DANIELA TIEMI MYAMOTO

**Caracterização e análise filogenética dos genes que codificam para
os componentes C3 e fator B do Sistema Complemento das
glândulas de veneno de aranhas *Loxosceles***

Tese apresentada ao Programa de Pós-Graduação em Imunologia do Instituto de Ciências Biomédicas da Universidade de São Paulo, para obtenção do Título de Doutor em Ciências.

São Paulo
2015

DANIELA TIEMI MYAMOTO

**Caracterização e análise filogenética dos genes que codificam para
os componentes C3 e fator B do Sistema Complemento das
glândulas de veneno de aranhas *Loxosceles***

Tese apresentada ao Programa de Pós-Graduação em Imunologia do Instituto de Ciências Biomédicas da Universidade de São Paulo, para obtenção do Título de Doutor em Ciências.

Área de Concentração: Imunologia

Orientadora: Prof. Dra. Denise V. Tambourgi

Versão original

São Paulo
2015

DADOS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
Serviço de Biblioteca e Informação Biomédica do
Instituto de Ciências Biomédicas da Universidade de São Paulo

reprodução não autorizada pelo autor

Myamoto, Daniela Tiemi.

Caracterização e análise filogenética dos genes que codificam para os componentes C3 e fator B do sistema complemento das glândulas de veneno das aranhas *Loxosceles* / Daniela Tiemi Myamoto. -- São Paulo, 2015.

Orientador: Profa. Dra. Denise Vilarinho Tambourgi.

Tese (Doutorado) – Universidade de São Paulo. Instituto de Ciências Biomédicas. Departamento de Imunologia. Área de concentração: Imunologia. Linha de pesquisa: Filogenia do sistema completo.

Versão do título para o inglês: Characterization of genes coding for the components C3 and factor B of complement system from *Loxosceles* spiders venom gland.

1. Filogenia 2. Complemento imunológico 3. Aranhas 4. Genes da resposta imune 5. Reação em cadeia por polimerase 6. Clonagem
I. Tambourgi, Profa. Dra. Denise Vilarinho II. Universidade de São Paulo. Instituto de Ciências Biomédicas. Programa de Pós-Graduação em Imunologia III. Título.

ICB/SBIB0121/2015

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS

Candidato(a): Daniela Tiemi Myamoto.

Título da Tese: Caracterização e análise filogenética dos genes que codificam para os componentes C3 e fator B do sistema complemento das glândulas de veneno das aranhas *Loxosceles*.

Orientador(a): Profa. Dra. Denise Vilarinho Tambourgi.

A Comissão Julgadora dos trabalhos de Defesa da Tese de Doutorado, em sessão pública realizada a/...../....., considerou

() **Aprovado(a)**

() **Reprovado(a)**

Examinador(a): Assinatura:
Nome:
Instituição:

Examinador(a): Assinatura:
Nome:
Instituição:

Examinador(a): Assinatura:
Nome:
Instituição:

Examinador(a): Assinatura:
Nome:
Instituição:

Presidente: Assinatura:
Nome:
Instituição:

SECRETARIA DE ESTADO DA SAÚDE
INSTITUTO BUTANTAN

São Paulo, 18 de junho de 2010.

Ilma. Sra.
Dra. Denise Vilarinho Tambourgi
Diretora do Laboratório de Imunoquímica do Instituto Butantan

Prezado Dra. Denise,

Comunicamos que a Comissão Interna de Biosegurança do Instituto Butantan -CIBio-IB, autoriza o desenvolvimento do projeto de pesquisa "Clonagem e caracterização dos genes que codificam para os componentes C3 e fator B do Sistema Complemento das glândulas de veneno de aranhas *Loxosceles laeta*", dentro do Certificado de Qualidade em Biosegurança 039/98 – CQB 039/98, ao Laboratório de Imunoquímica com Nível de Biosegurança 2, NB 2 (<http://www.ctnbio.gov.br/index.php/content/view/3143.html> na seção de Documentos Publicados no DOU, extrato N° 075 de 2003, publicado em 09/01/2004), por ter sido o projeto classificado como sendo de Nível de Biossegurança I (NBI).

Atenciosamente,

Viviane Fongaro Botosso
Presidente
CIBio-Instituto Butantan

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS

Cidade Universitária "Armando de Salles Oliveira"
Av. Prof. Lineu Prestes, 2415 – CEP. 05508-000 São Paulo, SP Brasil
Telefone : (55) (011) 3091.7733 – e-mail: cep@icb.usp.br

CERTIFICADO

Certificamos que o protocolo registrado sob nº **189** nas fls. **139** do livro **02** para uso de animais em Aula, sob a responsabilidade do Prof(a) Dr(a) **Denise Vilarinho Tambourgi**, Coordenador (a) da Linha de pesquisa "*Clonagem e caracterização dos genes que codificam para os componentes C3 e fator B do Sistema Complemento das glândulas de veneno de aranhas Loxosceles*" do qual participam o(s) aluno(s), **Daniela Tiemi Myamoto**, está de acordo com os Princípios Éticos de Experimentação Animal adotado pela Sociedade Brasileira de Ciência de Animais de Laboratório (SBCAL) e foi aprovado pela COMISSÃO DE ÉTICA NO USO DE ANIMAIS (CEUA) em **03.12.2012, com validade de 4 anos**.

São Paulo, 04 de dezembro de 2012.

Prof. Dr. WOTHAN TAVARES DE LIMA
Coordenador-CEUA - ICB/USP

Profa. Dra. ANA PAULA LEPIQUE
Secretária- CEUA - ICB/USP

COMISSÃO DE ÉTICA NO USO DE ANIMAIS
INSTITUTO BUTANTAN

Av. Dr. Vital Brazil, 1500, CEP 05503-900, São Paulo, SP, Brazil
Telefone: (55) (011) 2627-9585 - Fax: (55) (011) 2627-9505
ceuaib@butantan.gov.br

CERTIFICADO

Certificamos que o projeto intitulado "Caracterização e análise filogenética dos genes que codificam para os componentes C3 e fator B do sistema complemento das aranhas *Loxosceles*", **protocolo nº 938/12**, sob a responsabilidade de Denise Vilarinho Tambourgi e Daniela Tiemi Myamoto – que envolve a criação e/ou utilização de animais pertencentes ao filo Chordata, subfilo Vertebrata (exceto o homem), para fins de pesquisa científica – está de acordo com os preceitos da Lei nº 11.794, de 8 de outubro de 2008, do Decreto 6.899, de 15 de julho de 2009 e de normas complementares, bem como está de acordo com os Princípios Éticos na Experimentação Animal adotado pelo Colégio Brasileiro de Experimentação Animal (COBEA), e foi aprovado pela COMISSÃO DE ÉTICA NO USO DE ANIMAIS DO INSTITUTO BUTANTAN (CEUAIB) em reunião de 20/6/12.

We certify that the research entitled "Characterization and phylogenetic analysis of genes coding for C3 and factor B complement system components from *Loxosceles laeta* spider venom glands", **protocol number 938/12**, under the responsibility of Denise Vilarinho Tambourgi and Daniela Tiemi Myamoto – which involves the breeding and/or use of animals belonging to phylum Chordata, subphylum Vertebrata (except human beings), for scientific research – is in agreement with Brazilian laws for use of experimental animals and the Ethical Principles in Animal Research adopted by the Brazilian College of Animal Experimentation, and was approved by the ETHICAL COMMITTEE FOR ANIMAL RESEARCH of BUTANTAN INSTITUTE in the meeting of 6/20/12.

Vigência do Projeto:	Nº de animais/espécie
06/2012 - 02/2014	
Laboratório de Imunoquímica	20 camundongos Balb/c (M)

São Paulo, 22 de junho de 2012

Dr. Marcelo L. Santoro
Coordenador da CEUAIB

**UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS**

Cidade Universitária "Armando de Salles Oliveira"
Av. Prof. Lineu Prestes, 2415 - cep. 05508-000 São Paulo, SP - Brasil
Telefone : (55) (11) 3091.7733 telefax : (55) (11) 3091-8405
e-mail: cep@icb.usp.br

São Paulo, 25 de março de 2013.

PARECER 1106/CEP

A Comissão de *Ética em Pesquisas em Seres Humanos* do ICB, na sessão realizada no dia 27.02.2013, **APROVOU** o projeto intitulado: "*Clonagem e Caracterização dos Gens que Codificam para os Componentes C3 e fator B do Sistema Complemento das Glândulas de Veneno das aranhas Loxosceles*" das pesquisadoras **DENISE VILARINHO TAMBOURGI** e aluna **DANIELA TIEMI MYAMOTO**.

Cabe aos Pesquisadores executantes elaborar e apresentar a este Comitê, relatórios anuais (parciais ou final), de acordo com a resolução 196/06 do Conselho Nacional da Saúde, item IX. 2 letra c. conforme modelo constante no site.icb.usp.br.

Ao pesquisador cabe também finalizar o processo junto à Plataforma Brasil quando do encerramento deste.

O primeiro relatório deverá ser encaminhado à Secretaria deste CEP em **27.02.2014**.

Atenciosamente,

Prof. Dra. PAOLO M.A. ZANOTTO
Coordenador da Comissão de Ética em
Pesquisas com Seres Humanos - ICB/USP

AUTORIZAÇÃO DE ACESSO E DE REMESSA DE AMOSTRA DE COMPONENTE DO PATRIMÔNIO GENÉTICO nº 010328/2014-9

O CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO - CNPq, credenciado pelo Conselho de Gestão do Patrimônio Genético (CGEN/MMA), por meio da Deliberação CGEN nº 246, de 27 de agosto de 2009, para autorizar instituições nacionais, públicas ou privadas, que exerçam atividades de pesquisa e desenvolvimento nas áreas biológicas e afins, a acessar e remeter amostras de componente do patrimônio genético para fins de pesquisa científica sem potencial de uso econômico, neste ato representado pelo seu Diretor de Ciências Agrárias, Biológicas e da Saúde, nos termos da Portaria CNPq nº 104/2011, autoriza a instituição abaixo qualificada a acessar e remeter amostras de componentes do patrimônio genético.

Instituição: INSTITUTO BUTANTAN - IBU

CNPJ: 618.213.440/0001-56

Representante Legal: JORGE ELIAS KALIL FILHO

Cargo/Função: Diretor

CPF: 177.516.590-68 **RG:** 8004961771

Projeto: Clonagem e caracterização dos genes que codificam para os componentes C3 e fator B do Sistema Complemento das glândulas de veneno de aranhas *Loxosceles laeta*

Coordenador do Projeto: Denise Vilarinho Tambourgi

CPF: 050.043.288-03 **RG:** 76676948 - SSP / SP

Finalidade do projeto: O sistema complemento humano é composto por mais de 30 proteínas e muitas destas apresentam domínios conservados o que possibilita traçar a evolução filogenética com base em suas estruturas. Os estudos dos genomas de vertebrados e invertebrados mostraram que muitos dos domínios presentes em componentes do sistema complemento de mamíferos estavam presentes em deutetostômios e também em protostômios. A origem do sistema se dá provavelmente com o surgimento de C3 e fator B (fB), os únicos encontrados em alguns protostômios. Recentemente, foram identificadas no transcriptoma da glândula de veneno da aranha *Loxosceles laeta*, sequências gênicas homólogas as de C3 e fator B de *Carcinocorpius rotundicauda* e *Strongylocentrotus purpuratus*, respectivamente. Tais achados sugerem que os componentes centrais do Complemento possam estar expressos na glândula de veneno das aranhas marrom, desempenhando talvez um papel na defesa.

