VYCIQAEVL	LSTVYONHDSQ			
6CGT.pdb	HGWARDFKRTN	GIRVILDFARNHTSP	MSVDGI	RV	DA	VFTFGGEWF	QVTVYONHDMQ			
7TAA.pdb	HGWRQDIYSLN	GMVIVD/VVNRMCY	YSIDGL	RI	DT	VYCIQAEVL	LSTVYONHDSQ			

**Figura 5. Alinhamento estrutural de seqüências baseadas em 24 estruturas PDB de alfa-amilase.** Proteínas analisadas com o programa MUSTANG. Os números dos resíduos importantes na atividade alfa-glucosidase são marcados de acordo com a estrutura 1E3Z. Resíduos que interagem com o substrato (Y54, D98), resíduos que estabilizam o estado de transição (H103, H327), tríade catalítica (D231, E261, D328) (Kumar, 2010).

A partir de estudos estruturais já publicados (McGregor et al., 2001) com enzimas mutantes, substratos e inibidores, foram identificados os aminoácidos conservados que são característicos de especificidade da família alfa-amilase (**Tabela 1**). A **Tabela 1b** mostra o alinhamento estrutural de várias proteínas da família alfa amilase, as letras X significam que o resíduo nas proteínas é hidrofóbico, B o resíduo é hidrofílico e Z indica que o resíduo é importante para a especificidade da proteína. A **Tabela 1a** mostra a posição dos resíduos da **Tabela 1b** na posição relativa da proteína Taka-amilase (XP\_001821436.1).

a.							
I	116	VDVVANH	122	II	202	GLRIDTVKH	210
III	226	YCIGEVLD	233	IV	292	FVENHD	297
b. More generally							
I		XDXXXNH		II		GXRDXZZ	
III		XXX(G/A)EZZZ		IV		XXBBHD	

**Tabela 1. Aminoácidos conservados das enzimas da família alfa-amilase:** a. para a proteína Taka amilase A (XP\_001821436.1). b. Alinhamento estrutural de um conjunto de proteínas da família alfa-amilase com resíduos conservados indicados alternativamente com as letras **X**, resíduo hidrofóbico; **B**, resíduo hidrofílico; **Z**, resíduo importante para a especificidade (McGregor et al., 2001)

A maioria das alfa-amilases, têm resíduos aromáticos e hidrofóbicos na posição 210 equivalente na Taka-amilase (**Tabela 2**), que podem interagir com açúcares. As outras posições podem servir para identificar as proteínas de acordo a sua função (McGregor et al., 2001).

Glucosidases são as primeiras enzimas da digestão nos insetos sugadores de seiva (Morgan, 1975), mas a presença de uma alta atividade dessas enzimas em insetos hematófagos, como *R. prolixus*, sugere que estas são também importantes para insetos que se alimentam de sangue (Silva et al., 2004). Em *Phlebotomus langeroni* foi observado que a atividade de  $\alpha$ -glucosidase aumentou significativamente 1h após a alimentação com sangue (Dillon e Kordy, 1997).

Enzyme	Residues at position number (Taka-amylase A numbering)											No. of sequences	
	203	205	207	208	209	210	231	232	233	234	236		237
Enzymes acting on sucrose													
Amylosucrase			A	V	A	F							1
Glucan sucrose			A	V	D	N	M/D/A						14
Sucrose phosphorylase	A		A/V	G	Y/F								4
	A		F	A	Y								2
Cyclomaltodextrinase <sup>p</sup>			V	A	N	E						No K/L	4
			C	G	H	D							2
<b>α-Glucosidases</b>			A	L/V	P	Y/H		Y	S/T				5
			A/V	G/I	S	H/L/M							3
			O/V	A/L/V	H	G/L/M							4
Enzymes acting on α-1,6-bonds only													
Oligo-1,6-glucosidase			V	I	G/N	F/M/S							5
			V	L	W	L							1
Dextran glucosidase			V	I	D	L/M							2
	No L		V	I	N	L		A		I			1
Pullulanase			L	M	A	L							1
			V	P	N	E							1
(Limit dextrinase)			L/M/Q	L/M	G	D/H/I/L/Y							13
Isoamylase			L	A	S	I/V							6
Enzymes acting on α-1,4-bonds only													
Maltooligosyltrehalose trehalohydrolase			A	V	H	A							6
Maltosyl transferase			M	G	H	A	E						1
<b>CGTases</b>	I		A	V	K	H	W						18

**Tabela 2. Resíduos característicos das alfa-amilases com diferente especificidade.**  
Em destaque as alfa-glucosidases (McGregor et al., 2001).

As alfa-glucosidases são encontradas no intestino médio dos insetos, na forma solúvel ou ligadas a membranas. As α-glucosidases de membrana são proteínas integrais da membrana microvilar em células intestinais de Diptera (Espinoza-Fuentes & Terra 1986; Terra & Jordão, 1989, 1991). Na ordem Hemiptera, uma α-glucosidase é encontrada associada às membranas perimicrovilares, que envolvem as microvilosidades e se projetam para a luz do tubo digestivo, algumas vezes formando múltiplas camadas sobre as microvilosidades das células do intestino médio, provavelmente em todos os hemípteros (Terra, 1988; Silva *et al.*, 2004). Em alguns hematófagos, as membranas perimicrovilares parecem ser sintetizadas em resposta à alimentação com sangue. Em *R. prolixus* a atividade desta enzima aumenta significativamente nos quatro primeiros dias após o repasto

sangüíneo. Este aumento coincide com o aumento da Hz no lúmen intestinal desta espécie, sugerindo a participação dessas membranas na síntese do cristal de heme (Silva et al., 2007). Provavelmente, a pressão evolutiva determinou uma função diferente para esta enzima nestes animais, visto que, proteína, e não carboidrato é o nutriente majoritário presente nas dietas de insetos sugadores de sangue.

O papel da  $\alpha$ -glucosidase intestinal na síntese de Hz em *R. prolixus* foi demonstrado por nosso grupo (Mury et al, 2009). Este trabalho mostrou que a enzima promove a síntese de Hz, participando provavelmente na etapa de nucleação de heme, já que a adição de inibidores enzimáticos não bloquearam o processo depois de transcorridas 4 horas depois do início da síntese de hemozoina. Fêmeas de *R. prolixus* que foram injetadas com dupla fita de RNA, construída com base na seqüência de uma  $\alpha$ -glucosidase intestinal de *Anopheles aquasalis* apresentaram uma consistente redução na formação de Hz.

Adicionalmente nosso grupo também mostrou que uma  $\alpha$ -glucosidase da membrana perimicrovilar de *Dysdercus peruvianus*, um hemíptero sugador de sementes de algodão, é capaz de catalisar a síntese Hz *in vitro* e que ao ser artificialmente alimentado com heme, este hemíptero sintetiza cristais de heme com características espectroscópicas semelhantes à Hz, também *in vivo*. Além disso, se é oferecido sangue a fêmeas em jejum, este inseto é capaz de se alimentar de sangue sem alterações significativas de ovoposição e mortalidade, mostrando que o grupo apresenta as pré-adaptações necessárias à hematofagia (Calegário et al, em preparação). Entre estas pre-adaptações, é possível que  $\alpha$ -glucosidases deste grupo de insetos já apresentem características estruturais que permitem que elas sejam capazes de desempenhar uma nova função e capacitar estes organismos a vencerem o desafio evolutivo da hematofagia.

## 2. OBJETIVOS

### 2.1. Objetivo Geral

Como meta central do presente trabalho esta o análise computacional de alfa-glucosidases de insetos, com especial atenção aos grupos de insetos hematófagos.

### 2.2 Objetivos Específicos

- Análise de resíduos marcadores que identifiquem a atividade catalítica da proteína Rpro|ACI96031 de *R. prolixus* como parte da subfamília das alfa-glucosidases (EC. 3.2.1.20) das glucosilhidrolases.
- Análise da organização e distribuição da árvore filogenética das alfa-glucosidases homólogas à alfa-glucosidase Rpro|ACI96031 de *R. prolixus*.
- Análise do alinhamento múltiplo de seqüências das alfa-glucosidases de insetos hematófagos e não hematófagos com o fim de encontrar resíduos ou regiões conservadas.
- Análise do alinhamento múltiplo das estruturas das alfa-glucosidases de *R. prolixus* com o fim de encontrar estruturas conservadas.



### 3. MATERIAL E MÉTODOS

#### 3.1. Seqüências protéicas dos organismos selecionados

As duas principais características dos organismos selecionados para o presente trabalho são a posição filogenética e a disponibilidade de genoma sequenciado. Também foram escolhidos preferencialmente insetos de grupos hematófagos e não hematófagos e outros organismos não insetos. Os organismos selecionados e acrônimos foram:

1. *Aedes aegypti*, *Aaeg*.
2. *Anopheles gambiae*, *Agam*.
3. *Apis mellifera*, *Amel*.
4. *Caenorhabditis elegans*, *Cele*.
5. *Culex quinquefasciatus*, *Cqui*.
6. *Daphnia pulex*, *Dpul*.
7. *Drosophila melanogaster*, *Dmel*.
8. *Homo sapiens*, *Hsap*.
9. *Ixodes scapularis*, *Isap*.
10. *Nematostella vectensis*, *Nvec*.
11. *Pediculus humanus*, *Phum*.
12. *Saccharomyces cerevisiae*, *Scer*.

As sequências proteicas de cada um dos organismos foram descarregadas da base de dados do NCBI (National Center for Biotechnology Information, [www.ncbi.nlm.nih.gov/](http://www.ncbi.nlm.nih.gov/)) no formato FASTA.

#### 3.2. Busca de proteínas com alta similaridade com a alfa-glucosidase Rpro|ACI96031 de *R. prolixus*

HMMER (Hidden Markov Models, <http://hmmer.janelia.org/>) é uma ferramenta de bioinformática elaborada para detectar domínios conservados em seqüências protéicas usando o método estatístico de Hidden Markov Models (HMM). Em geral, tais domínios são provenientes de bancos de dados, como por exemplo o PFAM (<http://pfam.sanger.ac.uk/>). Neste trabalho foi utilizada a seqüência parcial da alfa-glucosidase de *R. prolixus* Rpro|ACI96031, para confirmar a família do PFAM que compreende as alfa-glucosidases (**PF00128**).

O domínio PFAM representante da família das alfa-glucosidases foi então utilizado para detectar enzimas desta família nos genomas selecionados (e-value < 0,01). Visando minimizar o risco de incluir isoformas do mesmo gene nas análises, foi utilizado o programa BLASTCLUST (identidade mínima de 95% e 80% de cobertura das seqüências agrupadas (<http://toolkit.tuebingen.mpg.de/blastclust>) e apenas uma proteína retida por cluster.

### 3.3. Predição de regiões transmembranares

Dois programas foram usados para realizar a predição de regiões transmembranares nas proteínas estudadas: TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>) e Phobius (<http://phobius.sbc.su.se/>). Estes programas estimam a probabilidade de ocorrência de regiões transmembranares na proteína a partir de cálculos de hidrofobicidade dos resíduos de acordo com a posição da seqüência na proteína.

### 3.4. Alinhamento múltiplo de seqüências

As proteínas selecionadas foram alinhadas utilizando o programa MUSCLE, (<http://www.ebi.ac.uk/Tools/msa/muscle/>). O alinhamento múltiplo resultante foi utilizado para a identificação de domínios conservados e reconstrução filogenética.

### 3.5. Construção da árvore filogenética

Foi usado o programa PHYML, “**PHY**logenetic estimation using **M**aximum Likelihood” (<http://www.phylogeny.fr/version2/>), executado com 100 Bootstraps e Likelihood-Ratio Test, modo SH-like.

### 3.6. Alinhamento múltiplo de estruturas

Foram localizados os melhores modelos estruturais para as 6 diferentes alfa-glucosidases de *R. prolixus* utilizando o site BLASTp do NCBI. So 4 alfa-glucosidases tiveram estruturas modelo, com a identidade suficiente (**Tabela 5**) para ser utilizados na modelagem comparativa, feita com o servidor Swiss model. As 4 alfa-glucosidases de *R. prolixus* simuladas foram alinhadas estruturalmente utilizando o servidor MISTRAL (**Figura 19**).

## 4. RESULTADOS

### 4.1. Resíduos marcadores das enzimas da família alfa-amilase

A família das alfa-amilases contém diversas glucosil-hidrolases, entre elas a subfamília das alfa-glucosidases. Assim a primeira questão que surgiu sobre a seqüência da alfa-glucosidase de *R. prolixus* Rpro|ACI96031.1 (Mury et al., 2009) que foi utilizada como padrão de seleção das alfa-glucosidases de outros organismos, foi se esta seqüência terá algum padrão característico de resíduos identificadores da subfamília das alfa-glucosidases ou algum padrão relacionado a outras subfamílias como as glucotransferases, ou isoamilases por exemplo, muito próximas evolutivamente das alfa-glucosidases.