Amostras a serem acessadas:

Grupos Taxonômicos: *Loxosceles laeta*

Tipo de material/quantidade de amostras: Nenhuma amostra será coletada atualmente. Foram coletados 100 pares de glândulas no ano de 2010 sob prévia autorização (Registro IBAMA 1/35/95/1561 Processo nº 02027016615/95-92). As amostras utilizadas são todas provenientes destas glândulas.

Local de depósito de subamostra: INSTITUTO BUTANTAN

Equipe do projeto: DENISE VILARINHO TAMBOURGI / CPF 050.043.288-03

DANIELA TIEMI MYAMOTO / CPF 317.238.288-58

Validade da Autorização: 29/05/2014 a 02/10/2015

A instituição acima qualificada deverá enviar ao CNPq, por meio do Coordenador do Projeto, relatório anual sobre o andamento do projeto de pesquisa, nos termos do Decreto nº. 4.946/2003. O roteiro para confecção do relatório está disponível em <http://www.cnpq.br/web/guest/relatorio-de-atividades>. Os relatórios devem ser enviados ao CNPq em meio eletrônico, para o endereço apg@cnpq.br e, preferencialmente, em formato .pdf.

Esta autorização está vinculada às informações, declarações e termos de compromisso firmados pelo coordenador do projeto e pelo representante legal, constantes do Processo nº 010328/2014-9. Atividades de acesso aos conhecimentos tradicionais associados, de acesso e de remessa de componente do patrimônio genético com finalidade comercial, aplicação industrial, bioprospecção ou desenvolvimento tecnológico não estão autorizadas.

Caso seja identificado uso econômico de produto ou processo, passível ou não de proteção intelectual, originado das

Página 1 de 2

amostras de componente do patrimônio genético acessado no âmbito desta autorização, a instituição beneficiada se compromete a adotar as providências cabíveis, nos termos da legislação vigente, junto ao CGEN/MMA.

Se ocorrer coleta de espécie não autorizada ou não identificada, deverá ser observado o que consta no Decreto nº 6.514, de 22/07/2008, no que refere à flora e fauna, e em particular sobre espécies ameaçadas de extinção ou de endemismo estrito.

A remessa de amostra de componente do patrimônio genético deverá ser precedida da assinatura do Termo de Transferência de Material (TTM) ou do Termo de Responsabilidade para Transporte de Amostra de Componente do Patrimônio Genético (TRTM). A remessa para instituições nacionais está isenta de autorização prévia. Contudo, a remessa para instituições sediadas no exterior depende de autorização prévia do CNPq, nos termos das resoluções do CGEN 15/2004 e 20/2006. Os modelos dos termos, assim como as citadas resoluções, estão disponíveis em <http://www.cnpq.br/web/guest/remessa-e-transporte> e devem ser enviados ao CNPq em meio eletrônico para o endereço apg@cnpq.br, preferencialmente em formato .pdf. Ainda, para a remessa de componente do patrimônio genético para instituição sediada no exterior, deverá ser solicitada ao Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis - IBAMA, por meio de formulário específico e mediante a apresentação de TTM ou TRTM, licença de exportação complementar a autorização de remessa, especialmente quando se tratar de remessa de espécies constantes nos Anexos da Convenção sobre o Comércio Internacional das Espécies da Flora e Fauna Selvagens em Perigo de Extinção (Cites).

Brasília, 29 de Maio de 2014

Marcelo Marcos Morales

Diretor de Ciências Agrárias, Biológicas e da Saúde

Para visualizar a versão digital da Autorização de Acesso e de Remessa de Amostra de Componente do Patrimônio Genético, V.Sa. poderá utilizar a ferramenta disponibilizada pelo CNPq para esse fim na página <http://servicosweb.cnpq.br/visualizador/> e informar o número do protocolo 2775021067439658 para recuperá-la do banco de dados do CNPq, ou poderá selecionar o arquivo salvo em seu computador (em formato PKCS7). V.Sa. pode também usar outro aplicativo disponível no mercado capaz de reconhecer arquivos no padrão PKCS7 para fazer a visualização e extração do documento.

Aos meus pais, irmãos e todos os amigos por
todo o apoio, amizade e ensinamentos.

AGRADECIMENTOS

À minha orientadora, Dra. Denise V. Tambourgi, primeiramente, por ter me aceitado a fazer parte do grupo de pesquisa e por toda orientação, dedicação, ensinamentos e paciência que teve comigo desde o começo, imprescindíveis para minha formação acadêmica. Obrigada pela oportunidade de vivenciar uma parte do universo da ciência e por todas as outras que me ajudaram a amadurecer pessoal e profissionalmente.

À minha família, pais e irmãos e ao Mingau (meu cachorro) por todo apoio, amor, incentivo e por ter me aguentado durante todos esses anos.

À Dra. Giselle Pidde Queiroz por todos os ensinamentos em biologia molecular, pelas discussões, sugestões que foram importantes para o andamento do projeto, por sempre estar disposta a me ajudar e pela amizade.

À Dra. Rute Maria Gonçalves de Andrade pelos ensinamentos em biologia de aranhas *Loxosceles*, por toda a parte referente à identificação, coleta e manutenção das aranhas e pela extração das glândulas de veneno, essenciais para o desenvolvimento do projeto.

Ao Dr. Aurélio Pedroso pelos ensinamentos na área de filogenia e biologia molecular, por ter oferecido muitas ideias e sugestões, essenciais para o refinamento do projeto, por todo o incentivo e amizade.

Às secretárias e grandes amigas do Laboratório de Imunoquímica, Elaine Rodrigues e Lia Aguiar, por todo o apoio e por ajudarem a deixar o ambiente do laboratório mais descontraído e alegre!

Aos pesquisadores do Laboratório de Imunoquímica do Instituto Butantan: Wilmar Dias da Silva, Osvaldo Augusto Sant'Anna, Carla Squaiella Baptisdão, Fernanda Calheta Vieira Portaro e Fábio Carlos Magnoli.

A toda a equipe de funcionários que já passaram e os que ainda estão do laboratório de Imunoquímica do Instituto Butantan: Sílvia, Maura, Márcia, Ana Freire, Ana Cláudia, Guilherme, Ricardo, Osmair, Alécio, Rosana, Severino Ramos, João e Cassiano. Obrigada por todo o apoio, dedicação e por terem contribuído para a realização da pesquisa!

À turma de amigos composta por pesquisadores, pós-doutorandos e estudantes do Laboratório de Imunoquímica: Ana Tung, Angela, Aurélio, Carla, Felipe França, Giselle, Isadora, Joel, Lygia, Marie e Priscila Hess. Obrigada por toda ajuda e apoio para minha formação e também por todas as conversas e momentos agradáveis e descontraídos, principalmente durante o almoço e nas paradas para o café após o almoço.

À turminha de amigos, em especial, que cultivei desde a minha entrada do laboratório e outros que entraram depois e que estão firmes e fortes até hoje, me apoiando, ajudando no que for preciso. São eles: Cinthya, Gabi, Elaine, Lia, Guilherme e Denise Santiago. Obrigada por todo o carinho, pelos momentos de convivência e pela amizade calorosa.

Aos ex-estudantes e amigos que já passaram pelo Laboratório e fazem muita falta por aqui: Daniel, Danielle, Estevam, Felipe Guidolin, Karina, Mariana, Priscila Lara, Verônica e Vivian.

À Tatiana Rodrigues Fraga, que conheci melhor durante as viagens para participação em congressos e que se tornou uma grande amiga. Obrigada por toda a ajuda e torcida!

Às grandes amigas de república que fiz durante a faculdade: Dani Hana, Marina, Mariana e Mawusi. Obrigada por continuarem fazendo parte da minha vida, mesmo que os momentos de convivência tenham diminuído! Mas é sempre muito especial os nossos reencontros!

Aos amigos da turma do cursinho da Poli de longa data que foram essenciais para minha formação e crescimento pessoal e também em muitas das minhas escolhas e que continuam me apoiando e me proporcionando muitas alegrias.

Aos amigos taiwaneses e agregados, em especial, a família Tseng e Lin que já considero como parte da minha família por todo o apoio, amizade e por serem grandes referências de valores que aprendi muito a apreciar.

Aos participantes da banca de qualificação pelas sugestões, conselhos e pela comissão julgadora da defesa de doutorado.

Ao Programa de Pós Graduação em Imunologia e a todos os professores que fazem parte dele, contribuindo muito para minha formação.

À secretária de Pós-Graduação do Departamento de Imunologia, Maria Eni, por todo o suporte durante meu doutorado.

Aos funcionários da Biblioteca do ICB-USP pelas orientações e revisões textuais.

À FAPESP, CNPq e INCTox pelo apoio financeiro e pela oportunidade de desenvolver o projeto de pesquisa.

Obrigada a todos!

“Ninguém ignora tudo, ninguém sabe tudo. Por
isso, aprendemos sempre.”

(Paulo Freire)

RESUMO

MYAMOTO, D. T. **Caracterização e análise filogenética dos genes que codificam para os componentes C3 e fator B do Sistema Complemento das glândulas de veneno de aranhas *Loxosceles***. 2015. 110 f. Tese (Doutorado em Imunologia) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2015.

O sistema complemento humano é composto por mais de 30 proteínas, sendo que muitas destas apresentam domínios conservados o que possibilita traçar a evolução filogenética dos componentes deste sistema. Tal sistema parece ter surgido com o aparecimento de C3 e fator B (FB), únicos componentes encontrados em alguns protostômios e cnidários, o que sugere que a via alternativa seja a mais primitiva. Recentemente, foram identificadas, no transcriptoma da glândula de veneno da aranha *Loxosceles laeta*, sequências gênicas similares as dos componentes C3 e FB de invertebrados das famílias Limulidae e Strongylocentrotidae, respectivamente. Tais achados sugerem que os componentes centrais do complemento possam estar presentes na glândula de veneno das aranhas marrom desempenhando, talvez, papel de defesa. Neste projeto, fragmentos de cDNA codificantes para os componentes fator B (Lox-FB) e C3 (Lox-C3) foram obtidos, a partir do RNA total isolado da glândula de veneno de *L. laeta*, que foram amplificados por técnicas de RACE-PCR, *primers* específicos e degenerados. As sequências de cDNA obtidas, e suas respectivas sequências deduzidas de aminoácidos, foram analisadas com ferramentas de bioinformática. Análises dos clones 5' e 3' de Lox-FB revelou um cDNA com 1953 pares de bases, codificante para uma proteína madura de 651 aminoácidos. Lox-FB apresenta uma organização de domínios clássica, com dois domínios CCP (*Complement Control Proteins*), um domínio vWFA (*von Willebrand factor domain type A*) e um domínio de serino protease (SP). Os aminoácidos envolvidos na ligação ao C3b, dependente de íons Mg^{+2} (MIDAS), foram bem conservados, no entanto, a tríade catalítica composta pelos aminoácidos histidina, ácido aspártico e serina não foi encontrada no domínio SP. Para Lox-C3, foi identificado um cDNA com 5178 pares de bases, codificante para uma proteína de 1726 aminoácidos, com uma configuração de domínios similar àquela da sequência primária do C3 humano. O alinhamento múltiplo entre as proteínas C3-like revelaram a presença de um sítio de processamento entre as cadeias α e γ , além do sítio de processamento entre as cadeias α e β , indicando que Lox-C3 seja composto por três cadeias. Para melhor entender a história evolutiva dos componentes Lox-C3 e Lox-FB de *L. laeta*, foi realizada uma análise filogenética que indicou que Lox-C3 pertence ao grupo em que estão também presentes duas isoformas de C3 da aranha *Hasarius adansoni* (Família Salcitidae), com a qual obteve maior similaridade (53%). Com relação às proteínas pertencentes à família C2/FB, Lox-FB foi incluído no grupo em que estão presentes componentes FB-like de anêmonas, anfioxos, moluscos e outras três espécies do Filo Arthropoda, com as quais apresentou uma maior similaridade (43%) e uma relação evolutiva mais próxima com a terceira isoforma da aranha *Hasarius adansoni*, aparentemente inativa. Este trabalho representa a primeira caracterização de dois genes que compõem o sistema complemento das aranhas *Loxosceles* que poderá contribuir para melhor entendimento do sistema imune inato dessas aranhas.