Com o alinhamento entre as seqüências da alfa-glucosidase referência de *R. prolixus* (**Rpro|ACI96031.1**) com a proteína Taka-amilase A (XP\_001821436.1), obtivemos a localização dos resíduos hidrofóbicos **X** e hidrofílicos **B**, mostrada na **Tabela 1b**. O resultado indica que a proteína Rpro|ACI96031.1 não perdeu os resíduos hidrofóbicos e hidrofílicos característicos das proteínas da família alfa-amilase.

A **Tabela 2** mostra os resíduos caraterísticos do domínio A de alfa – amilases com diferentes especificidades. Cinco das seis posições características das alfa-glucosidases (posições encerradas em caixas) aparecem na enzima referência de *R. prolixus*, enquanto que somente duas das seis posições, coincidiram com a seqüência da Ciclodextrina Glucanotransferases, confirmando que esta proteína pertence a subfamília das alfa-glucosidases.

### 4.2. Árvore filogenética das alfa-glucosidases

A seqüência da alfa-glucosidase de *R. prolixus* **Rpro|ACI96031.1** foi utilizada para encontrar o PFAM **PF00128**, que foi usado para localizar as alfa-glucosidases dos organismos selecionados utilizando o programa **HMMER**. Possíveis isoformas do mesmo gene foram eliminadas utilizando o programa **BLASTClust**. As alfa-glucosidases foram alinhadas usando o programa **MUSCLE** e a construção filogenética

foi feita com o software **PHYML**. A representação gráfica da árvore filogenética obtida foi feita usando o programa **FIGTREE (figura 6)**.

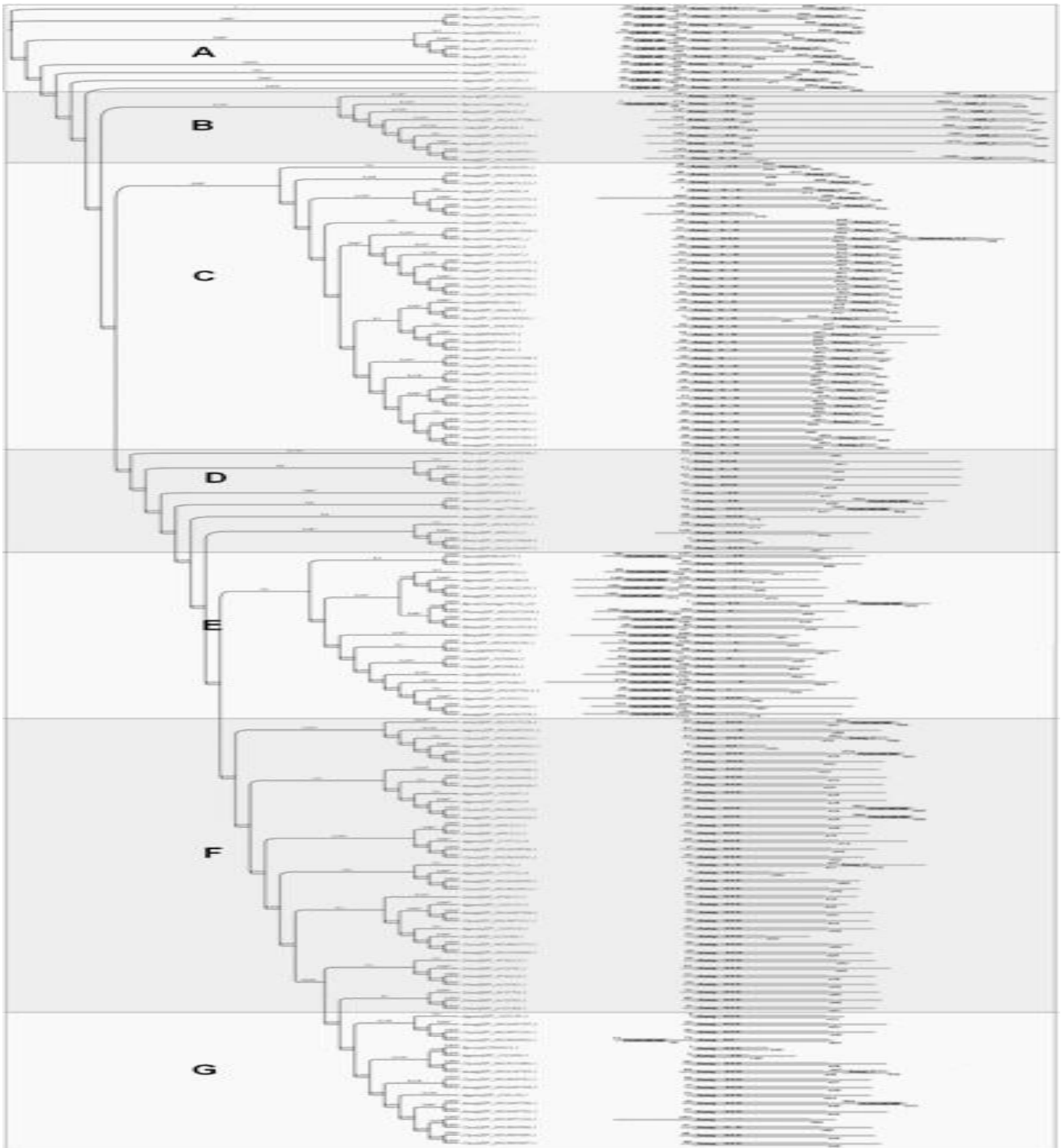
A árvore foi organizada com as ramificações A, B, C, D, E, F e G; cada ramificação foi classificada de acordo as características predominantes como tipo de domínio, solubilidade da proteína, presença ou ausência de resíduos específicos e outros.

A **figura 7** mostra apenas a região A da árvore filogenética. Esta região se caracteriza pela presença dos domínios ligantes de carboidratos CBM<sub>48</sub> (PF02922), alfa-amilase (PF00128) e alfa-amilase C (PF02806) em todas as proteínas. Esta região apresenta proteínas solúveis com domínios alfa amilase (PF00128) de tamanhos variados (93 – 290 aminoácidos) e só duas proteínas têm atividade catalítica de acordo a presença da tríade catalítica característica da família alfa-amilase.

A segunda ramificação da árvore filogenética, **região B**, mostrada na **figura 8**, se caracteriza pela presença do domínio alfa-amilo-1,6-glucosidase (PF06202). Inclui proteínas solúveis, com exceção da proteína de *R. prolixus* Contig15916\_1, com uma região transmembranar na porção N terminal. Os domínios alfa amilase (PF00128) têm tamanhos mais uniformes (100 – 138 aminoácidos), mas nenhuma tem atividade alfa-glucosidase.

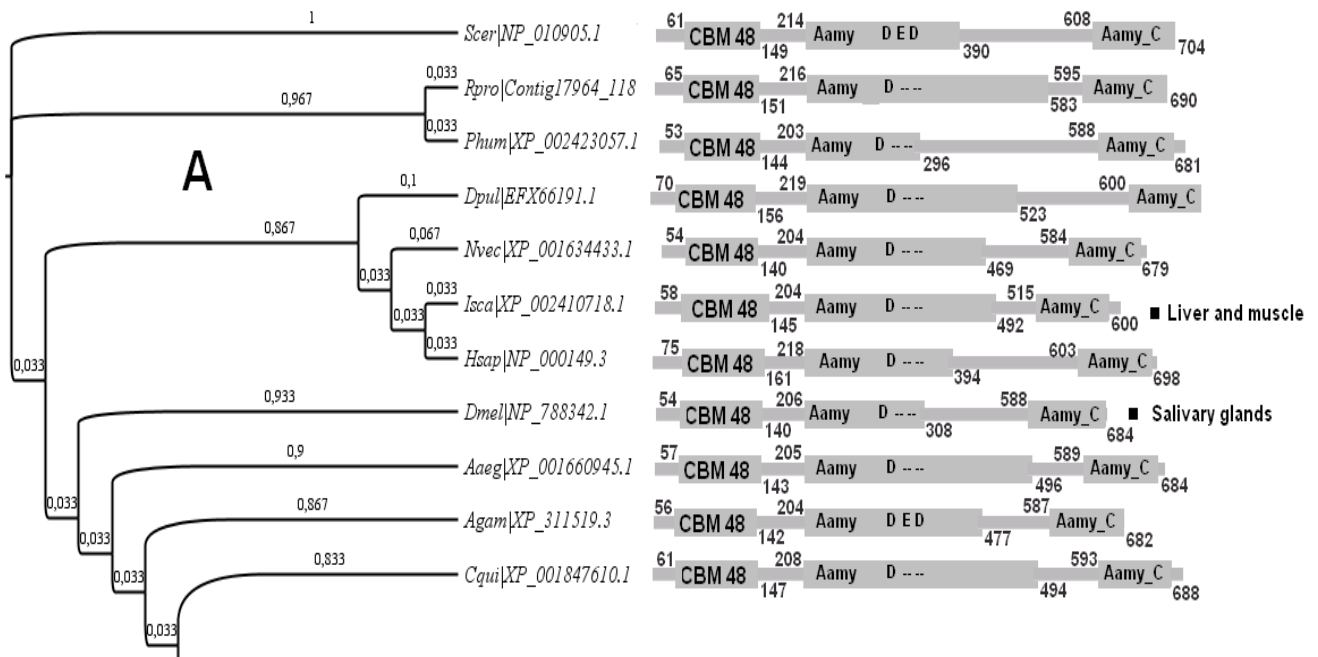
A terceira região da árvore filogenética, **região C**, é mostrada na **figura 9**. Esta se caracteriza pela presença do domínio alfa amilase (PF00128) e alfa amilase C (PF02806) muito próximos. Esta região só contém proteínas solúveis, com exceção da proteína de *R. prolixus* Contig15916\_1, que tem a região transmembranar na porção N terminal. Os domínios alfa amilase (PF00128) das proteínas nesta região têm tamanhos mais uniformes (100 – 138 aminoácidos) e os resíduos catalíticos se caracterizam pela ausência do ácido glutâmico **E**, só a proteína de *R. prolixus* Contig15916\_1, tem a tríade catalítica.

A quarta região da árvore filogenética, região **D**, é mostrada na **figura 10**. Esta se caracteriza pela ausência do domínio alfa amilase C (PF02806). Duas proteínas têm região transmembrana na porção C terminal, Amel|XP\_624736.2 e Rpro|Contig17590\_29. Nesta região, sete das treze proteínas têm a tríade catalítica das alfas glucosidases completa.

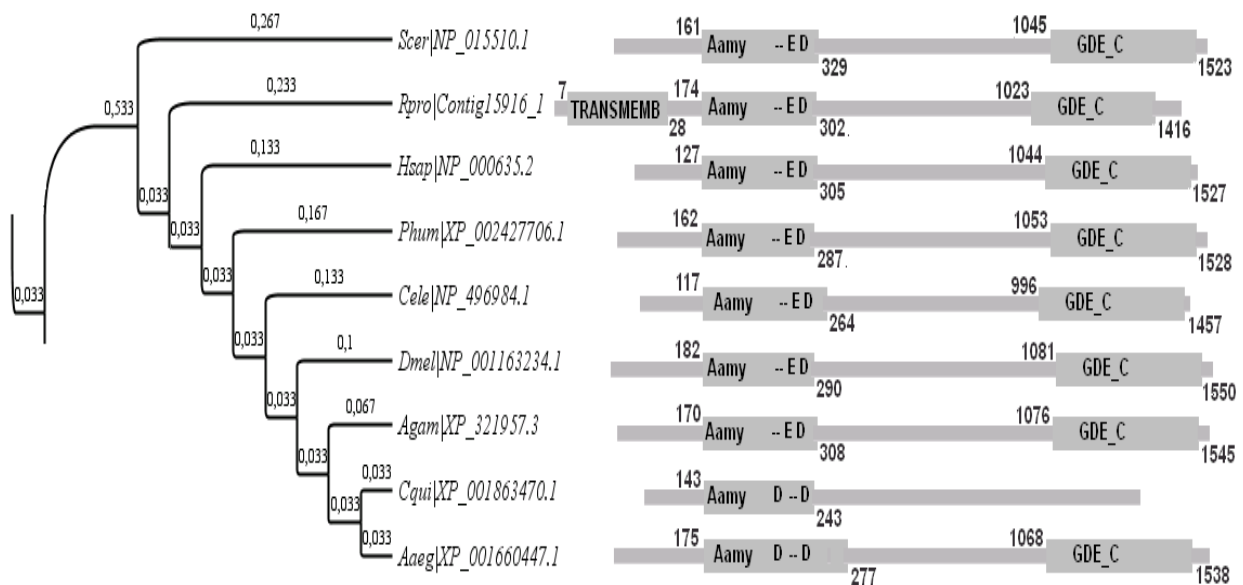


**Figura 6. Organização dos domínios na árvore filogenética das alfa-glucosidases.**

Cada região indicada pelas letras do alfabeto corresponde a ramificações com características comuns como resíduos catalíticos, domínios, solubilidade e outros. Acrônimo dos organismos: *Aedes aegypti*: **Aaeg**, *Anopheles gambiae*: **Agam**, *Apis Mellifer*: **Amel**, *Caenorhabditis elegans*: **Cele**, *Culex quinquefasciatus*: **Cqui**, *Daphnia pulex*: **Dpul**, *Drosophila melanogaster*: **Dmel**, *Homo sapiens*: **Hsap**, *Ixodes scapularis*: **Isca**, *Nematostella vectensis*: **Nvec**, *Pediculus humanus*: **Phum**, *Saccharomyces cerevisiae*: **Scer**. Os códigos em frente a cada organismo correspondem a identificadores obtidos da base de dados de NCBI, com exceção das proteínas de *R. prolixus*, cujos identificadores são provenientes do projeto genoma (em andamento). Domínios preditos usando SMART (<http://smart.embl-heidelberg.de/>).