Palavras-chave: *Loxosceles laeta*. Sistema complemento C3. FB. Filogenia. Clonagem.

ABSTRACT

MYAMOTO, D. T. **Characterization and phylogenetic analysis of genes coding for the components C3 and factor B of Complement System from *Loxosceles* spiders venom glands.** 2015. 110 p. Ph.D. Thesis (Immunology) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2015.

The human complement system is composed by more than 30 proteins and many of them have conserved domains that allow tracing the phylogenetic evolution. The complement system seems to be initiated with the appearance of C3 and factor B (FB), the only components found in some protostomes and cnidarians, suggesting that the alternative pathway is more ancient. Recently, they have been identified, in the transcriptome of *Loxosceles laeta* spiders venom glands, C3- and FB-like genes of invertebrates of Limulidae and Strongylocentrotidae Family, respectively. These findings suggest that the central components of the complement system can be expressed on the venom gland from the brown spiders, maybe playing a role in the defense mechanisms. In this project, cDNA fragments coding for the complement factor B (Lox-FB) and C3 (Lox-C3) were synthesized from total RNA *L. laeta* spiders venom gland and amplified using RACE-PCR techniques, specific and degenerate primers. The resulting cDNA sequences and their respective protein sequences were analyzed using bioinformatics tools. Analyses of 5' and 3' clones of Lox-FB allowed identifying a cDNA composed by 1953 base pairs that coding for a mature protein of 651 amino acids. Lox-FB has a classical domain organization in mosaic, composed by two CCPs (Control Complement Proteins), a vWFA domain (von Willebrand Factor domain) and a SP domain (Serine Protease). The amino acids involved in binding to C3b dependent of Mg^{2+} metal ion (MIDAS) are well conserved, however, the catalytic triad composed of amino acids histidine (H), aspartic acid (D) and serine (S) were not found in SP domain. For Lox-C3, it was identified a cDNA composed by 5178 base pairs that coding for a protein of 1726 amino acids with a configuration of domains similar to those primary sequence of Human C3. The multiple alignment between C3-like proteins revealed the presence of a processing site between the α and β chains, indicating that Lox-C3 is composed by three chains. To better understand the evolutionary history of Lox-C3 and Lox-FB components, the phylogenetic analysis was performed. Lox-C3 placed inside the group where others two isoforms of C3 from spider *Hasarius adansoni* (Família Salticidae), with which obtained major similarity (53%). With respect to proteins belonging to C2/FB Family, Lox-FB grouped inside the clade that FB-like from anemones, amphioxus, mollusks and others three species from Arthropoda Phylum are present, obtaining a major similarity (43%) and a evolutionary close relationship with the third isoform of spider *H. adansoni*, apparently inactive. This work represents the first characterization of genes coding for two components that comprise the *Loxosceles* spiders complement system that can contribute for best understanding of innate immune system of these spiders.

Keywords: *Loxosceles laeta*. Complement system. C3. FB. Phylogeny. Cloning.

LISTA DE ILUSTRAÇÕES

FIGURA 1 – Aranhas <i>Loxosceles</i>	21
FIGURA 2 – Esquema das três vias de ativação do sistema complemento	28
FIGURA 3 – Perspectiva evolutiva para as moléculas de reconhecimento e de serino proteases envolvidas na ativação da cascata do sistema complemento	30
FIGURA 4 – Componente Fator B Humano	32
FIGURA 5 – Componente C3 Humano	34
FIGURA 6 – Seleção de <i>primers</i> degenerados C3-like	39
FIGURA 7 – Análise da sequência EST Lox-FB	45
FIGURA 8 – Amplificação e clonagem do produto Lox-FB da reação 5' RACE-PCR.....	46
FIGURA 9 – Amplificação e clonagem do produto Lox-FB da reação 3' RACE-PCR.....	48
FIGURA 10 – Sequência de nucleotídeos e deduzida de aminoácidos de Lox-FB	50
FIGURA 11 – Sequência deduzida de aminoácidos de Lox-FB	52
FIGURA 12 – Alinhamento dos domínios CCPs do fator B humano (Hu-FB) e Lox-FB	53
FIGURA 13 – Alinhamento múltiplo de Lox-FB	55
FIGURA 14 – Modelo tridimensional de Lox-FB	57
FIGURA 15 – Comparação da estrutura terciária dos domínios vWFA e SP do FB humano ...	58
FIGURA 16 – Árvore filogenética das sequências similares ao Lox-FB.....	61
FIGURA 17 – Análise da sequência EST Lox-C3.	63
FIGURA 18 – Amplificação da região que contém o motivo tioéster de Lox-C3.....	64
FIGURA 19 – Componente C3-like – Lox-C3_TEP	65
FIGURA 20 – Análise da sequência Lox-C3_TEP	66
FIGURA 21 – Sequência parcial Lox-C3	67
FIGURA 22 – Amplificação da região que contém o domínio ANATO de Lox-C3	69
FIGURA 23 – Componente C3-like – Lox-C3_ANATO	70
FIGURA 24 – Amplificação da região terminal de Lox-C3 5'	71
FIGURA 25 – Amplificação e clonagem do produto Lox-C3 da reação 5' RACE-PCR.	72
FIGURA 26 – Amplificação da região Lox-C3 3'	73
FIGURA 27 – Amplificação e clonagem do produto Lox-C3 da reação de 3' RACE-PCR.....	74
FIGURA 28 – Sequência Lox-C3 3'	76
FIGURA 29 – Alinhamento de Lox-C3 5' e Lox-C3 3'	77

FIGURA 30 – Sequências de cDNA e de aminoácidos deduzida de Lox-C3.....	79
FIGURA 31 – Sequência deduzida de aminoácidos de Lox-C3.....	83
FIGURA 32 – Alinhamento múltiplo de Lox-C3	86
FIGURA 33 – Modelo tridimensional de Lox-C3.....	90
FIGURA 34 – Árvore filogenética das sequências similares ao componente C3	93

LISTA DE TABELAS

TABELA 1 – Análise comparativa da sequência Lox-FB com outras sequências similares	60
TABELA 2 – Valores de identidade (%) com base no alinhamento para a par de diferentes clones Lox-C3.....	76
TABELA 3 – Análise comparativa da sequência Lox-C3 com outras sequências similares	92

SUMÁRIO

1 INTRODUÇÃO	21
1.1 Aranhas <i>Loxosceles</i>	21
1.1.1 <i>Biologia e distribuição geográfica do gênero Loxosceles (Heinecken e Lowe, 1835)</i>	21
1.1.2 <i>Loxoscelismo e veneno</i>	22
1.2 Sistema Complemento	25
1.2.1 <i>Vias de ativação</i>	25
1.2.2 <i>Aspectos evolutivos do Sistema Complemento</i>	28
1.2.3 <i>Componentes FB/C2</i>	31
1.2.4 <i>Componentes C3/C4/C5</i>	33
1.2.5 <i>Sistema Complemento em aranhas Loxosceles</i>	35
2 OBJETIVOS	36
3 MATERIAL E MÉTODOS	37
3.1 <i>Extração de RNA total da glândula de veneno das aranhas L. laeta</i>	37
3.2 <i>Síntese de cDNA a partir do RNA total extraído das glândulas de veneno das aranhas L. laeta</i>	37
3.3 <i>Amplificação e clonagem das sequências C3-like de L. laeta</i>	37
3.4 <i>5' e 3' RACE (Rapid Amplification of cDNA ends)</i>	39
3.5 <i>Amplificação e clonagem das sequências terminais 3' e 5' de FB- e C3-like de L. laeta</i>	39
3.6 <i>Análise e clonagem dos produtos C3-like e FB-like de Loxosceles laeta</i>	40
3.7 <i>Preparo de células competentes de Escherichia coli</i>	40
3.8 <i>Transformação de E. coli XL1blue por choque térmico</i>	41
3.9 <i>Extração, purificação e sequenciamento do DNA plasmidial</i>	42
3.10 <i>Análise das sequências por bioinformática e modelagem molecular</i>	42
3.11 <i>Análise filogenética</i>	43
4 RESULTADOS	44
4.1 <i>Clonagem do fator B de Loxosceles laeta (Lox-FB)</i>	44
4.1.1 <i>Amplificação do cDNA 5' de Lox-FB</i>	44
4.1.2 <i>Amplificação do cDNA 3' de Lox-FB</i>	47
4.2 <i>Caracterização do fator B-like de L. laeta</i>	49
4.3 <i>Alinhamento múltiplo de Lox-FB com outras proteínas FB/C2</i>	53

4.4 Análises de similaridade e filogenia de Lox-FB	59
4.5 Clonagem do componente C3-like de <i>L. laeta</i>	62
4.5.1 Amplificação da região que contém o motivo tioéster da sequência Lox-C3.....	62
4.5.2 Amplificação da região que contém o domínio ANATO da sequência Lox-C3.	67
4.5.3 Amplificação da região 5' do cDNA de Lox-C3.....	71
4.5.4 Amplificação da região 3' do cDNA de Lox-C3.....	73
4.6 Caracterização do C3-like de <i>L. laeta</i>	78
4.7 Alinhamento múltiplo de Lox-C3 com outras proteínas C3/C4/C5	84
4.8 Análises de similaridade e filogenia da sequência Lox-C3	91
5 DISCUSSÃO.	94
6 CONCLUSÕES.	102
REFERÊNCIAS	103

1 INTRODUÇÃO

1.1 Aranhas *Loxosceles* sp

1.1.1 Biologia e distribuição geográfica do gênero *Loxosceles* (Heinecken e Lowe, 1835)

Aranhas do gênero *Loxosceles* (Aranae, Sicariidae) são animais de pequeno porte; medem de 8 a 15 mm de corpo e suas pernas de 8 a 30 mm. Apresentam coloração que varia do marrom claro (*L. laeta*) ao escuro (*L. gaucho*) sendo, por isso, vulgarmente conhecidas como aranhas marrons (**FIGURA 1A**). Uma estrutura característica das aranhas desse gênero é o cefalotórax que se assemelha a um violino e possuem seis olhos arranjados em três pares (**FIGURA 1B**). Constroem teias irregulares que lembram chumaços de algodão, possuindo hábitos sedentários e noturnos (FUTRELL, 1992; GONÇALVES-DE-ANDRADE et al., 1999). Aranhas *Loxosceles* não são agressivas ao homem; os acidentes ocorrem apenas quando a aranha é comprimida contra o corpo. Seu habitat natural são as frestas e cascas soltas de árvores, folhas secas e cavidades do solo e de pedras, tendo se adaptado, no entanto, aos domicílios, onde se abrigam atrás de móveis, nos sapatos e em roupas penduradas (BÜCHERL, 1969).

FIGURA 1 - Aranhas *Loxosceles*

A

B

Loxosceles laeta coletadas no município do Rio de Janeiro. Fonte: Silveira, Revista da Sociedade Brasileira de Medicina Tropical, 2009 [A]. Visão anterior do cefalotórax de *Loxosceles reclusa* mostrando o padrão de seis olhos arranjados em três pares. Fonte: Swanson e Vetter, Clinics in Dermatology, 2006 [B].

O habitat e o substrato ocupado por essas aranhas favorecem sua dispersão passiva, resultando em transporte de algumas espécies para outras regiões do mundo (FISCHER; VASCONCELOS NETO, 2005). O sucesso na ocupação de novos ambientes é favorecido por fatores como alta longevidade e resistência a longos períodos sem alimentação e água (LOWRIE, 1980).

Essas aranhas possuem uma ampla distribuição mundial em regiões temperadas e tropicais, com diferentes espécies registradas na América do Sul, América Central, América do Norte, Europa, África, Oriente Médio, Oceania e Ásia (FUTRELL, 1992). Na América do Norte, *L. reclusa* e *L. deserta* são endêmicas e responsáveis pela maioria dos acidentes documentados. Dentre as espécies da América do Sul, *L. laeta* é, possivelmente, a que possui o veneno mais tóxico, possuindo uma ampla distribuição geográfica (FUTRELL, 1992). É endêmica no Chile e Peru, mas também tem sido encontrada na Austrália, Finlândia, Rússia, Mediterrâneo e nos Estados Unidos (HOGAN; BARBARO; WINKEL, 2004). Além disso, *L. laeta* apresenta um grande potencial reprodutivo: cada fêmea pode produzir 15 ootecas durante toda a sua vida (GONÇALVES-DE-ANDRADE; LOURENÇO; TAMBOURGI, 2000) e cada uma delas pode conter até 138 ovos (GALIANO; HALL, 1973).

No Brasil, esta espécie ocupa ecótopos urbanos e sua presença já foi registrada desde o estado da Paraíba até o Rio Grande do Sul e, foram também encontradas na cidade de São Paulo (GONÇALVES-DE-ANDRADE; TAMBOURGI, 2003).

1.1.2 *Loxoscelismo e veneno*

Pelo menos três diferentes espécies de *Loxosceles*, de importância para a saúde pública, são conhecidas no Brasil (*L. intermedia*, *L. gaucha*, *L. laeta*) e mais de 7000 casos de envenenamento por estas aranhas foram registrados no ano de 2014 (SINAN - Ministério da Saúde). O estado do Paraná, em especial o município de Curitiba, apresenta uma situação atípica em termos mundiais, quanto ao número de acidentes loxoscélicos (MARQUES-DASILVA; FISCHER, 2005) e com relação ao tamanho da população dessas aranhas.

O envenenamento em humanos pode resultar em duas variantes clínicas: o loxoscelismo cutâneo e o loxoscelismo viscero-cutâneo. O primeiro é mais comum e é caracterizado por uma lesão eritematosa acompanhada de queimação, que se desenvolve no

intervalo de 2 a 8 horas (FUTRELL, 1992). Subsequentemente, uma vesícula com conteúdo soro-hemorrágico aparece e evolui para uma extensa área de dermonecrose (SEZERINO et al., 1998).

O loxoscelismo viscero-cutâneo ou sistêmico é a forma mais grave do envenenamento e, além da lesão necrótica, os indivíduos podem apresentar febre, icterícia, trombocitopenia, coagulação intravascular, hemólise e falência renal (FUTRELL, 1992). Embora seja menos comum, o loxoscelismo sistêmico é a principal causa de morte associada ao envenenamento por *Loxosceles*. Muitos dos casos graves, relatados em crianças, na América do Sul, estão relacionados às espécies *L. laeta* (FUTRELL, 1992).

O grau de severidade do envenenamento parece estar relacionado a diversos fatores associados, por um lado, às aranhas, como as variações intra- e interespecíficas dos venenos (BÁRBARO et al., 1996; DE OLIVEIRA et al., 2005; PRETEL et al., 2005; TAMBOURGI et al., 1995), ao estágio do desenvolvimento (GONÇALVES-DE-ANDRADE et al., 1999), ao sexo (DE OLIVEIRA et al., 1999, 2005) e a quantidade de veneno injetada (BERNSTEIN; EHRlich, 1986), quanto por características do próprio paciente, como idade, constituição genética (BARRETO et al., 1990; NANCE, 1961; TAMBOURGI et al., 1998) e região do corpo aonde ocorreu a picada (FUTRELL, 1992).

Várias intervenções têm sido propostas e testadas para o tratamento do loxoscelismo, no entanto, não existe uma terapia definitiva e ainda há muitas divergências quanto à eficácia das drogas, do antiveneno ou de suas combinações (PAULI et al., 2006). Considerando todas as intervenções avaliadas, o antiveneno específico foi o mais indicado devido ao seu grande potencial quando administrado nas primeiras horas após o envenenamento, durante a fase aguda (FUTRELL, 1992). Contudo, ainda existem divergências quanto à efetividade do soro em neutralizar efeitos locais (dermonecrose) e o período ideal para sua administração (HOGAN; BARBARO; WINKEL, 2004; PAULI et al., 2006).

Assim como os venenos de serpentes e escorpiões, os de aranhas são misturas complexas de substâncias biologicamente ativas e inativas. As atividades dermonecrotica e letal dos venenos de *Loxosceles* foram relacionadas a componentes com peso molecular entre 32 e 35 kDa, com atividade esfingomielinásica e, por isso, denominados como esfingomielinases D (SMases D). As SMases D são capazes de hidrolisar esfingomielinina e lisofosfatidilcolina, induzir dermonecrose em coelhos e tornar eritrócitos humanos suscetíveis à lise por complemento (TAMBOURGI et al., 1995, 1998, 2000, 2002, 2005; VAN

MEETEREN et al., 2004). Foi proposto que a suscetibilidade à ação lítica do complemento autólogo esteja relacionada à atividade esfingomielinásica e à ativação de metaloproteases endógenas, responsáveis pela clivagem das glicoforinas presentes na membrana dos eritrócitos, com consequente remoção de ácido siálico, permitindo a deposição de componentes do complemento, principalmente, da via alternativa (TAMBOURGI et al., 1995, 1998, 2000).

Chonn, Cullis e Devine (1991), demonstraram que a carga dos fosfolipídios expostos na superfície celular pode direcionar a ativação do complemento para uma determinada via. A via alternativa é ativada por fosfolipídios carregados positivamente, enquanto a via clássica é ativada por fosfolipídios aniônicos como a fosfatidilserina. Considerando estes dados, Tambourgi et al. (2002, 2007), demonstraram que em eritrócitos humanos tratados com as SMases de *L. intermedia* havia perda da assimetria da membrana, com consequente exposição de fosfatidilserina, permitindo assim a ligação de C1q e ativação da via clássica do complemento. Estas mesmas atividades foram também, posteriormente, confirmadas com SMases D isoladas da glândula de veneno de *L. laeta* (FERNANDES-PEDROSA et al., 2002; SANTI-FERRARA et al., 2009). O efeito da SMase D tem sido bem caracterizado em mamíferos, mas como não tem sido relatado qualquer registro que envolva a alimentação de vertebrados por essas aranhas da Família Sicariidae, acredita-se que esta toxina tenha evoluído primariamente para auxiliar na imobilização das presas (artrópodes) ou na sua pré-digestão e que as funções dermonecróticas ocasionadas em humanos sejam um resultado evolutivo accidental. Partindo desta hipótese, Zobel-Thropp, Kerins e Binford (2012), testaram o efeito da toxina SicTox (SMase D assim denominada pelo grupo) proveniente da aranha *L. arizonica* após injeção em grilos. A resposta dos grilos à injeção de doses parálíticas da toxina putificada foi quase imediata, levando à extensão das pernas dianteiras, elevação das pernas acima e para longe do corpo. Os animais continuavam paralisados na mesma posição com a observação de espasmos ocasionais até morrerem dentro de 24 horas.

1.2 Sistema Complemento

1.2.1 Vias de ativação

Durante a evolução, dois sistemas de imunidade emergiram: a imunidade inata ou natural e a imunidade adquirida. O sistema imune inato é filogeneticamente mais antigo e é encontrado em todos os organismos multicelulares, enquanto a resposta adquirida, que surgiu há cerca de 450 milhões de anos, é encontrada em todos os vertebrados, com exceção dos Agnatha (NONAKA; KIMURA, 2006; ZARKADIS; MASTELLOS; LAMBRIS, 2001). O sistema complemento em vertebrados superiores desempenha um importante papel efetor tanto na imunidade inata quanto na adquirida, além de participar de vários processos imunorregulatórios (ZARKADIS; MASTELLOS; LAMBRIS, 2001).

O sistema complemento de humanos é formado por mais de 30 proteínas que podem ser encontradas no plasma ou ligadas à superfície de células. Este sistema pode reconhecer patógenos e células tumorais, ser ativado e gerar moléculas que participam da resposta inflamatória, da opsonização de partículas, para posterior eliminação por meio da lise celular ou fagocitose e que podem, ainda, estimular linfócitos B e T (CAROLL, 2004; GROS; MILDER; JANSSEN, 2008). A ativação do complemento pode ocorrer por três vias, ou seja, clássica, alternativa e das lectinas (**FIGURA 2**).

A ativação da via clássica pode ser iniciada pela ligação direta da subunidade C1q do componente C1 à superfície do patógeno. C1q é composto por seis unidades de três cadeias polipeptídicas (A, B e C) e, cada um dos trímeros, possui uma região CLR (*colagen-like region*) na região N-terminal e um domínio gC1q (domínio globular) na região C-terminal. Alguns ligantes, como as moléculas IgM e IgG, complexadas ao antígeno, interagem com o domínio globular de C1q e esse evento representa a ligação-chave entre os mecanismos da imunidade inata e adquirida. Outros ligantes, incluindo a proteína C-reativa (CRP), uma proteína de fase aguda do plasma humano e que se liga a polissacarídeos bacterianos (VOLANAKIS; KAPLAN, 1974), lipoproteínas de baixa densidade oxidadas (oxLDL) e células apoptóticas, interagem com a região CLR de C1q (KISHORE et al., 2004). Roumenina et al. (2008) demonstraram a interação do domínio globular do C1q humano com o lipopolissacarídeo de *Salmonella typhimurium*. A interação de C1q com diferentes ligantes gera uma mudança conformacional nas demais subunidades de C1 (C1r:C1s)₂, levando a ativação da ação autocatalítica de C1r e

este ativado pode agir hidroliticamente sobre C1s. C1s apresenta um sítio catalítico de serinoprotease que o capacita a atuar sobre os componentes C2 e C4, clivando-os, resultando na formação da C3 convertase C4b2a (ARLAUD et al., 2002).

A via das lectinas é ativada pelo reconhecimento de resíduos de carboidratos, presentes em microrganismos, pelas lectinas ligantes de manose (MBL) e ficolinas, que se encontram associadas às serinoproteases, denominadas MASPs (MASP-1, MASP-2 e MASP-3). As MBL e ficolinas apresentam uma estrutura semelhante à de C1q, apresentando os mesmos domínios colágeno-*like* e globular, o que indica origem evolutiva comum (FUJITA; ENDO; NONAKA, 2004). Quando a MBL reconhece motivos de carboidratos presentes em certos microrganismos, ocorre uma ativação das serino proteases associadas à MBL: MASP-1 e MASP-2, sendo que esta última, assim como o componente C1s da via clássica, cliva os componentes C2 e C4, formando a C3 convertase C4b2a. MASP-1 não é capaz de clivar C4 diretamente, somente C2 e C3, sendo que para este último, a hidrólise ocorre em uma razão muito baixa, tornando questionável seu papel essencial em ativar o sistema complemento. No entanto, experimentos utilizando inibidores específicos para MASP-1 demonstraram a não clivagem de C4 mediada por MASP-2 e que a ativação das vias das lectinas era completa e permanentemente bloqueada; o mesmo ocorria quando foram usados inibidores específicos para MASP-2, revelando o papel essencial de ambas serino proteases (DOBÓ et al., 2014).