**Figura 7 . Região A da árvore filogenética das alfa-glucosidases.** Na direita da ramificação aparece a organização dos domínios, indicando os resíduos onde iniciam e terminam. O domínio ligante de carboidratos (Carbohydrate-binding module family 48, pfam **PF02922**) é marcado como **CBM\_48**. O domínio alfa-amilase (Alpha-amylase family 13, pfam **PF00128**) é marcado como **Amy**. O domínio alfa-amilase C terminal (alpha-amylase, C terminal all-beta domain, pfam **PF02806**) é marcado como **Amy\_C**. A presença da tríade catalítica **D, E e D** (**D**: ácido aspártico, **E** ácido glutâmico) do domínio alfa-amilase **Amy** (pfam **PF00128**) indica que domínios são ativos ou inativos, a marcação -- indica ausência do resíduo catalítico por deleção evolutiva, a marcação \*\* indica ausência do resíduo catalítico por sequência incompleta na proteína.



**Figura 8. Região B da árvore filogenética das alfa-glucosidases.** O domínio alfa amilo-1,6-glucosidase (Glycogen debranching enzyme, pfam **PF06202**) é marcado como **GDE\_C**. As regiões hidrofóbicas da proteína (Transmembrana) são marcadas como **TRANSMEMB**.



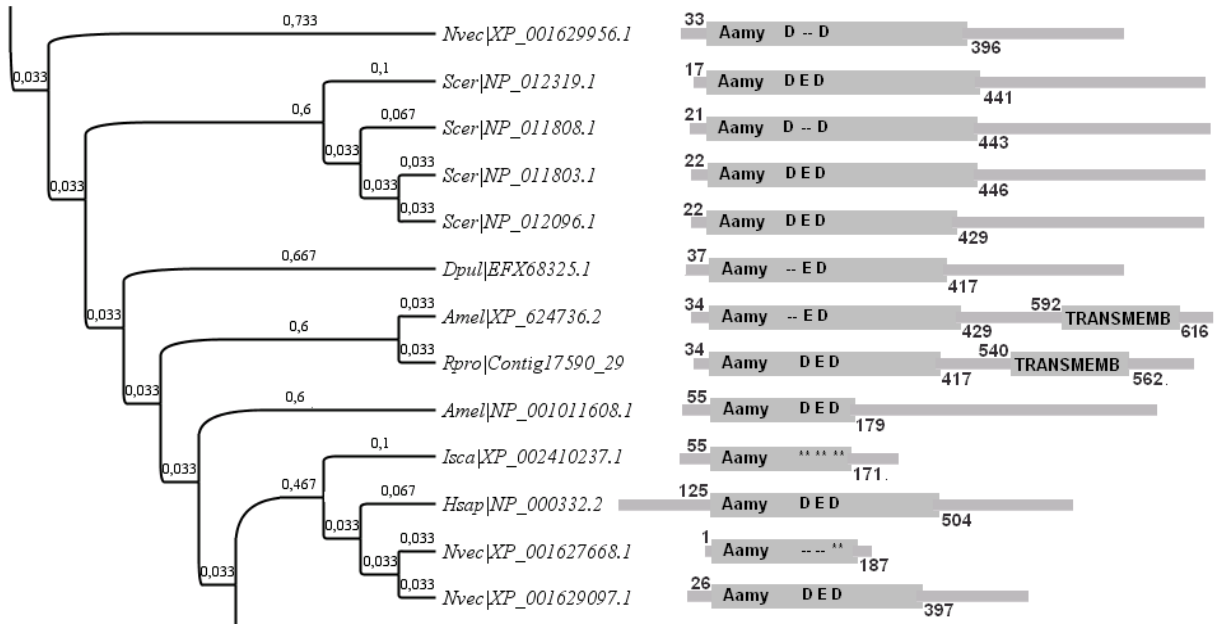


Figura 10. Região D da árvore filogenética das alfa-glucosidases. Ver legenda figuras 7, 8 e 9.

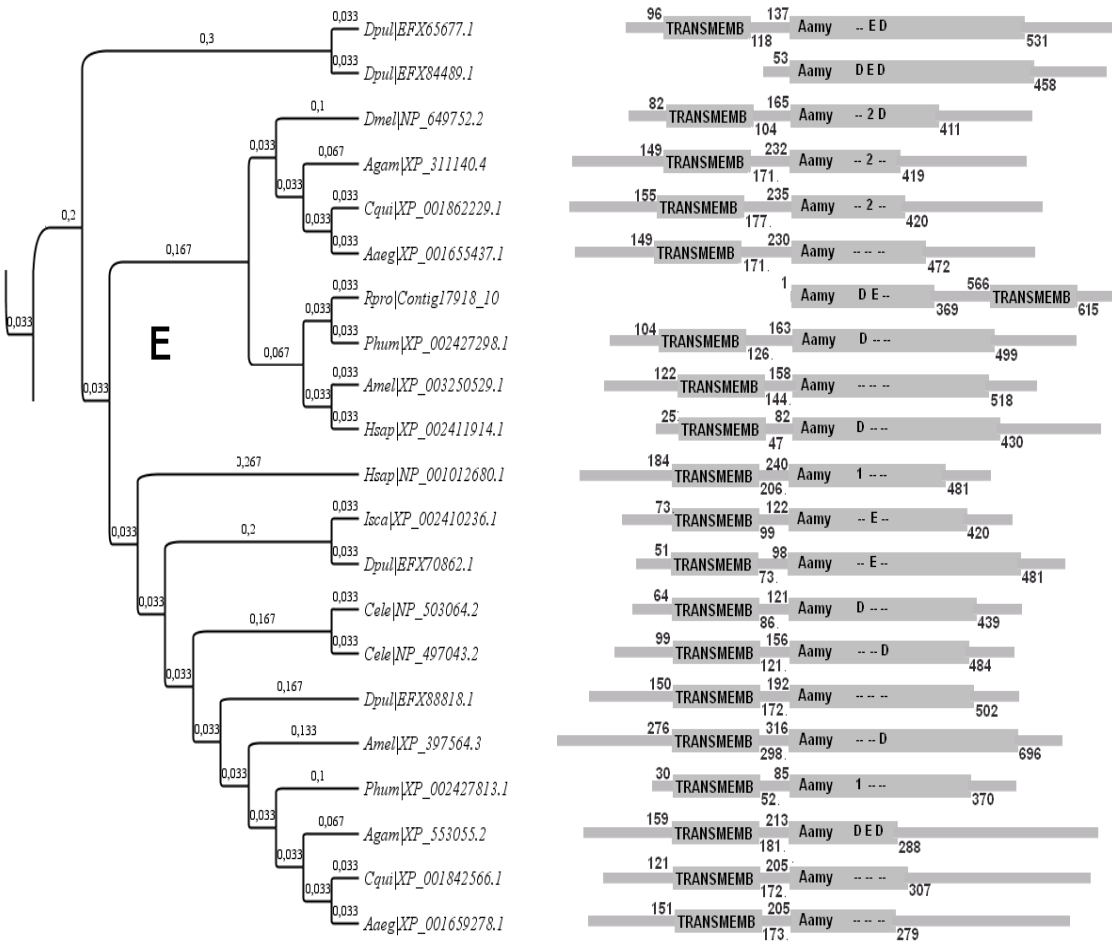
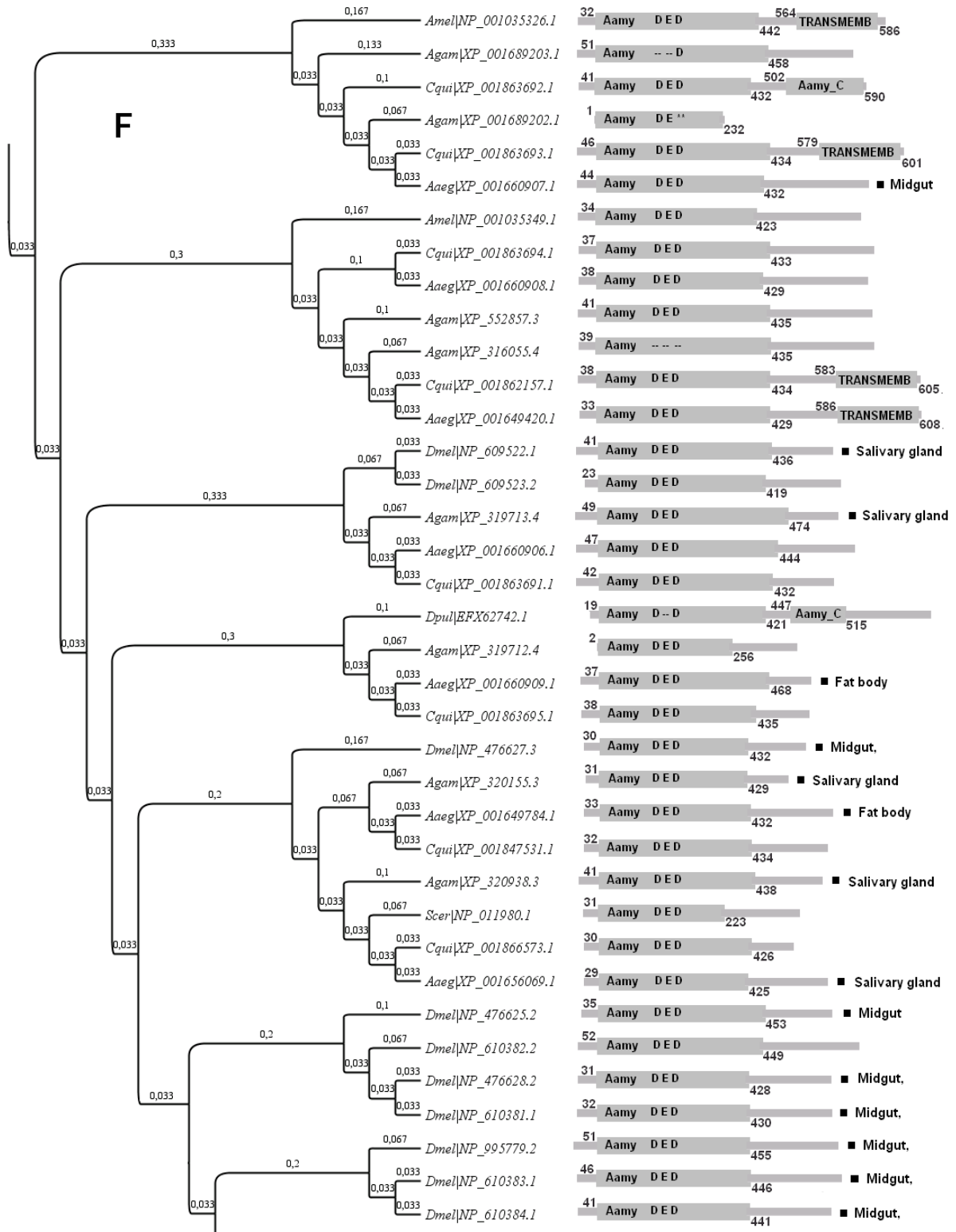


Figura 11. Região E da árvore filogenética das alfa-glucosidases. Ver legenda figuras 7, 8 e 9.