A via alternativa pode ser ativada lentamente por hidrólise espontânea da ligação tioéster, presente nas moléculas de C3 (mecanismo de *tickover*), resultando em um componente C3 modificado (C3(H₂O)) (BEXBORN et al., 2008) que, por sua vez, pode se ligar ao fator B (componente homólogo ao componente C2 da via clássica e das lectinas), formando o complexo C3(H₂O)B. Nesta condição, o fator B pode ser proteoliticamente ativado por uma serinoprotease, o fator D, originando fragmentos Ba e Bb. O fragmento Ba se dissocia do complexo, enquanto Bb permanece ligado ao C3(H₂O)B, formando a C3 convertase iniciadora da via alternativa C3(H₂O)Bb. O fragmento C3b, assim como C3(H₂O), pode se associar ao fator B e, posteriormente, ser clivado pelo fator D gerando a C3 convertase C3bBb (HOURCADE, 2006). O complexo C3bBb é relativamente instável com uma meia vida de 90 segundos sob condições fisiológicas, mas quando este se associa à properdina (P), sua estabilidade na superfície de patógenos e de imunocomplexos aumenta em pelo menos 10 vezes (FEARON; AUSTEN, 1975).

A ativação de cada uma dessas três vias converge para a formação da C3 convertase que cliva C3, gerando o fragmento C3a, um pequeno peptídeo também conhecido como anafilatoxina, capaz de recrutar células inflamatórias para os sítios de infecção. A clivagem de C3 gera um outro fragmento maior, o C3b, que se liga covalentemente à superfície de patógenos por meio da região tioéster e, dessa forma, opsoniza o microrganismo, promovendo a fagocitose dependente de receptores para C3b presentes em leucócitos (LAMBRIS, 1988; RICKLIN et al., 2010).

Além de opsonizar microrganismos, o C3b pode interagir com muitas outras proteínas como os reguladores do complemento, o que leva à dissociação das convertases, facilitando a sua clivagem pelo fator I. As primeiras duas clivagens geram iC3b, que ao contrário de C3b não se liga ao fator B e não forma convertases, mas pode mediar a fagocitose por meio da ligação a CR1, CR3 e CR4, receptores presentes nos leucócitos. A terceira clivagem libera o fragmento C3c, deixando o fragmento C3dg ligado à superfície. C3dg não induz fagocitose, mas, assim como iC3b, é capaz de estimular células B por meio dos receptores CR2 (GROS; MILDRE; JANSSEN, 2008). O fragmento C3dg pode, ainda, ser clivado pela tripsina (LACHMANN; MÜLLER-EBERHARD, 1968), gerando C3d e C3g. O C3d pode se ligar aos receptores CR2, presentes em linfócitos B e esta interação funciona como um sinal co-estimulatório (*co-signal 1*) para ativação destas células em conjunção com a interação do antígeno ao receptor BCR (DEMPSEY et al., 1996), apresentando, portanto, uma propriedade adjuvante.

Quando C3b se liga à C3 convertase forma outra serinoprotease, a C5 convertase (C4b2a3b, na via clássica e das lectinas, e C3bBb3b, na via alternativa) que cliva moléculas de C5, gerando C5a. C5a é uma potente anafilatoxina, responsável por muitas das reações inflamatórias induzidas pela ativação inicial, por meio da interação com receptores para C5a (C5aR) presentes em células, como por exemplo, endoteliais e mastócitos (HARBOE; MOLLNES, 2008). Há formação também de um fragmento maior, C5b, que se liga aos componentes finais da cascata do complemento (C6-9) formando o complexo de ataque à membrana (MAC), C5b-9, responsável pela lise celular (WALPORT, 2001).

FIGURA 2 - Esquema das três vias de ativação do sistema complemento

Fonte: Dzik, *Acta Biochimica Polonica*, 2010

1.2.2 Aspectos evolutivos do Sistema Complemento

O sistema complemento tem sido extensivamente estudado do ponto de vista evolutivo. Há mais de 30 anos, acreditava-se que o complemento estava presente somente em vertebrados, uma vez que todas as tentativas de identificar elementos deste sistema em invertebrados haviam falhado (NONAKA; KIMURA, 2006). Com o surgimento de técnicas moleculares mais sofisticadas, incluindo-se as análises genômicas e transcriptômicas, foi possível a identificação de genes homólogos a C3 e fator B de mamíferos em invertebrados deuterostômios como o SpC3 (AL SHARIF et al., 1998) e SpFB (SMITH; SHIH; DACHENHAUSEN, 1998) no ouriço-do-mar *Strongylocentrotus purpuratus* e CiC3 (MARINO et al., 2002) e CiFB (YOSHIZAKI et al., 2005) na ascidia *Ciona intestinalis*.

Estudos com *Drosophila melanogaster* (ADAMS et al., 2000) e *Caenorhabditis elegans* (The *C. elegans* sequencing Consortium, 1998) revelaram a ausência desses genes nesses

organismos, sugerindo que o sistema complemento tenha sido estabelecido a partir dos deuterostômios. Contudo, com o isolamento de homólogos funcionais dos componentes C3 e fator B de vertebrados nos límulus *Carcinoscorpius rotundicauda* (CrC3 e CrC2/FB) (ZHU et al., 2005) e *Tachypleus tridentatus* (TtC3) (ARIK et al., 2008), na aranha *Hasarius adansoni* (Had-C3-1 e Had-C3-2) (SEKIGUCHI; FUJITO; NONAKA, 2011), no molusco da Classe Cephalopoda *Euprymna scolopes* (Es-C3) (CASTILLO; GOODSON; MCFALL-NGAI, 2009), no bivalve *Ruditapes decussatus* (Rd-C3 e Rd-FB) (PRADO-ALVAREZ et al., 2009), no coral *Swiftia exserta* (SeC3) (DISHAW; SMITH; BIGGER, 2005) foi possível demonstrar que os componentes centrais do complemento já existiam nos protostômios primitivos, antes mesmo da divergência das linhagens dos protostômios e deuterostômios. Portanto, a ausência de genes relacionados ao sistema complemento em *Drosophila melanogaster* e *C. elegans* pode ser devida a uma perda secundária (NONAKA; KIMURA, 2006).

Como sugerido por Nonaka e Kimura (2006), o sistema complemento teria se originado a partir dos componentes C3 e fator B, os únicos encontrados em cnidários e em alguns protostômios. Estes componentes permitiriam o funcionamento de uma via, semelhante à alternativa dos mamíferos, que se originou a partir de um ancestral comum de Cnidaria e Bilateralia, há mais de um bilhão e 300 milhões de anos. Enquanto os genes C3 e FB foram mantidos nos deuterostômios, estes foram perdidos, muitas vezes e independentemente, na linhagem dos protostômios. Como os genes da MASP, MBL e ficolinas não tinham sido identificados em protostômios e nos equinodermos, foi sugerido que estes componentes teriam sido recrutados para o sistema complemento após a emergência dos Cordados, cerca de 900 milhões de anos atrás. No entanto, Kimura, Sakaguchi e Nonaka (2009), posteriormente, conseguiram identificar genes *MASP-like* na anêmona-do-mar *Nematostella vectensis*, além dos genes que codificam para C3 e FB. Esta descoberta indica que a origem da via das lectinas do sistema complemento possa ser traçada antes da emergência dos cordados, possivelmente antes da divergência dos Cnidaria e Bilateralia na era Pré-Cambriana. Os ágnatos, que são animais desprovidos de mandíbula e os mais primitivos entre os organismos pertencentes ao Subfilo Vertebrata, parecem ter desenvolvido somente as vias alternativa e das lectinas do sistema complemento, provavelmente devido à ausência de imunoglobulinas nesses animais (SUNYER; LAMBRIS, 1998).

Finalmente, os eventos de duplicação gênica que aconteceram entre C3/C4/C5, entre FB/C2 e MASPs/C1r/s, antes da emergência dos peixes cartilaginosos, cerca de 600 milhões de anos atrás, contribuíram para o estabelecimento da terceira via de ativação, a clássica. Esta última via parece ter se estabelecido, simultaneamente, com o aparecimento do sistema imune adquirido, baseado em linfócitos e MHC (NONAKA; KIMURA, 2006). A identificação de proteínas relacionadas ao complemento em organismos filogeneticamente distintos está ilustrada na **FIGURA 3**.

FIGURA 3 - Perspectiva evolutiva para as moléculas de reconhecimento e de serino proteases envolvidas na ativação da cascata do sistema complemento

Circulos abertos e fechados representam a ausência e a presença do componente nas espécies correspondentes, respectivamente.

1.2.3 Componentes FB/C2

O componente C2 participa da formação das C3 convertases das vias clássica e das lectinas, enquanto o fator B é uma serinoprotease equivalente, atuante na via alternativa. Fator B e C2 são muito similares e, possivelmente, divergiram na linhagem dos vertebrados mandibulados sendo, frequentemente, denominados como C2/FB. O ancestral de C2/FB desempenhava, provavelmente, funções tanto de C2 como de FB de mamíferos (ZARKADIS; MASTELLOS; LAMBRIS, 2001).

C2 e fator B são serinoproteases codificadas por *loci* presentes em uma região próxima ao complexo principal de histocompatibilidade, no cromossomo seis em humanos. Estas duas proteínas compartilham cerca de 40% de similaridade. Seus sítios catalíticos, C2a (509 resíduos) e Bb (505 resíduos), apresentam similaridade com as porções catalíticas de outras serinoproteases. Os sítios não catalíticos, C2b (223 resíduos) e Ba (234 resíduos), contêm três motivos de repetição de aproximadamente 60 aminoácidos cada (BENTLEY; CAMPBELL, 1986).

O fator B, em humanos, é uma molécula única contendo cinco domínios: três CCPs (proteínas de controle do complemento) na região N-terminal, que estão conectados ao domínio VWA (fator de *von Willebrand* do tipo A) por um resíduo *linker* (α L) e um domínio serinoprotease na região C-terminal (SP) que apresenta o sítio catalítico (**FIGURA 4**). Os domínios VWA e SP formam o fragmento Bb e os domínios CCP1-3 e o *linker* formam o fragmento Ba. A ligação do fator B ao C3b depende de elementos estruturais do fragmento Ba e de um sítio de adesão dependente do íon Mg^{+2} (MIDAS), encontrado no domínio VWA (MILDER et al., 2007).

Os domínios CCP em componentes do complemento estão envolvidos na interação entre proteínas. As análises da estrutura tridimensional e por microscopia eletrônica dos CCPs indicaram a existência de pequenos domínios globulares e que os “bolsões de ligação” são formados entre dois CCPs. Dentre todos os CCPs presentes no FB humano, o terceiro contém elementos essenciais que proporcionam a interação com o fragmento C3b e o primeiro, aparentemente, é o menos imprescindível, de acordo com as análises de mutação sítio dirigida para esta região (HOURCADE; WAGNER; OGLESBY, 1995).

Genes que codificam para proteínas *FB-like* têm sido caracterizados em diferentes organismos que pertencem à linhagem anterior a dos cordados. Sequências *FB-like* do límulo

C. rotundicauda (CrC2/FB) (ZHU et al., 2005) e do ouriço-do-mar *S. purpuratus* (SpFB) (SMITH et al., 2001) apresentam algumas particularidades, tais como a presença de 5 domínios CCP, ao invés dos 3 domínios encontrados na sequência do FB humano, sugerindo que a presença de domínios CCP adicionais seja uma característica primitiva do gene C2/FB. Em contraste, Kimura et al. (2009) caracterizaram duas isoformas para os genes *FB-like* da anêmona-do-mar *N. vectensis* (Nv-FB-1 e Nv-FB-2), sendo que a primeira é constituída por 3 domínios CCP, enquanto a segunda apresenta dois domínios CCP adicionais como no *FB-like* encontrados em *C. rotundicauda* e *S. purpuratus*. Portanto, o Nv-FB-1 é o primeiro gene *FB-like* identificado em invertebrados que apresenta a mesma composição de domínios presente no FB de vertebrados. Recentemente, algumas sequências *FB-like* dos organismos que pertencem ao Filo Arthropoda foram identificados e todos eles apresentam mais de uma isoforma que variam entre dois a sete CCPs (SEKIGUCHI; NONAKA, 2015). É possível que a presença de somente três domínios CCP, na porção N-terminal, seja um fator ancestral e que esta região, no fator B de invertebrados, tenha sofrido duplicação gênica ou deleção. De qualquer forma, parece provável que genes com os mesmos domínios estruturais, encontrados em límulos e na anêmona-do-mar, sejam ortólogos dos genes FB e C2 dos mamíferos, o que sugere que estes tenham aparecido antes da divergência dos Cnidaria/Bilateria (NONAKA; KIMURA, 2006).

FIGURA 4 - Componente Fator B humano

Representação esquemática da estrutura tridimensional do fator B humano e sua sequência de domínios. Fonte: Milder et al., *Nature Structural & Molecular Biology* 2007.

1.2.4 Componentes C3/C4/C5

Os componentes C3 e C4 humanos possuem estruturas e funções similares. Ambos desempenham papéis análogos nas vias alternativa e clássica da cascata do complemento e são sintetizados como polipeptídeos compostos por uma única cadeia de aproximadamente 185 kDa. O processamento das pró-moléculas é similar e, neste, incluem a remoção do peptídeo sinal, glicosilação e clivagem proteolítica. Seus níveis séricos são mantidos primariamente pelo fígado, mas outros estudos identificaram sítios extrahepáticos relacionados à síntese de proteínas do complemento. Dentre estes sítios, estão incluídos os fagócitos mononucleares, fibroblastos, células mesangiais glomerulares, células mielóides, células epiteliais do pulmão, intestino e rim (MITCHELL et al., 1996).

A similaridade entre as estruturas primárias de C3, C4 e C5 é baixa, no entanto, é significativa quando comparada às do inibidor α -2 macroglobulina (A2M) e do CD109 (SOLOMON et al., 2004). Além disso, os componentes C3 e C4 compartilham uma estrutura única, *i.e.*, uma região tioéster intramolecular, ausente em C5, o que pode ser explicado pela perda secundária de bases que codificam para esta região (NONAKA; KIMURA, 2006). A superfamília que apresenta essa região é chamada de TEP (proteínas que contêm tioéster). A superfamília TEP é subdividida em duas famílias: (1) família C3 que compreende os componentes C3/C4/C5 de vários eumetazoários e a (2) família α -2 Macroglobulina (A2M), composta por α 2M, proteína de zona de gravidez (PZP), CD109, CPDAM8 de humanos e seus ortólogos como também iTEPs presentes em insetos (NONAKA, 2014; SEKIGUCHI et al., 2011).

A molécula madura do C3 humano (**FIGURA 5**) é composta por duas cadeias – α (resíduos 1-645) e β (resíduos 650-1641), que juntas formam 13 domínios, isto é, oito domínios macroglobulina (MG1-8), um domínio linker (LNK), um domínio anafilatoxina (ANATO), um domínio CUB, um domínio contendo tioéster (TED) e um domínio C345c. Os membros da família A2M apresentam a mesma estrutura básica de domínios encontrada no componente C3, no entanto, os domínios ANATO e C345c estão ausentes (BAXTER et al., 2007; JANSSEN et al., 2005).

No C3 nativo, o domínio ANATO estabiliza a interface TED-MG8, mas quando este é clivado e o domínio ANATO é removido, ocorrem mudanças conformacionais na estrutura da molécula. Com a remoção do domínio ANATO, há o deslocamento dos domínios MG7 e MG8,

o domínio α -NT assume uma nova posição e os domínios CUB e TED giram em direção à região exterior, expondo a ligação tioéster. Devido a ausência do domínio ANATO, as moléculas TEP1r dos insetos são mais estáveis e, portanto, estão mais estruturalmente relacionadas ao pro-C3 (GROS; MILDER; JANSSEN, 2008).

O componente denominado *Cobra Venom Factor* (CVF), isolado do veneno da serpente *Naja naja kouthia*, é um análogo funcional do C3b humano. Similarmente ao C3b dos mamíferos, o CVF se liga ao fator B que é, então, clivado pelo fator D, resultando no complexo CVFBb, uma eficiente C3 convertase e C5 convertase. Diferente das convertases C3bBb e C4bC2a, CVFBb é um complexo mais estável, já que sua estrutura, um intermediário entre C3b e C3, não é constituída pelo domínio TED e apresenta um domínio CUB numa posição idêntica àquela identificada no C3b (KRISHNAN et al., 2009).

FIGURA 5 - Componente C3 Humano

Representação esquemática da estrutura tridimensional do C3 humano. Sequências de domínios de C3 e seus respectivos produtos de clivagem após ativação do sistema complemento. Fonte: Janssen et al., Nature 2005.

Com relação aos genes *C3-like*, que foram recentemente caracterizados em organismos da linhagem dos protostômios como, por exemplo, CrC3 (*C. rotundicauda*), Rd-C3 (*R. decussatus*) e Haad-C3 (*Hasarius adansonii*), e comparando suas sequências de aminoácidos com as do C3 humano, foram encontradas, pelo menos, 4 regiões conservadas: a região de ligação da cadeia α e β , as pontes de dissulfeto formadas entre as cadeias α e β , o resíduo catalítico caracterizado pela presença do aminoácido histidina e, finalmente, o sítio tioéster (NONAKA, 2011; PRADO-ALVAREZ et al., 2009; SEKIGUCHI; FUJITO; NONAKA, 2011; ZHU et al., 2005). Portanto, apesar de algumas diferenças, a presença de domínios conservados sugere que elementos estruturais de C3, essenciais à sua ação, foram mantidos ao longo da evolução. Além disso, foi detectado um domínio de junção das cadeias α e γ composto por quatro aminoácidos RKKR (*R. decussatus*) (PRADO-ALVAREZ et al., 2009), RQKR (*H. adansonii*) (SEKIGUCHI; FUJITO; NONAKA, 2011), RVKR (*C. rotundicauda*) (ZHU et al., 2005) que se alinham próximo à região do motivo RRRR presente no C4 humano, sugerindo que a estrutura das proteínas *C3-like*, de organismos da linhagem dos protostômios, seja composta por 3 cadeias.

1.2.5 Sistema complemento em aranhas *Loxosceles*

Um dos métodos utilizados para o estudo da diversidade molecular e da expressão gênica de uma célula ou tecido específico é a caracterização do transcriptoma, por meio da determinação randômica de sequências expressas (ESTs – *expressed sequence tags*). A análise do transcriptoma da glândula de veneno de *L. laeta* (Araneae, Sicariidae) possibilitou a identificação de sequências similares às de C3 (LLAE0374S – GenBank: EY188746.1) de *Carcinoscorpius rotundicauda* e fator B (LLAE0899S – GenBank: EY189271.1) de *Strongylocentrotus purpuratus* (FERNANDES-PEDROSA et al., 2008). Tal achado abriu portas para que novos estudos fossem conduzidos, com objetivo de caracterizar as sequências gênicas e de aminoácidos desses componentes, o que certamente poderá contribuir para uma melhor compreensão da evolução do complemento e do papel desse sistema na glândula de veneno de *Loxosceles*.

2 OBJETIVOS

Os objetivos do presente projeto são:

1. Identificar os genes que codificam para os componentes FB e C3 presentes nas glândulas de veneno das aranhas *L. laeta*;
2. Determinar as sequências deduzidas de aminoácidos;
3. Analisar a estrutura e determinar os possíveis domínios funcionais;
4. Avaliar do ponto de vista evolutivo as sequências.

3 MATERIAL E MÉTODOS

3.1 Extração de RNA total das glândulas de veneno das aranhas *L. laeta*

Aranhas *L. laeta*, oriundas da cidade de Campo Alegre/SC, foram criadas e mantidas pelo “Biotério de Criação e Manutenção de Aranhas *Loxosceles*” do Laboratório de Imunoquímica do Instituto Butantan. Oitenta aranhas fêmeas de *L. laeta* foram submetidas à restrição alimentar para estimular a produção de RNA mensageiro nas glândulas de veneno. Após cinco dias, as glândulas foram coletadas, sendo a extração do RNA total realizada com o uso do reagente Trizol, seguindo as instruções do fabricante (Gibco-BRL Life Technologies, MD, EUA).

3.2 Síntese de cDNA a partir do RNA total extraído das glândulas de veneno das aranhas *L. laeta*

Para a síntese dos cDNAs, foi empregada a reação de RT-PCR (Transcriptase Reversa) utilizando o *ProtoScript® M-MuLV First Strand cDNA Synthesis kit* (New England Biolabs, MA, EUA), seguindo as descrições do fabricante. Para a reação de RT-PCR foram utilizados 5 µL de RNA total, 2 µL de *Random Primer mix* (6 µM) ou oligo-dT₂₃ (50 µM) e 1 µL de água tratada com DEPC (dietilpirocarbonato), sendo as amostras previamente incubadas a 70 °C por 5 minutos. Após o período de incubação, as amostras foram deixadas no gelo durante um minuto e, em seguida, foram adicionados 10 µL de M-MuLV *reaction mix* e 2 µL M-MuLV *Enzyme mix*. As amostras foram, então, incubadas a 42 °C durante 1 hora e, após esse período, a reação de transcriptase reversa foi inativada a 80 °C por 5 minutos, sendo as amostras mantidas a -20 °C até a sua utilização.

3.3 Amplificação e clonagem das sequências C3-like de *L. laeta*

Para identificar a porção 5' do gene C3-like de *L. laeta*, foram utilizados dois *primers* degenerados desenhados a partir das regiões mais conservadas com base no alinhamento de proteínas C3-like representativas de animais invertebrados utilizando o programa BioEdit (HALL, 1999) (**FIGURA 6**). A seguir, foi feita uma reação de PCR com o objetivo de amplificar

a região que compreende o motivo tioéster, utilizando o *primer* C3 degenerado C3_TEP_2 *sense* - 5' AGYACNTGGCTNACNGCWTTY 3' (Integrated DNA Technologies, IA, EUA) - e o *primer* C3 *antisense* - 5' GTTCCGCACTCTTGACAAGC 3' (Invitrogen, CA, EUA). Para a reação de PCR, foram utilizados 2 µL de cDNA obtido da amplificação com *Random Primers*, 2 µL do tampão 10X *Standard Buffer* (New England Biolabs, MA, EUA), 0,5 µL do *primer* C3_TEP_2 (100 µM), 1 µL do *primer antisense* C3 (10 µM), 0,5 µL de DNTP mix (20 mM), e 0,12 µL de *Taq* (New England Biolabs, MA, EUA), completando o volume da reação para 20 µL com água Mili-Q. As amostras foram incubadas no termociclador *Veriti® 96-Well Fast Thermal Cycler* (Applied Biosystems, CA, EUA), utilizando as seguintes condições de amplificação: 95 °C por 5 minutos, 30 ciclos de 95 °C por 30 segundos, 60 °C por 30 segundos e 68 °C por 1 minuto, terminando a reação com uma extensão adicional a 68 °C por 7 minutos.