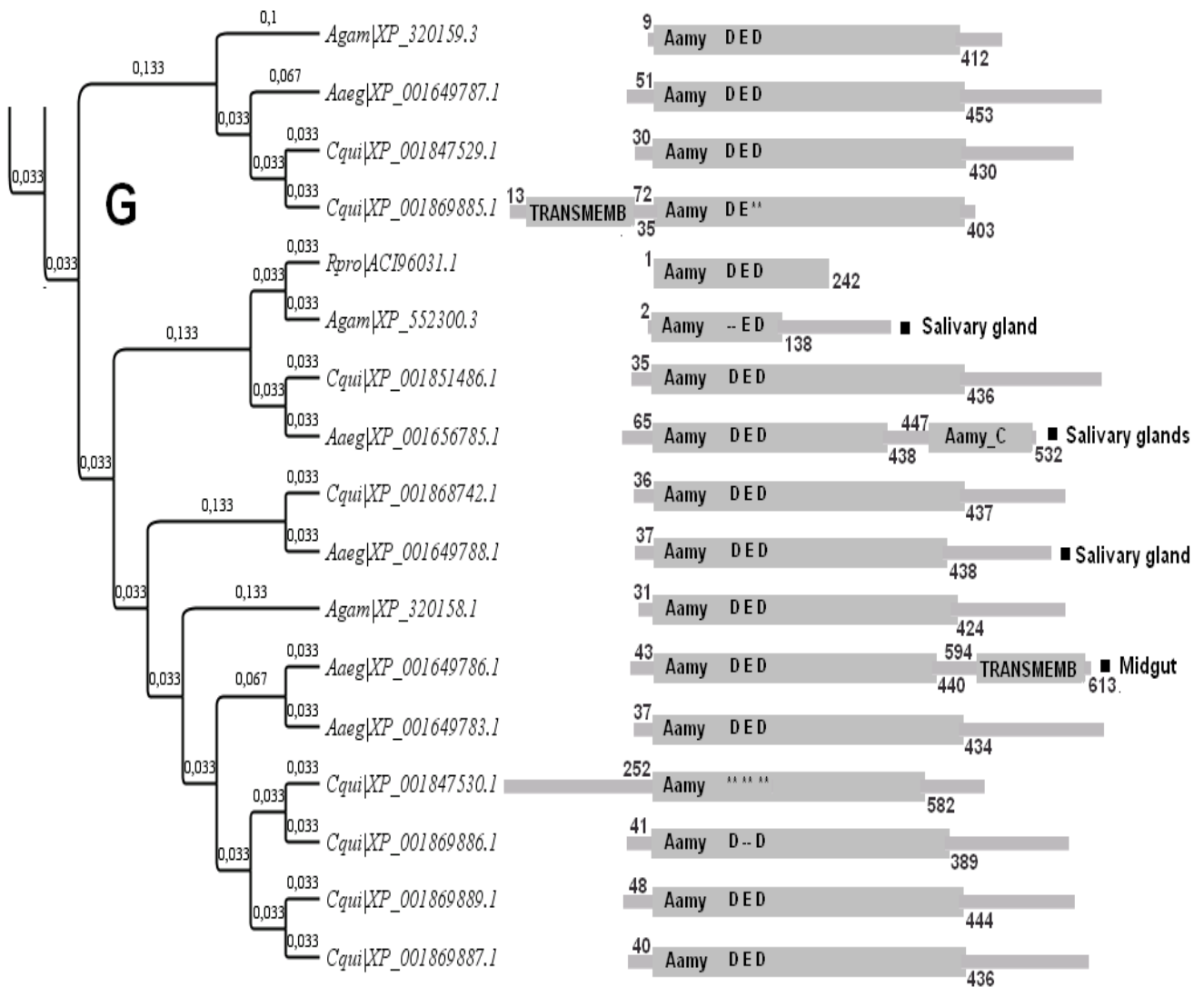
A **região E** da árvore filogenética é mostrada na **figura 11**. Esta se caracteriza pela região transmembrana na extremidade N terminal, próxima ao domínio alfa-glucosidase (PF00128). Só duas proteínas têm a tríade catalítica completa.-



**Figura 12.** Região F da árvore filogenética das alfa-glucosidases. Ver legenda figuras 7, 8 e 9.

A região **F** da árvore filogenética é mostrada na **figura 12**. Nesta figura pode se observar que somente duas proteínas têm o domínio alfa-glucosidase C terminal e quatro proteínas têm região transmembrana na extremidade C terminal.

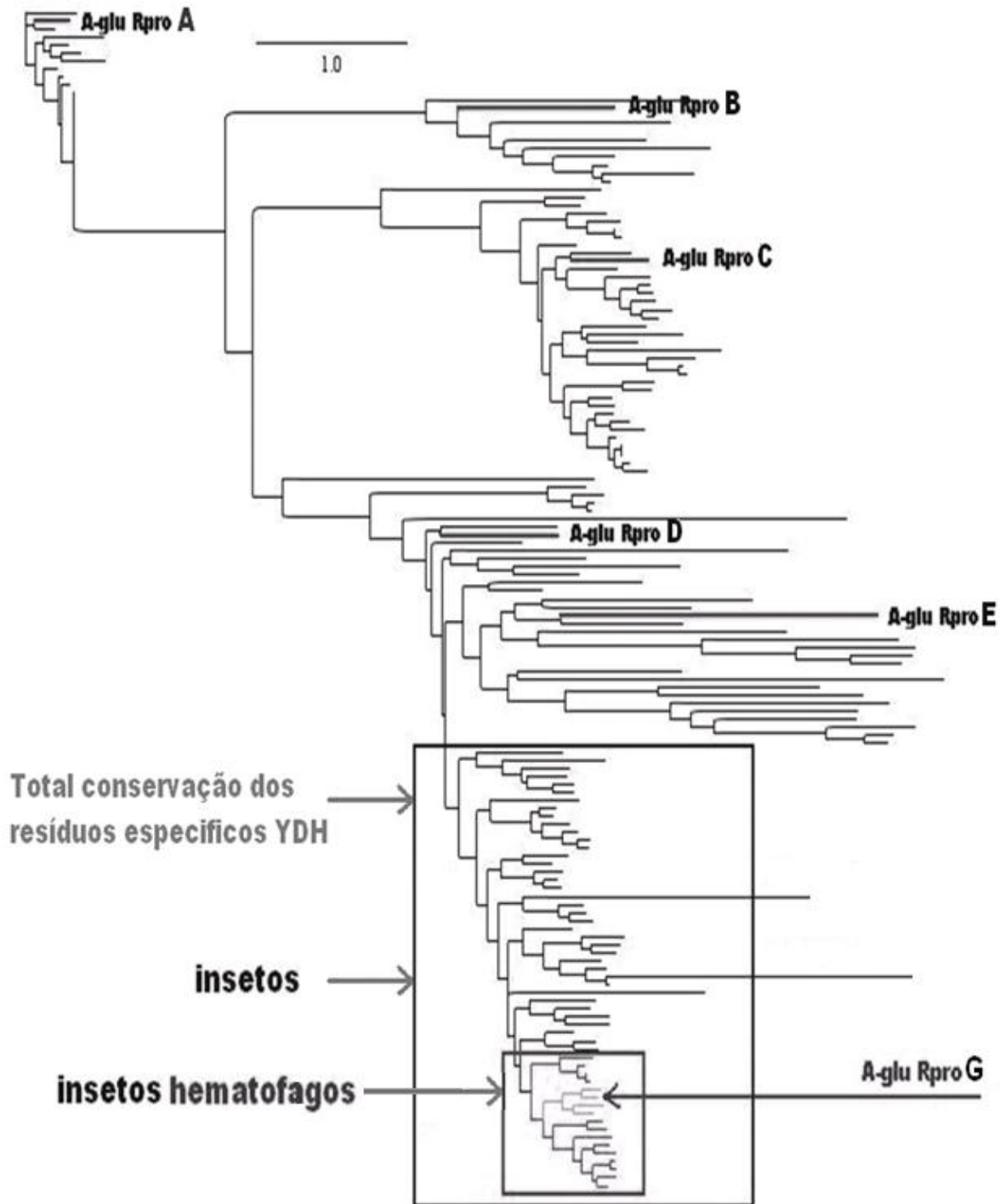
Na **região G** da árvore filogenética, (**figura 15**) somente constam proteínas de insetos hematófagos e somente duas delas não têm a tríade catalítica completa.



**Figura 13. Região G da árvore filogenética das alfa-glucosidases.** Ver legenda figuras 7, 8 e 9.

A árvore filogenética das alfa-glucosidases selecionadas de cada organismo resultou em uma distribuição de alfa-glucosidases de insetos hematófagos concentradas apenas no final da árvore, região marcado na caixa pequena (**Figura 14**). Na mesma ramificação está a alfa-glucosidase referência de *R. prolixus*, identificada como A-glu Rpro A.

Em uma maior ramificação, estão às alfa-glucosidases de insetos, região marcada em caixa maior na **Figura 14**, exceto para a alfa-glucosidase de *Daphnia Pulex* e *Saccharomyces cerevisiae*.



**Figura 14. Conservação dos resíduos Y, D, H das alfas glucosidases na árvore filogenética.** As alfa-glucosidases de *R. prolixus* estão marcadas como A-glu Rpro. As letras A, B, C, D, E, G correspondem às regiões da árvore filogenética onde elas se encontram. A caixa maior destaca alfa-glucosidases de insetos e a caixa menor destaca um clado contendo apenas alfa-glucosidases de insetos hematófagos.

Esta ramificação que tem insetos hematófagos e não hematófagos é caracterizada pela conservação total de resíduos de ácido aspártico, tirosina e histidina (Kumar, 2010). Estes resíduos são característicos das alfa-glucosidases do clã GH-H das glucosilhidrolases (<http://www.cazy.org/Glycoside-Hydrolases.html>). A conservação destes resíduos desaparece fora desta região, indicando, que estes resíduos altamente conservados nestas alfa-glucosidases desempenham uma função importante e específica neste clã.

A partir da sequência da alfa-glucosidase referência de *R. prolixus*, **AgluRproG** (Rpro|ACI96031.1) (**Figura 14**), foram procuradas outras alfa - glucosidases no banco de dados de proteínas de *R. prolixus*, usando a ferramenta de busca HMMER, resultando em proteínas marcadas por **Aglu Rpro.A, B, C, D, E e F**. Estas proteínas estão distribuídas ao longo de vários clados da árvore filogenética. Vale destacar que a proteína **A-gluRProG** aparentemente tem maior afinidade filogenética por alfa-glucosidases contidas no clado contendo apenas proteínas de insetos hematófagos.

Finalmente, da árvore filogenética é possível fazer uma tabela das características mais gerais de cada região da árvore (**tabela 3**).

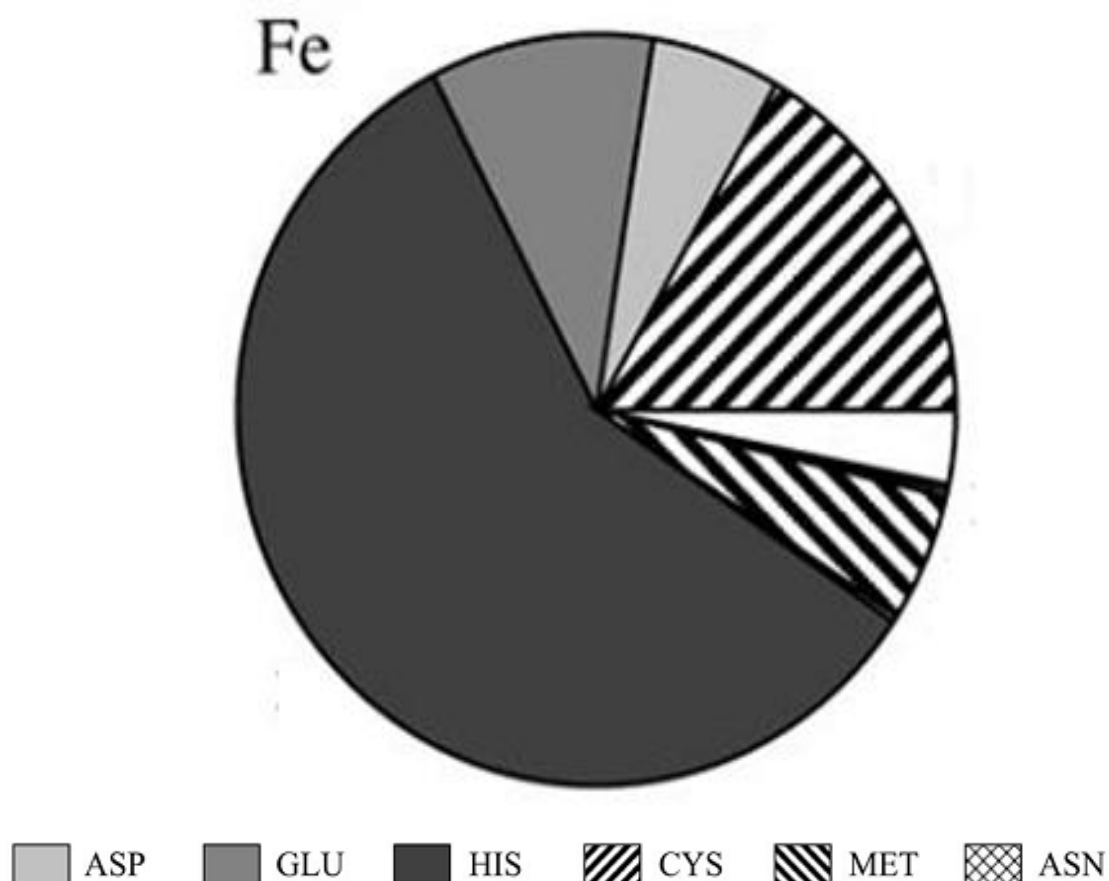
**Tabela 3.** Características gerais das regiões da árvore filogenética das alfa-glucosidases.

Região	Quantidade de proteínas	alfa-glucosidases ativas	Presença de domínio C terminal
A	11	2	11
B	9	0	0
C	36	1	34
D	13	7	0
E	21	2	0
F	37	34	1
G	17	14	1

### 4.3. Alinhamento múltiplo das sequências das alfa-glucosidases

A ligação de metal em proteínas normalmente exige um arranjo geométrico específico de relativamente poucos resíduos, cuja identidade depende fortemente do tipo de ligação metálica (Brylinski & Skolnick, 2011). Alguns resíduos de aminoácidos

interagem com maior intensidade com íons metálicos. Isto pode ser visto na **Figura 15**, onde os resíduos de cisteína e histidina interagem mais fortemente com o átomo de ferro, enquanto que os resíduos de ácido aspártico e glutâmico o fazem com menor intensidade.

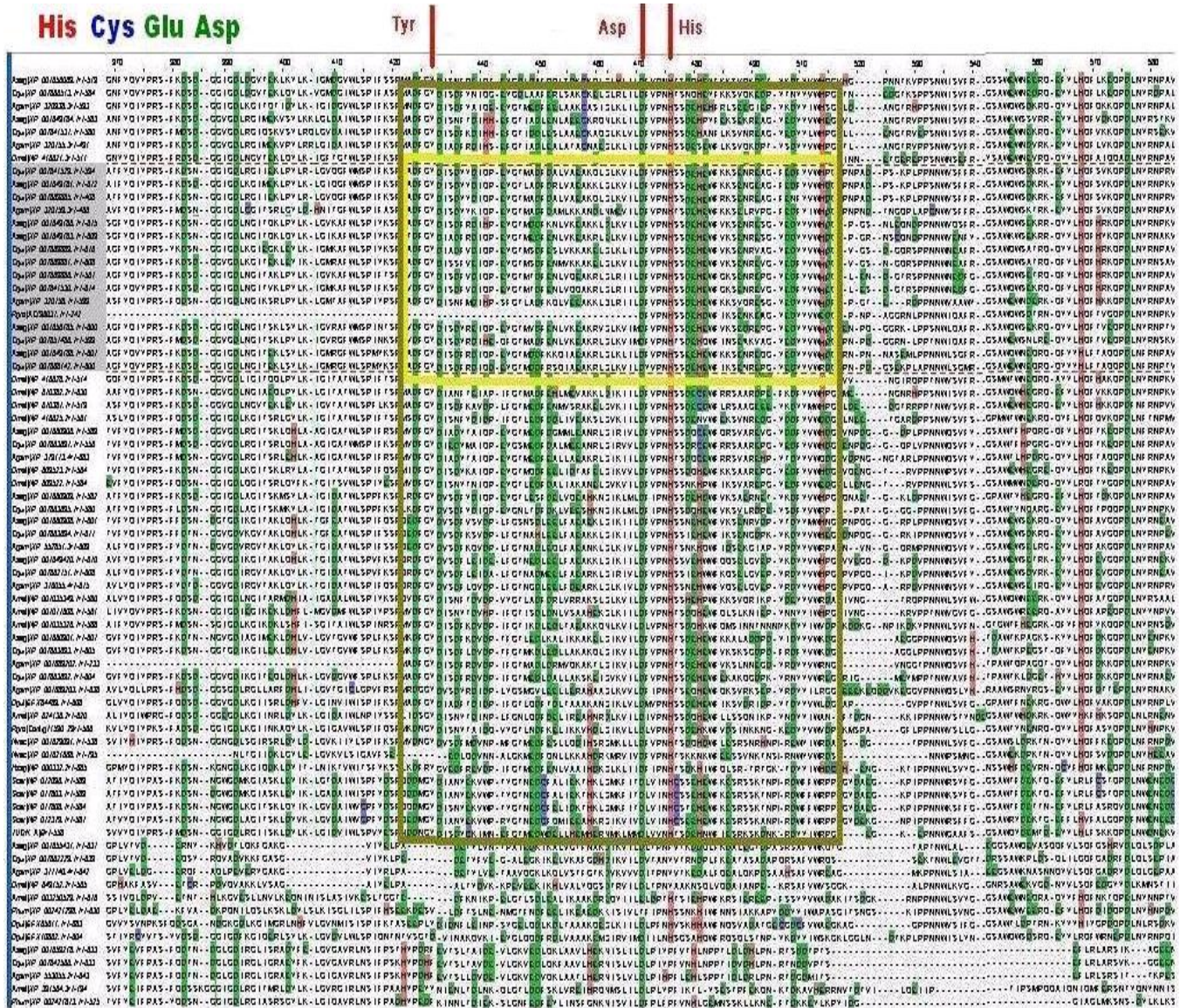


**Figura 15.** Preferências de aminoácido em relação a ligação de íons ferro. (Brylinski & Skolnick, 2011).

A partir desta informação, introduzimos diferentes cores para reconhecer os quatro resíduos que melhor se adequam à ligação com metais no alinhamento múltiplo e podemos observar uma pequena região altamente conservada (**Figura 16**). Esta região contém os três resíduos Y, D, H (ácido aspártico, tirosina e histidina) característicos das alfa-glucosidases do clã GH-H acima mencionados na árvore filogenética.



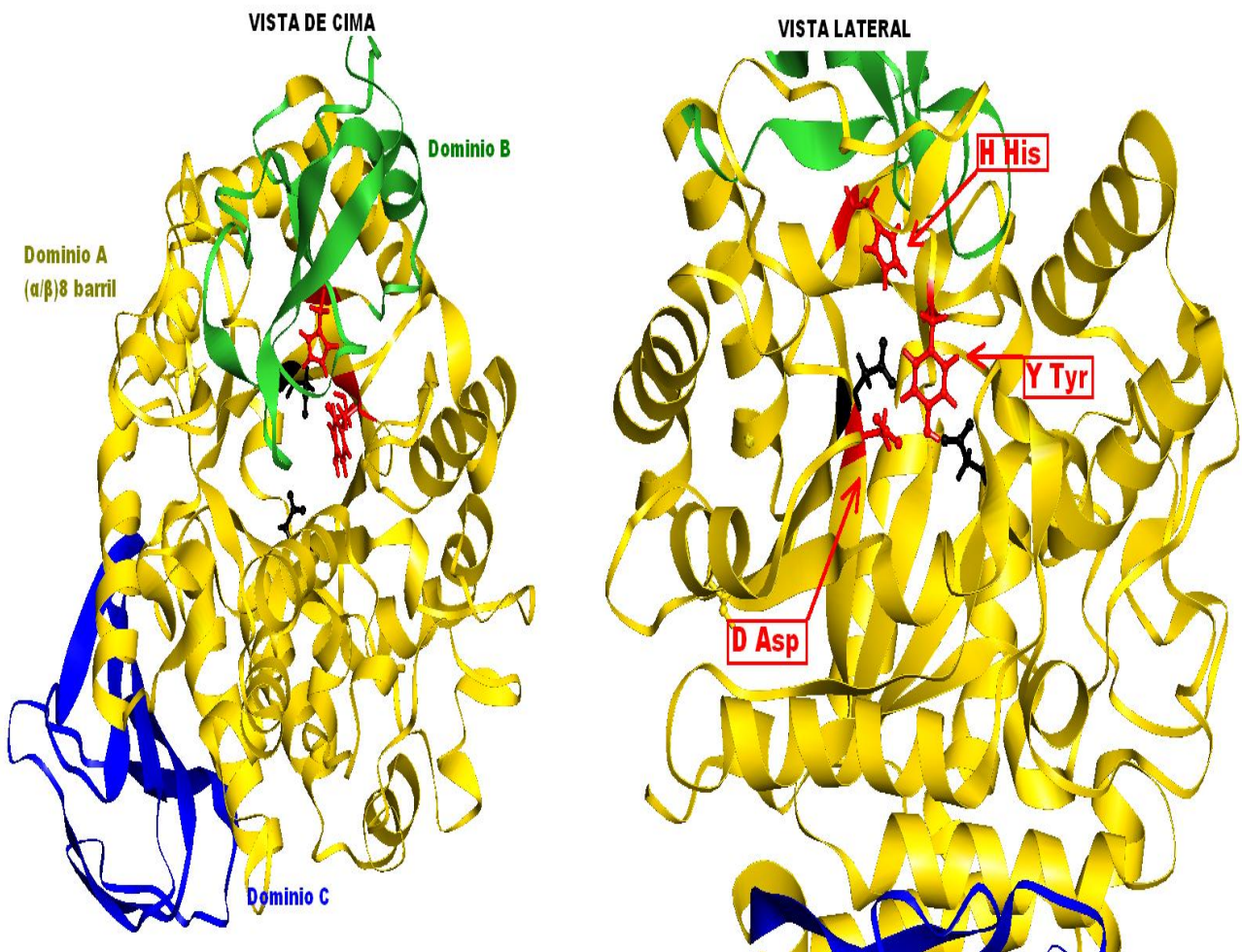
começam a desaparecer quando observamos as seqüências de proteínas de organismos que não pertencem à classe Insecta, como *Daphnia pulex* (Arthropoda: Crustacea) e *Saccharomyces cerevisiae* (Fungi, Saccharomycetes)



**Figura 17.** Distribuição dos resíduos H (vermelho), C (azul), D e E (verde) no alinhamento múltiplo das alfa-glucosidases, correspondente a região 370 – 570 do alinhamento. As proteínas dentro da caixa de marcação amarela mais clara correspondem a proteínas de insetos hematófagos e dentro da caixa de marcação amarela mais escura correspondem a proteínas de insetos não hematófagos. Resíduos (427Y, 471D, 476H) característicos do clã GH-H.

O padrão característico mostrado na **Figura 17** pode indicar que resíduos de ácido aspártico e ácido glutâmico são candidatos a interagir com o átomo de ferro,

sugerindo que possam estar envolvidos no passo de formação do dímero de heme no processo de formação de H<sub>z</sub>, já que os resíduos conservados Y, D, e H estão em uma posição tridimensionalmente privilegiada, uma vez que os três resíduos, estão localizados na parte superior interna na estrutura barril (alfa / beta) 8 mostrada na **Figura 18**. Isto pode ter sido utilizado pelos hemípteros ao longo da evolução para, por exemplo, ligar heme.



**Figura 20.** Estrutura da alfa-glucosidase de *R. prolixus* Rpro|ACI90631.1. Estrutura feita utilizando modelagem comparativa (Swiss protein Server) com a estrutura cristalina PDB 1UOK utilizando o programa para visualização molecular NOC (<http://noch.sourceforge.net/#Snapshots>). Os domínios A, B e C da alfa-glucosidase são mostrados em amarelo, verde e azul, respectivamente. Os resíduos catalíticos (D, E, D) são mostradas em preto e resíduos característicos de clã GH-H (Y, D, H) em vermelho.



#### 4.4. Alinhamento múltiplo das estruturas das alfa-glucosidases de *R. prolixus*

As alfa-glucosidases de *R. prolixus* não mostraram correlação entre a atividade alfa-glucosidase e a localização na árvore filogenética, conforme mostram os dados na **tabela 4**.

**Tabela 4.** Predição de propriedades bioquímicas das alfa-glucosidases de *R. prolixus*.

Proteína	Região na árvore	Resíduos catalíticos	Solubilidade
Rpro 17964_118	A	D -- --	Solúvel
Rpro 15916_1	B	-- E D	Membrana
Rpro 16092_1	C	D E D	Solúvel
Rpro 17590_29	D	D E D	Membrana
Rpro 17918_10	E	D E --	Membrana
Rpro ACI96031.1	G	D E D	Solúvel

A alfa-glucosidase Rpro|ACI96031.1 é considerada solúvel, já que é conhecida 50% de sua sequência (**Tabela 4**).

No sentido de encontrar correlação entre as estruturas das alfa-glucosidases de *R. prolixus* foi feita a modelagem comparativa destas utilizando o melhor modelo encontrado a partir do BLASTp do site NCBI, com a base de dados PDB (**tabela 5**).