Para amplificar a região que compreende o domínio ANATO de C3-like de *L. laeta*, foi feita outra reação de PCR utilizando um segundo *primer* degenerado denominado C3_A2M_2 *sense* - 5' GAYYTNGGHTGYGGHCCWGGH 3' (Integrated DNA Technologies, IA, EUA) - e o *primer antisense* 1 C3_TEP_2 - 5' GAAGACGTGATGTGCAGTGGTGTG 3' (Invitrogen, CA, EUA). O tamanho esperado para a amplificação deste produto é de 1500 pares de bases e, para este fim, foi realizada uma reação de PCR, utilizando 3 µL de cDNA (amplificado com *Random Primers* ou com oligo-dT), 5 µL de 5X *Kappa Hi Fi Fidelity Buffer*, 0,5 µL de *primer* C3 *sense* degenerado C3_A2M_2 (100 µM), 1 µL de *primer antisense* 1 C3_TEP_2 (10 µM), 0,75 µL de DNTP mix KAPPA (10 mM) e 0,5 µL de enzima *KAPPA HiFi™* (KAPABIOSYSTEMS, MA, EUA). As amostras foram incubadas no termociclador *Veriti® 96-Well Fast Thermal Cycler* (Applied Biosystems, CA, EUA), utilizando as seguintes condições de amplificação: 95 °C por 5 minutos, 30 ciclos de 98 °C por 20 segundos, 63 °C por 30 segundos e 72 °C por 1 minuto e 30 segundos, terminando a reação com uma extensão adicional a 72 °C por 7 minutos. Após verificar a amplificação do produto esperado, foi feita uma reação 2X PCR, utilizando 5 µL do produto PCR anterior diluído [1: 100], 5 µL de 10X *Standard Buffer*, 0,5 µL de *primer* C3 *sense* degenerado C3_A2M_2 (100 µM), 1 µL de *primer antisense* 1 C3_TEP_2 (10 µM), 0,5 µL DNTP mix (20 mM) e 0,25 µL de *Taq* (New England Biolabs, MA, EUA), completando o volume da reação para 50 µL com água Mili-Q.

FIGURA 6 – Seleção de *primers* degenerados C3-like

Alinhamento múltiplo entre as sequências C3-like de invertebrados e seleção das regiões mais conservadas para desenho de *primers* degenerados. *Primer* degenerado 1 C3_TEP_2 sense e *primer* degenerado 2 C3_A2M_2 sense.

3.4 5' e 3' RACE (*Rapid Amplification of cDNA ends*)

Devido a dificuldade em desenhar *primers* degenerados para a região da extremidade 5' e 3' dos genes C3 e fator B de *L. laeta*, foi empregada a técnica de RACE-PCR que consiste na incorporação de um oligonucleotídeo específico na porção 5' ou 3' do RNA mensageiro e na reação direta de transcrição reversa realizada pela enzima *SMARTScribe Reverse transcriptase* sem necessidade de tratamento prévio do RNA. A reação foi realizada seguindo as orientações do fabricante (Clontech, CA, EUA), utilizando um volume de 2,75 µL de RNA total (338,3 ng/µL).

3.5 Amplificação e clonagem das sequências terminais 3' e 5' de FB-like e C3-like de *L. laeta*

Para amplificar a região 5' e 3' dos genes correspondentes ao fator B e C3 de *L. laeta*, foi utilizado um volume de 2,5 µL do cDNA resultante da reação RT-PCR, 5 µL de *Universal Primer Mix* (UPM) [10X], 34,5 µL de água Mili-Q, 5 µL de Tampão *Advantage 2 PCR* [10 x], 1 µL de DNTP mix [10 mM] e 1 µL da enzima *Advantage 2 Polymerase*. Para amplificar a porção 5'

do gene fator B, foi utilizado 1 µL de *primer* FB *antisense* [10 µM] - 5' CCTTCATCCATGCGACCAC 3' (Invitrogen, CA, EUA) e para a porção 3', foi utilizado 1 µL de *primer* FB *sense* [10 µM] - 5' CGAAGCAGCTCAAGGACCAC 3' (Invitrogen CA, EUA). Com relação à porção 5' do gene C3-like, foi utilizado 1 µL do *primer* ANATO *antisense* [10 µM] - 5' GTGCTCTGGGAATGAATCACGATCG 3' (Integrated DNA Technologies, IA, EUA) e para a porção 3', três *primers sense* localizados nas regiões que contêm os domínios ANATO - 5' GGTCTGTGGAGGTTTGACAAC 3' (*primer* anato *sense*) (Síntese Biotecnologia, MG, BR), tioéster 5' GAAGACGTGATGTGCAGTGGTGTG 3' (*primer* tioéster *sense*) (Invitrogen, CA, EUA) - 5' e o domínio A2M_receptor - 5' CGTCTTGCTTCAAGATGACC 3' (*primer* C3 *sense*) foram empregados. As amostras foram incubadas no termociclador Veriti® 96-Well Fast Thermal Cycler (Applied Biosystems, CA, EUA), utilizando as seguintes condições de amplificação: 5 ciclos de 94 °C por 30 segundos, 72 °C por 3 minutos, 5 ciclos de 94 °C por 30 segundos, 70 °C por 30 segundos e 72 °C por 3 minutos e 27 ciclos de 94 °C por 30 segundos, 68 °C por 30 segundos e 72 °C por 3 minutos.

3.6 Análise e clonagem dos produtos C3 e fator B de *L. laeta*

Para análise dos produtos de PCR, um volume de 5 µL de cada reação foi submetido à eletroforese em gel de agarose a 1% em tampão TAE [1 X], sendo as bandas visualizadas sob luz ultravioleta e coloração com gel red diluído [1: 500] (New England Biolabs, MA, EUA). Os produtos de PCR foram purificados, utilizando o *PureLink™ PCR Purification Kit* (Invitrogen, CA, EUA) e, posteriormente, clonados no vetor *pGEM-Teasy* seguindo as instruções do fabricante (Promega, WI, EUA).

3.7 Preparo de células competentes de *Escherichia coli*

Células competentes foram preparadas a partir de culturas da linhagem *XL1 blue* de *E. coli*. Para tanto, inoculou-se uma colônia isolada em amostras contendo 3 mL de meio SOB (bactotripton 2%, extrato de levedura 0,5%, NaCl 10 mM, KCl 2,5 mM, MgCl₂ 10 mM e MgSO₄ 10 mM). A amostra foi incubada com agitação a 37 °C por 18 horas. Em seguida, foram adicionados 100 µL da cultura anterior em 50 mL de meio SOB, seguida da incubação por mais 2 ou 3 horas até atingir D.O._{550 nm} = 0.5-0.6. A cultura foi transferida para tubos de

polipropileno (previamente gelados), aos quais foi adicionado 0,5 mL de MgCl_2 [1 M] e estes foram incubados no gelo por 15 minutos. Após a centrifugação a $1200 \times g$ por 12 minutos e a 4°C , os sobrenadantes foram completamente descartados, os precipitados ressuspensos em 10 mL de solução RFI pH 5,8 (KCl 10 mM, MnCl_2 50 mM, acetato de potássio 30 mM, CaCl_2 10 mM e glicerol 15%) e incubados, novamente, em gelo por 15 minutos. As suspensões foram centrifugadas e ressuspensas em 2 mL de solução RFII pH 6,8 (Na-MOPS 10 mM, KCl 10 mM, CaCl_2 75 mM e glicerol 15%) e, em seguida, distribuídas em alíquotas de 100 μL em tubos, congeladas em banho de gelo seco com etanol e estocadas a -80°C .

3.8 Transformação de *E. coli XL1blue* por choque térmico

Amostras de 5 μL , contendo os genes C3- e FB-*like* ligados aos plasmídeos *pGEM-Teasy* foram adicionadas aos tubos contendo 100 μL das suspensões de bactérias competentes *XL1 blue* e, em seguida, as amostras foram mantidas no gelo por 30 minutos. Para induzir a transformação das bactérias, as amostras foram incubadas em banho-maria a 42°C por 1 minuto e meio e, em seguida, colocado no gelo durante 3 minutos. A seguir, foi adicionado 1 mL de meio SOC (bactotripton 2%; extrato de levedura 0,5%; NaCl 10 mM; KCl 2,5 mM; MgCl_2 10 mM; MgSO_4 10 mM e glicose 20 mM) e as amostras foram mantidas sob agitação (200 rpm) por 1 hora a 37°C . As amostras foram, então, centrifugadas ($16000 \times g$) a 23°C por 2 minutos. Os sobrenadantes foram descartados e os precipitados bacterianos plaqueados em meio ágar LB (triptona 1%, NaCl 1%, extrato de levedura 0,5% e ágar 2%) contendo 100 $\mu\text{g/mL}$ de ampicilina. As placas, previamente incubadas com 40 μL de X-Gal [5-bromo-4-chloro-3-indolyl- β -galactoside] (20 mg/ml) e 4 μL de IPTG [*isopropylthio-B-D-galactoside*] (200 mg/ml), foram mantidas na estufa a 37°C durante a noite.

No dia seguinte, as bactérias transformadas (colônias brancas) foram selecionadas e inoculadas em meio LB líquido (triptona 1%; NaCl 1%; extrato de levedura 0,5%) contendo ampicilina (100 $\mu\text{g/mL}$), os quais foram submetidos à agitação (100 rpm) a 37°C durante a noite.

3.9 Extração, purificação e sequenciamento do DNA plasmidial

Das culturas frescas de bactérias, foi extraído o DNA plasmidial pelo método *Boil-mini prep* (HOLMES; QUIGLEY, 1981), sendo as amostras transferidas para tubos e centrifugadas por 2 minutos, a 16.000 x g e a 4 °C. Os sobrenadantes foram descartados, os precipitados ressuspendidos em 700 µL de solução STET (Sacarose 20%, Tris-HCl 1 M, EDTA 0,5M, Triton-X 100, RNase 10 mg/ml) e, em seguida, foi adicionado um volume de 25 µL de lisozima (10 mg/mL) em cada uma das amostras que foram, posteriormente, incubadas por 2 minutos a 95 °C. Após centrifugação por 8 minutos a 16.000 x g e a 4 °C, os precipitados (restos bacterianos e DNA cromossômico) foram descartados, enquanto aos sobrenadantes, que continham os DNAs plasmidiais, foram adicionados 80 µL de acetato de amônio (3 M) e 800 µL de isopropanol para precipitar o DNA. Após a etapa de lavagem do DNA, pela adição de 1 mL de etanol 70%, o precipitado foi seco em *speed-vac* por 10 minutos e ressuspendido em 50 µL de solução T.E. (Tris-HCl 10 mM, pH 8; EDTA 1 mM). As amostras de DNA plasmidial foram submetidas à digestão com a enzima de restrição EcoRI, conforme as instruções do fabricante (New England Biolabs, MA, EUA) e a análise do tamanho dos insertos, correspondentes às sequências C3- e FB-like, foi efetuada por meio da técnica de eletroforese em gel de agarose a 1%.