**Tabela 5.** Dados da procura do modelo para geração das estruturas das alfa-glucosidases de *R. prolixus*

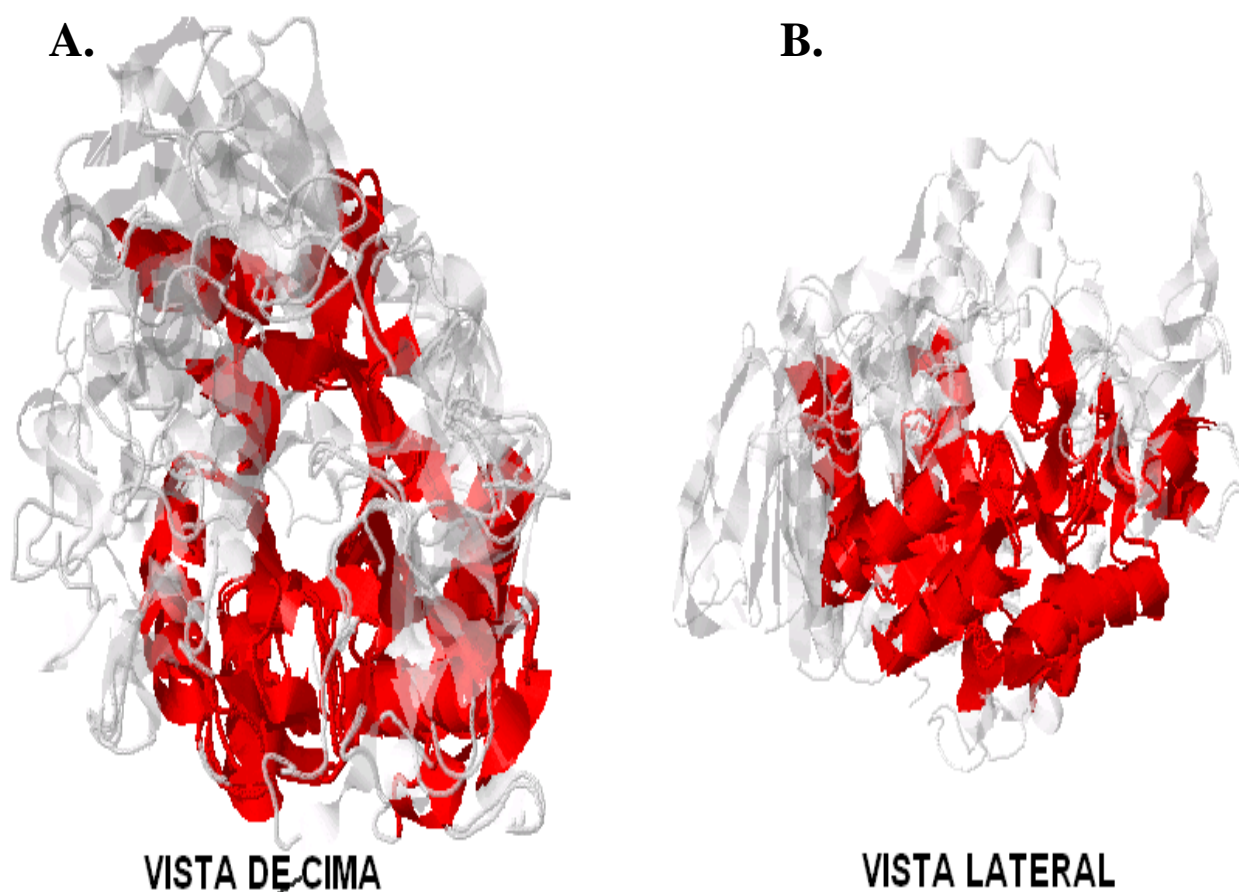
Proteína	Resíduos simulados	Identidade	PDB modelo
Rpro 17964_118	98%	52%	3AMK_A
Rpro 15916_1	6%	21%	+++
Rpro 16092_1	54%	57%	1VIW_A
Rpro 17590_29	83%	35%	4AIE_A
Rpro 17918_10	59%	23%	+++
Rpro ACI96031.1	83%	34%	3A47_A

+++ Indica que não foi simulado por falta de estruturas PDB que correspondam com a identidade mínima para servir como modelo.

Posteriormente a estrutura foi feita utilizando-se o servidor Swiss model (<http://swissmodel.expasy.org/>) e um alinhamento múltiplo das estruturas das alfa-glucosidases geradas pelo servidor MISTRAL (multiple protein structural alignment, <http://eole2.lscce.ipsl.fr/ipht/mistral/protein.php>) (**figura 19**).

No alinhamento múltiplo estrutural das alfa-glucosidases de *R. prolixus* (**figura 19**), se aprecia parte da estrutura em barril (alfa/beta)<sub>8</sub>. Em cor vermelha estão representados os resíduos que coincidem estruturalmente entre as proteínas.

Por ultimo, foram alinhadas as sequências das alfa-glucosidases de *R. prolixus* (**figura 20**) e este alinhamento mostra baixa similaridade entre as proteínas.



**Figura 19.** Alinhamento múltiplo das estruturas das alfa-glucosidases de *R. prolixus*, obtidas por modelagem comparativa dos modelos da **tabela 5**, com uma tolerância para alinhamento de 3 angstroms de distância.

Nas duas figuras de cima (**Figuras 19A e 19B**), em vermelho estão os resíduos das alfa-glucosidases que coincidem no alinhamento estrutural, em cor

cinza os resíduos que não coincidem. Nas duas figuras de baixo (**Figuras 19C e 19D**), as cores preta, amarela, azul e rosa correspondem as alfa-glucosidases Rpro|17964, Rpro|16092, Rpro|17590 e Rpro|ACI96031, respectivamente.

```

Rpro|ACI96031.1          -----D-----FVPNHS-----SNLHD 12
Rpro|Contig17590_29    FKELIKKVHDLGKIKVLLD-----LVPNHT-----SDEHK 130
Rpro|Contig17964_118  YVKHDFGIWKLHIPPNDGTAQINHLSEIKIAVRTTNGVLDRI-----SPWAT 152
Rpro|Contig16092_1    QEVIDYGGEAIAKATEYTD FGHVTEFRYGIELSRAFRGKNLLKWLRSFGPDWNLMPKSKAL 300
                        *                               :   *

Rpro|ACI96031.1          WFIKSENKVAGYEDYYVWED-GKPNPAGGRNLPNN----WIAQFRGSAWQWSDTRNQY 66
Rpro|Contig17590_29    WfvdsInkknGkEdyyVwVdPdkMSPAG--FLPPNN----WVSLGGSAWTYNENRQOY 183
Rpro|Contig17964_118  YVVKPKPTEGLNPKHLFWNPSPQKYLTKERRPKKPDTLRIYECHIGIATSEHKVGTYSE 212
Rpro|Contig16092_1    VFIDNHDTRQSSGDFLTYKDSRKYKMAVGFMLSWPYG-TKRVMSSDFDSSRDQGPHPNPD 359
                        ..  ..  ..  :   :   :   :   :   :   :   :   :   :

Rpro|ACI96031.1          YLHQFTVGGQPDNLNRYNPKVQEMKDILT-----FULDKGVGDGLRIDAVP 110
Rpro|Contig17590_29    YLHQFASYQPDlnYhNPKVLEeIKNVmR-----YULDIGVDGFRVDAVP 227
Rpro|Contig17964_118  FAKEIVPRIVKLGYNAVQIMAMEHAYYGSFGYQ-----VTSFFAASSRYGTPEELKQ 265
Rpro|Contig16092_1    MTIKDVTINPDMTCNNGWICEHRWRQIYMMVGFSNIVKGTfVENfWDNGSNQIAYCRGNK 419
                        :   .   :   .   :   :   :   :   :   :   :

Rpro|ACI96031.1          TLYEDTQ-----LRDEPASGLTEDTETN---CLLHIYTQDLPETVE----- 149
Rpro|Contig17590_29    FLVEDKE-----LVSEDQIQGCTDWSYN---CFRHDKTKNLKETYM----- 266
Rpro|Contig17964_118  LIDChKngLYVLLDlVhShASKNVLdGLNEFDGTDACyFhSGSRGThPLWdSRLFDYSQ 325
Rpro|Contig16092_1    GFIAINGNNYDLNQTlMTCLPAGLYCDIISGLKREHKCTGKSVEVDDNGKAQIHLSTHDE 479
                        :   :   :   :   :   :   :   :   :   :

Rpro|ACI96031.1          -----MVYEWRELVDAY---QREHGdTRVIMTEAYSSLEIIQTYYSASGR--LGSH 197
Rpro|Contig17590_29    -----ILRELGTVFEEN---KYKD--QPRKLFLEAYITLNDTIKYYSKHA----- 306
Rpro|Contig17964_118  IEVQRfLLSNlCWYYEYGFdGFRFDGVTsILYhSRGMEsFGDYNDYfGLNVDNDGLTYL 385
Rpro|Contig16092_1    DGILAIHIQEFSSLRDLK-----EEEENGYLPLDPNFATRAIHVGQDPDKWSSGIV 530
                        :   :   :   :   :   :   :   :   :   :

Rpro|ACI96031.1          MPFNFRITTELNEQSKPSDYVK-----VVQDWSIIVPAGNVP 234
Rpro|Contig17590_29    TPFNFNLLLLTRDVN-ATEVEK-----LLKDWTDVLPKENRT 342
Rpro|Contig17964_118  MLANyVlHnLYPNaITIAEDVSGMPGLCRPTEEAGVGFYRLAMaIPDKWIEllKTKRDE 445
Rpro|Contig16092_1    IPPIHMGTTfKQDAACKHRGyEYGRSGNPSRAVLEECLASLDDAKYGLAFSSGLAAVTGS 590
                        .   :   .   .   :   :

Rpro|ACI96031.1          NWVLGNHD----- 242
Rpro|Contig17590_29    IWVTGNHDNPRIGTKMGEDMIDAVSMLSFMAPGNTVTTYGEEIGHVDTYMG---RNYKS 398
Rpro|Contig17964_118  DWNVGNIVHTLSNRRWLEKNIAyceSHDQALVGDkTIAfWlMDKEMyTGMS---LIYGS 501
Rpro|Contig16092_1    CALLESGDHIITGDDIYGgTNRyLRNVaKkMGVDtTFFDmTAPeTLNEVlQPnTRlVWVE 650
                        .

Rpro|ACI96031.1          ----- 441
Rpro|Contig17590_29    DPRNPE-----RTPMQWNSNTSAGFSNITKTWLPVNPNYWDINQEKEA 441
Rpro|Contig17964_118  NDIIERGIALHKIIRLITHALGGEAYLNFMGNEFGHPeWLDfPREGNSesyQYARRQWNL 561
Rpro|Contig16092_1    SPTNPMKVMdIPRISEIIKRHNQDIIFVVDNTfLTPYfQKPLALGADISMYSLTKYMMNG 710

```

**Figura 20.** Alinhamento múltiplo das sequências das alfa-glucosidases de *R. prolixus*. O alinhamento foi feito com o servidor clustalw2 (<http://www.ebi.ac.uk/Tools/msa/clustalw2>).

## 5. DISCUSSÃO

A análise computacional das alfa-glucosidases de insetos hematófagos realizada neste trabalho, tem como objetivo encontrar evidências de adaptações bioquímicas que permitiram o aparecimento da hematofagia entre estes insetos. Foi assim que, a partir de uma sequência de alfa-glucosidase de *R. prolixus* (Rpro|ACI96031.1) previamente conhecida (Mury, 2009), foram localizadas outras alfa-glucosidases nas famílias de proteínas de *R. prolixus* e outros organismos utilizando o programa HMMER

A partir das sequências encontradas, uma árvore filogenética sem raiz, com 144 sequências de proteínas foi construída (**figura 6**). Nesta árvore observaram-se diversos domínios (**CBM\_48**, domínio ligante de carboidratos; **Aamy**, domínio alfa amilase da família 13; **Aamy\_C**, domínio C terminal da alfa amilase; **GDE\_C**, domínio desligante de glicogênio; **Aminotran\_1\_2**, domínio aminotransferase classe I e II), distribuídos nas diferentes regiões da árvore. Esta diversidade de domínios é comum nas glucosilhidrolases da família 13, GH13 (Stam et al., 2006). A família GH13, também chamada de família alfa-amilase, é conhecida por ter pouca identidade nas sequências das proteínas homólogas mais próximas (Kumar, 2011), como mostra a **figura 20**.

As glucosilhidrolases são proteínas muito antigas e são encontradas em quase todos os organismos vivos (Janecek et al., 2003). Estas enzimas parecem ter sofrido modificações estruturais mais cedo que outras proteínas, como sugerem alguns estudos de mutação de resíduos de um motif nas posições equivalentes de duas alfas amilases, gerando mudanças na especificidade da enzima (Kumar, 2011). Isso pode explicar a baixa identidade das alfa-glucosidases da árvore filogenética e entre as alfa-glucosidases de *R. prolixus* (**figura 14**).

A tríade catalítica das alfa-glucosidases (Asp107, Glu175, Asp242 na proteína de *R. prolixus* Rpro|ACI96031.1) são resíduos invariáveis entre os membros do clã GH-H (famílias GH13, GH70 e GH 77) das glucosilhidrolases (Stam, 2006). A ausência de algum destes resíduos indicam, que a proteína pode ser considerada inativa. Assim, as alfa-glucosidases que apresentam a tríade catalítica completa, se encontram nas regiões F e G da árvore filogenética-(**tabela 1**). A região G tem alfa-

glucosidases de insetos hematófagos e a região F tem alfa-glucosidases de insetos tanto hematófagos como não hematófagos (**figura 14**).

A região conservada de resíduos de ácidos aspártico e ácido glutâmico no alinhamento múltiplo de sequências de insetos hematófagos (**figuras 18 e 19**) é um dado muito importante para se comparar com as regiões conservadas do alinhamento múltiplo das alfa-glucosidases de *R. prolixus*, combinado com a atividade formadora de hemozoina silenciada via RNAi nas alfa-glucosidases de *R. prolixus*. Com isto, pode ser feita uma predição estrutural de resíduos que têm preferência de ligação a íons ferro, com programas de predição que têm de 70% até 90% de precisão ou outros programas de predição estrutural para a interação de resíduos com o grupo propiônico do grupo heme (Ylinski & Skolnick, 2010).

Testes bioquímicos e de biologia molecular indicam a presença das alfa-glucosidases nas membranas perimicrovilares do intestino médio do *R. prolixus*. Com o interesse de confirmar o tipo de atividade da sequência da alfa-glucosidase de *R. prolixus* Rpro|ACI96031 encontrada nos mesmos testes (Mury et. al, 2009), foram comparados os resíduos do motif 4Beta a 5Alfa do domínio A característicos para as diferentes enzimas do clã GH-H (McGregor et. al, 2001), coincidindo com os resíduos característicos da família alfa amilase (**Tabela 1**) e da subfamília alfa-glucosidase (**Tabela 2**).

O domínio alfa-amilase PF00128 das alfa-glucosidases tem diferentes tamanhos na árvore filogenética devido ao fato de que o domínio B, (o domínio B é um loop localizado entre a terceira folha Beta e a terceira alfa hélice do barril  $(\alpha/\beta)_8$ , (**figura 4**) em animais e plantas, pode ter tamanhos de 40 até 250 resíduos; apesar do grande tamanho do domínio B, este não é considerado um domínio independente (Pujadas & Palau, 2002). É por este motivo que só aparecem nas bases de dados os domínios A e C.

Em algumas proteínas da família GH13, o domínio alfa amilase C terminal estabiliza o domínio catalítico, domínio A, criando um escudo de resíduos hidrofóbicos, além de contribuir na ligação de substrato (McGregor et. al, 2001). Mas, no caso das alfa-glucosidases homólogas à proteína Rpro|ACI96031.1, não existe relação direta com a atividade alfa-glucosidase. O domínio alfa amilase C terminal se encontra em maior proporção nas regiões A e C da árvore filogenética (**figuras 7 e 9**). Mas, nestas regiões apenas aparecem três alfa-glucosidases com a tríade catalítica completa (**Tabela 3**), indicando que não há correlação entre a

presença do domínio alfa amilase C com a possível atividade catalítica nas proteínas destas regiões.

Os resíduos Y54, D98 e H103, da estrutura PDB 1E3Z, são característicos da família GH13 (**figura 5**). Estes resíduos acomodam o substrato e estabilizam o estado de transição durante a atividade glucosilhidrolase (Kumar, 2010). Estes resíduos são totalmente conservados nas regiões F e G da árvore filogenética (**figura 14**). Estas duas regiões também apresentam a maior quantidade de proteínas com a tríade catalítica completa.

A pouca informação obtida dos tecidos nos quais foram expressas cada uma das alfa-glucosidases nos organismos da árvore filogenética, foi outro fator que impediu analisar a diversidade funcional das proteínas presentes na árvore filogenética.

A evolução das enzimas tem três caminhos diferentes, nas quais a especificidade do substrato, o mecanismo de reação ou a arquitetura do sitio ativo, é reutilizado na nova enzima (Wise & Rayment, 2004). No alinhamento múltiplo das estruturas das alfa-glucosidases de *R. prolixus* (**figura 19**), se pode observar a similaridade do domínio A marcado em vermelho. Este resultado coincide com a evolução divergente de enzimas devido à reutilização da arquitetura do sitio ativo, apesar que o alinhamento das seqüências das alfa-glucosidases de *R. prolixus* mostrar que estas compartilham baixa identidade (**figura 20**).

Pesquisas desenvolvidas em biologia estrutural respeito enzimas com estrutura de barril ( $\alpha/\beta$ )<sub>8</sub> divergentemente relacionadas revelaram que a conservação de estruturas muito estáveis são elementos indispensáveis na evolução divergente (Wierenga, 2001). Estes estudos mostraram que, ainda quando a evolução ocorreu em uma serie de mutações aleatórias, estruturas estáveis como a do barril ( $\alpha/\beta$ )<sub>8</sub> e as características estruturais do sitio ativo são frequentemente reutilizados ao longo da evolução e adaptação para novos propósitos catalíticos (Wierenga, 2001). Estas proteínas têm limitada identidade nas sequências, mecanismos de reação e substratos diferentes, mas adotam conformações estruturais de barril ( $\alpha/\beta$ )<sub>8</sub> com alta similaridade (Wise & Rayment, 2004).

Apesar das alfa-glucosidases de *R. prolixus* terem pouca similaridade nas suas sequências (**figura 20**), observou-se neste trabalho uma alta similaridade estrutural (**figura 19**). Isto sugere que a nova alfa-glucosidase de *R. prolixus* com atividade formadora de Hz evoluiu reutilizando as características estruturais da alfa-

glucosidase sem esta atividade e se adaptou para catalisar uma reação diferente da catalisada pela enzima progenitora, conservando a estrutura de barril ( $\alpha/\beta$ )<sub>8</sub> ou domínio catalítico alfa-glucosidase. Este domínio também executa a atividade formadora de hemozoina, já que inibidores específicos de alfa-glucosidase (Eritritol e castanoespermina) e um anticorpo policlonal obtido contra alfa-glucosidase, inibiu as atividades alfa-glucosidase e formação de hemozoina simultaneamente (Borgues et al., 2009).

A origem dos triatomíneos coincide com a formação do continente sulamericano, onde o ancestral (não hematófago) deste grupo de insetos evoluiu pouco depois da formação do continente (Patterson & Gaunt, 2010). Assim, mais mutações nas proteínas do organismo podem ter acontecido para executar atividades em diversas variedades de novos sistemas biológicos e ambientes (Janecek et al., 2003. Kumar, 2010). Este fato pode responder porque as alfa-glucosidases de *R. prolixus* estão dispersas nas diferentes regiões da árvore filogenética. Além disso, os insetos da ordem hemíptera são os insetos hematófagos mais antigos entre os insetos (**figura 1**). Isto pode refletir-se em proteínas com mais mudanças de resíduos e, portanto de funções.

Devido à baixa identidade das alfa-glucosidases, um teste de silenciamento da atividade formadora de Hz das cinco alfa-glucosidases de *R. prolixus* homólogas com a seqüência Rpro|ACI96031.1, será de grande utilidade para localizar a região conservada no alinhamento das seqüências das alfa-glucosidases que mostrem silenciamento, e assim gerar uma árvore filogenética a partir de uma região de interesse na atividade formadora de Hz.

Também é importante conhecer se a alfa-glucosidase de *R. prolixus* realiza a atividade formadora de Hz independente de outra molécula, ou precisa de algum cofator, por exemplo o ambiente lipídico da MPMV, como no caso do complexo protéico que dá a formação da Hz no *Plasmodium falciparum* (Chugh et al., 2013). Só assim pode-se atribuir as adaptações para hematofagia no *R. prolixus* exclusivamente à evolução da alfa-glucosidase no intestino médio.

## 6. CONCLUSÕES

- As regiões F e G da árvore filogenética caracterizam-se por apresentarem a maior quantidade de alfa-glucosidases com a tríade catalítica completa. Também, estas regiões têm totalmente conservados os resíduos YDH, característicos do clã H das glucosilhidrolases. A região F só tem insetos e a região G só tem insetos hematófagos. Estes resíduos conservados podem cumprir uma função importante além da atividade alfa-glucosidase nas proteínas destes insetos.
- Os resíduos marcadores estruturais da família alfa amilase e da subfamília alfa-glucosidase (Seção 4.1) indicam que a proteína Rpro|ACI96031 apresenta atividade alfa-glucosidase.
- A região conservada no alinhamento múltiplo de sequências que tem alta concentração de resíduos de ácido aspártico e ácido glutâmico é uma região importante para revisar nas próximas análises com programas de predição de interação estrutural das alfa-glucosidases (que tem atividade formadora de Hz) com o grupo heme.
- A conservação estrutural do domínio A das alfa-glucosidases de *R. prolixus*, indica que esta estrutura de barril (alfa/beta)<sub>8</sub> foi utilizada no curso da evolução, para gerar a nova atividade de formação de Hz.
- As alfa-glucosidases aqui estudadas caracterizam-se pela baixa similaridade entre si, impedindo uma análise mais detalhada das regiões conservadas. Além disso, apenas algumas proteínas têm dados de expressão em tecidos. Por isto, o mais indicado neste momento é fazer o silenciamento dos genes identificados e associar à atividade formadora de Hz via RNAi nas alfa-glucosidases de *R. prolixus* e assim obter alguma região conservada no alinhamento múltiplo destas proteínas silenciadas.



## 7. REFERÊNCIAS BIBLIOGRÁFICAS

**AFT, RL E MULLER, GC. (1984)** Heme – mediated DNA strand scission. *J. Biol. Chem.* 258: 12069-12072.

**AGGARWAL K, SILVERMAN N. (2008).** Positive and negative regulation of the *Drosophila* immune response. *BMB Re p* 41: 267 – 277

**ANDRIES, JC. E TORPIER, G. (1982)** An extracellular brush border coat of lipid membranes in the midgut of *Nepacineria* (Insecta:Heteroptera): ultrastructure and genesis. *Biol. Cell.* 46: 195-202.

**BAERWALD, JC E DECARPIO, JB. (1983)** Double membrane-bound intestinal microvilli in *Oncopeltus fasciatus*. *Cell. Tiss. Res.* 232:593-600 .

**BENDRAT, K, BERGGER, BJ, CERAMI, A. (1995)** Haem polymerization in malaria. *Nature.* 378: 138.

**BILLINGSLEY, PF E DOWNE, AER. (1989)** The effects of artificial diets on the interior intestinal cell ultrastructure of *Rhodnius prolixus*. *Inter. J. Parasitol.* 19 (3): 291-299.

**BORGES F.M., ROBERTO J., SOUZA L, SANTOS B, DANSA-PETRETSKI M, (2009)** Alpha-Glucosidase Promotes Hemozoin Formation in a Blood-Sucking Bug: An Evolutionary History, Vol. 4 Issue 9

**BOUDREAUX, H. B. (1979).** *Arthropod Phylogeny with Special Reference to Insects.*Wiley; New York; viii\_320 pp.

**BRYLINSKI M., AND SKOLNICK J.(2011).** FINDSITE-metal: Integrating evolutionary information and machine learning for structure-based metal-binding site prediction at the proteome level. *Proteins;* 79:735–751.

**CHOU, AC E FITCH, CD. (1992).** Heme polymerase: modulation by chloroquine treatment of a rodent malaria. *Life Sci.* 51 (26):.2073-8.

**CHUGH Monika, Vidhya Sundararamana, Saravanan Kumar, Vanga S. Reddy (2013)** Protein complex directs hemoglobin-to-hemozoin formation in *Plasmodium falciparum*. [www.pnas.org/cgi/doi/10.1073/pnas.1218412110](http://www.pnas.org/cgi/doi/10.1073/pnas.1218412110)

**DANSA-PETRETESKI, M; RIBEIRO, JMC.; ATELLA, GC MASSUDA, H & OLIVEIRA, PL. (1995).** Antioxidant role of *Rhodnius prolixus* heme-binding protein:

protection against heme-induced lipid peroxidation. *J. Biol. Chem.* 270(18): 10893-10896.

**DILLON, K. & KORDY E. (1997).** Carbohydrate digestion in sandflies: alphasglucosidase activity in the midgut of *Phlebotomus langeroni*. *Comp Biochem Physiol B Biochem Mol Biol.* 116(1): 35-40.

**DORN, A, STOFFEL, R, MATLLE, H, BUBENDORF, A., RIDLEY, RG. (1995).** Malarial haemozoin/ $\beta$ -haematin supports haem polymerization in the absence of protein. *Nature.* 374: 269-271.

**EGAN, TJ (2008)** Haemozoin formation. *Mol. & Biochem. Parasit.*157(2):127-136.

**EMMERSON, KC.; KIM, KC, PRICE, RD. (1973).** Lice. In R.J. Flynn (ed), *Parasites of Laboratory Animals.* Ames, Iowa State University Press.

**ESPINOZA-FUENTES FP & TERRA WR. (1986).** Properties of larval and imaginal membrane-bound digestive enzymes from *Trichosiu pubescens*. *Arch. Insect. Biochem. Physiol.* 3: 181-192.

**FITCH, CD, CAI, GZ, CHEN, YF, SHOEMAKER, JD. (1999).** Involvement of lipids in ferriprotoporphyrin IX polymerization in malaria. *Biochim. Biophys. Acta.* 1454:31-37.

**GOODCHILD, A. (1966).** Evolution of the alimentary canal in the Hemiptera. *Biol. Rev.* 41: 97-140.

**GRAÇA-SOUZA AV, SILVA-NETO MA, OLIVEIRA PL. (1999).** Urate synthesis in the blood-sucking insect *Rhodnius prolixus*. Stimulation by hemin is mediated by protein kinase C. *J. Biol. Chem.* 274(14):9673-6.

**GUTTERIDGE, JMC & SMITH, AJ. (1988).** Antioxidant protection by haemopexin of haem – stimulated lipid peroxidation. *Biochemistry.*, 256: 861-865.

**HOUSEMAN, JG, DOWNW, AER, MORRISON, PE. (1985).** Similarities in digestive proteinase production in *Rhodnius prolixus* (Hemiptera: Reduviidae) and *Stomoxys calcitrans* (Diptera: Muscidae). *Insect Biochem.* 15: 471-474.

**JANECEK Stefan, Birte Svensson and E. Ann MacGregor (2003)** Relation between domain evolution, specificity, and taxonomy of the alpha-amylase family members containing a C-terminal starch-binding domain. *Eur. J. Biochem.* 270, 635–645.

**JANI D, NAGARKATTI R, BEATTY W, ANGEL R, SLEBODNICK C ET AL. (2008)** HDP – a novel heme detoxification protein from the malaria parasite. *PLoS Pathog* 4(4):e1000053.

**KRISTENSEN, J. (1999).** *Handbuch der Zoologie: Eine Naturgeschichte der Stämme des Tierreiches: Band IV: Arthropoda: Insecta: Teilband 35: Lepidoptera, Moths and*

*Butterflies: Volume 1: Evolution, Systematics, and Biogeography.*; The non-glossatan moths. Pp. 41–50. In N. P. (ed.), Berlin, Germany; x\_491 pp.

**KROGH A.** and **SONNHAMMER E. (2004)** A combined transmembrane topology and signal peptide prediction method. *J. Mol. Biol.* 338, 1027–1036.

**KUMAR V., (2010)** Analysis of the key active subsites of glycoside hydrolase 13 family members. *Carboh. Res.* 345, 893–898

**KUMAR V. (2011)** Identification of the sequence motif of glycoside hydrolase 13 family members *Bioinformation* 6(2): 61-63.

**LANE, NJ & HARISSON, JB. (1979).** An unusual cell surface modification: a Double plasma membrane. *J. Cell. Sci.* 39: 355-372

**LEHANE, MJ. (1991)** Harper Collin Academic, 77-85, *Biology of blood-sucking insects.* Fulham Palace Road, Hammersmith, London, w6 8jB – UK.

**LEHANE, M. (2005).** *Biology of blood-sucking insects.* Harper Collins Academic, London, 427p.

**LIU, SC, ZHAI, S, LAWLER, J, PALEK, J. (1985).** Heme – mediated dissociation of erythrocytes membrane skeletal proteins. *J. Biol. Chem.* 260: 12234-12239.

**LOVERING, A. L.; LEE, S. S.; KIM, Y.; WITHERS, S. G. E STRYNADKA, C. J. (2005)** Mechanistic and Structural Analysis of a Family 31 Glycosidase and Its Glycosyl-enzyme Intermediate. *J. Biol. Chem.* v.280 n.3 p.2105–2115.

**LUKASHEVICH, ED & MOSTOVSKI, MB. (2003).** Hematophagous insects in the fóssil record. *Paleont. J.* 37: 153 – 161.

**MCGREGOR E., JANECEK B, SVENSSON B. (2001).** Relationship of sequence and structure to specificity in alpha-amylase. *Bioch. et Biophys. Acta* 1546, 1 – 20.

**MANS, BJ. & NEITZ, AWH (2004).** Adaptation of ticks to a blood-feeding environment: evolution from a functional perspective. *Ins. Bioch. Mol. Biol.* 34: 1-17

**MARSHALL, AG. (1981).** *The ecology of ectoparasitic insects.* Academic Press, London, 459p.

**MOHR, CO. (1943).** Cattle droppings as ecological units. *Ecol. Monographs* 13: 275

**MOREIRA, LM; RIBELATTO, JC; IMASATO, H. (2004)** Ruffled and planar conformations of the porphyrin ring in complexes and heme proteins: physical-chemistry properties and spectroscopic implications. *Quím. nova* 27(6): 958-963.

**MORGAN, MRJ. (1975)** A qualitative survey of the carbohydrases of the alimentary tract of the migratory locust, *Locusta migratoria migratorioides*. *J. Insect. Physiol.* 21: 1045–1053.

- MURY** FB, DA SILVA JR, FERREIRA LS, DANSA-PETRETSKI M. (2009). Alpha-glucosidase promotes hemozoin formation in a blood-sucking bug: an evolutionary history. PLoS One 4(9):.
- OLIVEIRA**, MF, SILVA, JR, DANSA-PETRETSKI, M, SOUZA, W, LINS, U, BRAGA, C M S, MASUDA, H, OLIVEIRA, PL. (1999). Haem detoxification by an insect. Nature. 400: 517-518
- OLIVEIRA**, MF, SILVA, JR, DANSA-PETRETSKI, M, DE SOUZA, W, LINS, U, BRAGA, CMS, MASUDA, H, OLIVEIRA, PL (2000). Hemozoin formation in the midgut of the blood – sucking insect *Rhodnius prolixus*. FEBS Lett. 477:95-98
- PAES**, MC, OLIVEIRA, PL. (1999) Extracellular glutathione peroxidase from the blood-sucking bug, *Rhodnius prolixus*.. Archives of Insect Biochemistry and Physiology. 41(4):171-177
- PAIVA-SILVA**, GO. SORGINE, MHF, BENEDETTI, C, MENEGHINI, R, ALMEIDA, IC, MACHADO, EA, PETRETSKI, MDA, LAW, JH, OLIVEIRA, PL, MASUDA, H. (2002). On the biosynthesis of *Rhodnius prolixus* heme-binding protein. Insect Biochemistry and Molecular Biology. 32:1534-1541.
- PAIVA-SILVA**, GO, OLIVEIRA, CC, MONTEIRO, CM, NAKAYASU, ES, DUNKOV, B, MASUDA, H, ALMEIDA, IC, OLIVEIRA, PL. (2006). A heme degradation pathway in a blood sucking insect. PNAS. Proceedings of the National Academy of Sciences of the United States of America. 103(21):8030-8035.
- PASCOA**, V, OLIVEIRA, PL, DANSA-PETRETSKI, M, SILVA, JR, ALVARENGA, PH, JACOBS-LORENA, M, LEMOS, FJ. (2002). *Aedes aegypti* peritrophic matrix and its interaction with heme during blood digestion. Insect Biochem Mol. Biol.32:517-523.
- PATTERSON** James S., Michael W. Gaunt (2010) Phylogenetic multi-locus codon models and molecular clocks reveal the monophyly of haematophagous reduviid bugs and their evolution at the formation of South America. Molecular Phylogenetics and Evolution 56 (2010) 608–621.
- PEARMAN**, JV. (1960). Some African Psocoptera found on rats. Entomologist 93:246-250.
- PUJADAS** Gerard \*&Jaume Palau (2002) Independent folding of the A and B domains In the-amylase family. Biologia, Bratislava, 57/Suppl. 11: 43-57.
- REGER**, JF. (1971). One structure of the surface coat of midgut epithelial cells in the homopteran *Phylloscelis atra* (Fulgorid). J. Submicrosc. Cytol. 3:353-358

- RIBEIRO, JM (1995).** Blood-feeding arthropods: Live syringes or invertebrate pharmacologists? *Insect. Agent. Dis.* 4:143-152
- SILVA, CP, RIBEIRO, AF, GULBENKIAN, S, TERRA, WR. (1995).** Organization, origin and function of the outer microvillar (perimicrovillar) membranes of *Dysdercus peruvianus* (Hemiptera) midgut cells. *J. Insect. Physiol.* 41:1093-1103
- SILVA, JR. (2001).** Associação das membranas perimicrovillares do intestino médio de *Rhodnius prolixus* com o processo de polimerização de heme. Dissertação de mestrado. Universidade Estadual do Norte Fluminense. Programa de Biociências e Biotecnologia.
- SILVA, CP, SILVA, JR, VASCONCELOS, FF, DANSA-PETRETSKI, M, DA MATTA, RA, RIBEIRO, AF. (2004).** Occurrence of midgut perimicrovillar membranes in paraneopteran insect orders with comments on their function and evolutionary significance. *Arthr. Struct. Develop.* 33:139-148
- SILVA, JR, MURY, FB, OLIVEIRA, MF, OLIVEIRA, PL, SILVA, CP, DANSA-PETRETSKI, M. (2007).** Perimicrovillar membranes promote hemozoin formation into *Rhodnius prolixus* midgut. *Insect Biochem. Molec. Biol.* 37:523-531.
- SLATER, AFG, SWIGGARD, WJ, ORTON, BR, FLITTER, WD, GOLDBERG, E, CERAMI, A, HENDERSON, GB. (1991).** An iron – carboxylate bond links the heme units of malaria pigment. *Proc. Natl. Acad. Sci. USA.* 88:325-329
- SLATER, AF E CERAMI, A. (1992).** Inhibition by cloroquine of a novel haem polymerase enzyme activity in a malaria trophozoites. *Nature.* 355:167-169
- SMITH, TS, MUNN, EA, GRAHAM, M. (1993).** Purification and evaluation of the integral membrane protein H11 as a protective antigen against *Haemonchus contortus*. *Int. J. Parasitol.*, 23:271-280.
- STAM Mark R., Etienn e G.J.Dan chin, Cor inne Rancurel, (2006)** Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of alpha-amylase related proteins. *Protein Engineering, Design & Selection* vol. 19 no. 12 pp. 555–562.
- SULLIVAN, DJ JR, GLUZMAN, IY, GOLDBERG, DE. (1996).** Hemozoin formation mediated by histidine – rich proteins. *Science.* 271:219-221.
- TEKWANI, BL E WALKER, LA (2005).** Targeting the hemozoin synthesis pathway for new antimalarial drug discovery: Technologies for In Vitro -Hematin Formation Assay. *Bent. Sci. Pub.* 8:63-79

- TERRA, WR, FERREIRA, C, DE BIANCHI, AG. (1979).** Distribution of digestive enzymes among the endo and ectoperitrophic spaces and midgut cells of *Rhynchosciara* and its physiological significance. *J. Insect Physiol.* 25:423 – 434
- TERRA, WR. (1988).** Physiology and biochemistry of insect digestion: an evolutionary perspective, *Brazilian J. Med. Biol. Reser.* 21: 675-734
- TERRA, WR E JORDÃO, BP. (1989).** Final digestion of starch in *Musca domestica* larvae. Distribution and properties of midgut alpha-glicosidases and gluco-amylase. *Insect Biochem.* 19:285-292
- TERRA, WR. (1990).** Evolution of digestive systems of insects. *Ann. Rev. of Entomol.* 35:181-200
- TERRA WR E JORDÃO BP. (1991).** Regional distribution and substrate specificity of digestive enzymes involved in terminal digestion in *Musca domestica* hind – midgut. *Arch. Insect Biochem. Physiol.* 17:157-168
- TERRA. WR E FERREIRA, C. (1994).** Insect digestive enzymes: properties, compartmentalization and function. *Comp. Biochem. Physiol.* 109b: 1-62.
- VINCENT, SH, GRADY, RW, SHAKLAI, N, SNIDER, JM, MULLER-EBERHARD, U. (1988).** The influence of heme-binding proteins in heme-catalyzed oxidations. *Arch. Biochem. Biophys.* 265:539-550.
- WIERENGA R.K. (2001)** The TIM-barrel fold : a versatile framework for efficient enzymes. *FEBS Letters* 492. 193-198.
- WISE ERIC L. AND IVAN RAYMENT. (2004)** Understanding the importance of protein structure to nature's routes for divergent evolution in tim barrel enzymes. *Acc. Chem. Res.*, 37, 149-158.
- YLINSKI Michael Br and Jeffrey Skolnick (2010)** FINDSITE-metal: Integrating evolutionary information and machine learning for structure-based metal-binding site prediction at the proteome level. *Proteins*; 79:735–751.
- ZEHEL, DL & WITHERS, SG. (2000).** glycosidase mechanisms: anatomy of a finely tuned Catalyst. *Acc. Chem. Res.* 33:11-18.