As amostras foram purificadas utilizando fenol e clorofórmio (AUSUBEL, 1995) e, posteriormente, submetidas ao sequenciamento de DNA, de acordo com o método de terminação de cadeia com dideoxinucleotídeos, utilizando-se o *Big Dye Terminator Cycle Sequencing Ready Reaction Kit* (Applied Biosystems, CA, EUA), *primers* T7 (5' TAATACGACTCACTATAGGG 3') e Sp6 (3' TAAATCCACTGTGATATCTT 5') específicos para o vetor, e o sequenciador de DNA, modelo ABI PRISMTM 3100 de eletroforese capilar (Applied Biosystems, CA, EUA). Para sequenciar as porções 5' e 3' do gene C3-like, *primers* específicos foram desenhados para determinar a sequência completa.

3.10 Análise das sequências por bioinformática e modelagem molecular

As sequências obtidas foram submetidas à análise de similaridade, utilizando o programa BLAST (*Basic Local Alignment Search Tool*) disponível, interativamente, no endereço <http://www.ncbi.nlm.nih.gov> (ALTSCHUL et al., 1997). A tradução e a análise das

sequências deduzidas de aminoácidos foram realizadas com o auxílio da ferramenta *TRANSLATE TOOL* (<http://www.expasy.ch>). Uma vez detectada a similaridade entre as amostras de interesse e outras sequências presentes no *Genbank*, foi feito um alinhamento de várias sequências para analisar a existência de sítios conservados, utilizando a ferramenta *MUSCLE (Multiple Sequence Comparison by Log-Expectation)* disponível no programa *MEGA 6 (Molecular Evolutionary Genetics Analysis)* (TAMURA et al., 2013). Com base na estrutura do fator B humano (código PDB 2ok5.1) e do C3 bovino (código PDB 2b39.2A), o programa *Swiss model* disponível no endereço eletrônico <http://www.expasy.ch>, foi utilizado para criar um modelo de homologia do fator B e C3 de *Loxosceles*, respectivamente (BORDOLI et al., 2009).

3.11 Análise filogenética

Todas as sequências *C3-like* e *FB-like* utilizadas para a análise filogenética foram obtidas do banco de dados do *GenBank* e submetidas a um alinhamento múltiplo utilizando a ferramenta *MUSCLE (Multiple Sequence Comparison by Log-Expectation)*. As árvores filogenéticas foram estimadas com o algoritmo Máxima Verossimilhança e a confiabilidade estatística das relações filogenéticas foi avaliada pelo método de *Bootstrap* de 1000 repetições, utilizando o programa *MEGA 6 (Molecular Evolutionary Genetics Analysis)* (TAMURA et al., 2013).

4 RESULTADOS

4.1 Clonagem do fator B de *Loxosceles laeta* (Lox-FB)

4.1.1 Amplificação do cDNA 5' de Lox-FB

A sequência EST correspondente ao FB-like (EY189271) da glândula de veneno de *L. laeta* foi submetida a uma análise para verificar com qual organismo era mais relacionada dentre as sequências depositadas no Banco de Dados de Genes (*GenBank*). O fator B/C2-like do anfioxo *Branchiostoma belcheri* foi o que apresentou maior identidade com a EST EY189271 (31%) e, portanto, sua sequência Bb FB/C2 (ABY28382.1) foi utilizada como base para as análises posteriores.

A EST FB-like de *L. laeta* apresenta uma sequência que codifica parte de um domínio reconhecido como VWA (fator de *von Willebrand* do tipo A), a qual se alinha aproximadamente na posição 377 a 532 da sequência de aminoácidos de Bb FB/C2. Sabendo que Bb FB/C2 é uma proteína com 752 aminoácidos e que o *primer antisense*, utilizado para amplificar FB-like de *L. laeta*, se alinharia próximo à posição 488 da sequência de aminoácidos, era esperado a amplificação de um produto de, aproximadamente, 1464 nucleotídeos a partir da reação 5' RACE-PCR (**FIGURA 7**).

Após a reação de PCR, na qual foram utilizados os *primers* FB *antisense* e o universal comercial UPM (Clontech, CA, USA), foi possível observar a amplificação de um produto de aproximadamente 1400 pb (**FIGURA 8A**). Este produto de PCR foi purificado, clonado no vetor *pGEM-Teasy*. Após a etapa de transformação, os cDNAs dos clones positivos foram extraídos e submetidos à digestão com a enzima EcoRI (**FIGURA 8B**) e, a partir da análise do gel de agarose 1%, foi observado que os 4 primeiros clones apresentavam o tamanho esperado, sendo o clone 2 selecionado para ser submetido ao sequenciamento.

FIGURA 7 - Análise da sequência EST Lox-FB**A**

Score: 93.6 bits (231), Expect: 6e-19,				
Identidade: 51/167(31%) Similaridade 88/167(59%) Gaps: 29/167(17%)				
EST Lox-FB	1	AMVPALTSVFLKVQPFRLERSKHALFLISDGEPNIGGDDITPEDISKQLKDHAAFEIFAV	60	
		A+ AL +V V P R +K ALF+I+DG+ N+G +P +++L++ AFE+FAV		
Bb C2/FB	377	ALTAALQTVLNDVLPVTRAGAKRALFIITDGKSNVGA---SPGVFARRLREEEAFEVFAV	433	
Primer FB antisense				
EST Lox-FB	61	GIGPDVQRKTLASVASEPVLHHVFM LDKFTDLEDMMAIMKNRPTEEPVSLDKCGRMDGR	120	
		G+G ++ + L SVAS+P HVF+++ F++ + ++ + E + ++CGR+		
Bb C2/FB	434	GVGANIDKNELNSVASQPFTSHVFLINDFSNFDTLVNTIA-----EKEIDYEQCGRVKQA	488	
EST Lox-FB	121	NESEG-----WPWVAAYVVSYEAXNSTQLSLCAG	149	
		++ G WPW+ A+Y E QL +C G		
Bb C2/FB	489	QQAAGQCGSSSRVVGTEAARGAWPWLVAIY--QERRQGVQL-VC GG	532	

B

Alinhamento de aminoácidos das sequências EST Lox-FB (EY189271) e FB/C2-like do anfióxico *Branchiostoma belcheri* (Bb FB/C2) [A] e arquitetura dos domínios de ambas as sequências com localização dos *primers* para amplificação da região 5' e 3' de Lox-FB [B] utilizando como base a ferramenta Blastp (<http://blast.ncbi.nlm.gov/Blast.cgi>)

FIGURA 8 - Amplificação e clonagem do produto Lox-FB resultante da reação de 5' RACE-PCR

Análise dos produtos amplificados após reação de PCR [A] e do tamanho dos fragmentos de cDNA, inseridos no vetor *pGEM-Teasy* após digestão com *EcoRI* [B], em géis de agarose a 1%.

4.1.2 Amplificação do cDNA 3' de Lox-FB

Com base na sequência do fator B/C2 do anfioxo *Branchiostoma belcheri* Bb Bf/C2, foi verificado que o *primer* FB *sense* utilizado para amplificar a porção 3' do gene FB-like de *L. laeta* se alinharia próximo à posição 419 da sequência de aminoácidos e que o tamanho esperado do produto resultante da reação 3' RACE-PCR seria de, aproximadamente, 1000 nucleotídeos (333 aminoácidos) (**FIGURA 7**).

Após a reação de PCR, na qual foram utilizados os *primers* FB *sense* e o universal comercial UPM, foi possível observar a amplificação de um produto em torno de 1300 pares de bases (pb) (**FIGURA 9A**). Este produto de PCR foi purificado, clonado ao vetor *pGEM-Teasy*. Após a etapa de transformação, os cDNAs dos clones positivos foram extraídos e submetidos à digestão com a enzima EcoRI (**FIGURA 9B**) e, a partir da análise do gel de agarose 1%, foi observado que os insertos dos clones 2-13 e 15-18 apresentavam o tamanho esperado, sendo que o clone 2 foi selecionado para ser submetido ao sequenciamento.

FIGURA 9 – Amplificação e clonagem do produto Lox-FB resultante da reação de 3' RACE-PCR

Análise dos produtos amplificados após reação de PCR [A] e do tamanho dos fragmentos de cDNA, inseridos no vetor *pGEM-Teasy* após digestão com *EcoRI* [B], em géis de agarose a 1%.

4.2 Caracterização do fator B-like de *L. laeta*

Os clones provenientes dos produtos FB 5' e FB 3' geraram um cDNA, com 1953 pares de bases, que codifica para uma proteína madura de 651 aminoácidos (**FIGURA 10**). O peptídeo sinal é composto por 25 aminoácidos e a proteína madura por 626 aminoácidos (**FIGURA 11A**). Dezoito cisteínas foram encontradas, oito destas localizadas nos domínios CCP, uma no domínio vWFA e nove cisteínas no domínio SP.

A massa molecular de Lox-FB, mensurada por métodos bioinformáticos foi de 72,38 kDa e o ponto isoelétrico de 5,73, não considerando os oito sítios putativos de glicosilação. Lox-FB apresenta uma organização de domínios clássica, ou seja, composta por dois domínios CCP, um domínio vWFA e um domínio serino protease (SP) (**FIGURA 11B**).

FIGURA 10 – Sequência de nucleotídeos e deduzida de aminoácidos de Lox-FB (Parte 1)

1	ATG	GCA	TCC	GTG	GAG	TGG	GCG	ATT	CTT	GGA	TTC	CTC	TTT	TGG	ATT	TCA	ACT	GTT	ATA	CGG	60
1	M	A	S	V	E	W	A	I	L	G	F	L	F	W	I	S	T	V	I	R	20
61	CAA	GCA	GGT	AGC	TCC	GTA	TGC	CGT	CCC	TTG	CCG	TAT	TCT	TCT	TTG	CGG	AAC	GGG	AAC	ATC	120
21	Q	A	G	S	S	V	C	R	P	L	P	Y	S	S	L	R	N	G	N	I	40
121	ATC	GCT	GAA	AGT	GGA	GGA	ACT	TTT	CGC	TTC	GAA	TGC	AAC	AAA	GGT	TAC	TTT	CTC	GGA	GGA	180
41	I	A	E	S	G	G	T	F	R	F	E	C	N	K	G	Y	F	L	G	G	60
181	CCC	TCA	AAA	ATA	CGT	TGC	TAT	AAG	GGA	AAC	TGG	ACG	TCG	GAA	AGA	GCA	CCC	GTG	TGC	ATC	240
61	P	S	K	I	R	C	Y	K	G	N	W	T	S	E	R	A	P	V	C	I	80
241	GAA	TCT	GGA	GAA	TTT	TGT	GAT	GCT	CCT	CAA	AAG	ATC	GCA	AAC	GCA	AAA	CTG	CAC	GGT	CCC	270
81	E	S	G	E	F	C	D	A	P	Q	K	I	A	N	A	K	L	H	G	P	90
271	ACC	GTG	TAT	GAA	ATT	GAT	TCT	GTC	ATA	CAG	TAC	ATG	TGC	CGT	CCA	GGA	TAT	GTG	CTG	GTG	360
91	T	V	Y	E	I	D	S	V	I	Q	Y	M	C	R	P	G	Y	V	L	V	120
361	GGC	AAC	GGA	CGG	AGG	ATA	TGC	CTC	AGG	AGT	GGC	TAT	TGG	AGC	GGT	ATC	ACT	CCC	ACA	TGC	420
121	G	N	G	R	R	I	C	L	R	S	G	Y	W	S	G	I	T	P	T	C	140
421	ATG	GAT	GAA	TCC	GAG	CCT	CTG	CAG	AAT	GTT	GCT	GAG	AGA	TTG	AAC	AAA	GAG	TTT	GTT	ACG	480
141	M	D	E	S	E	P	L	Q	N	V	A	E	R	L	N	K	E	F	V	T	160
481	AAA	ATG	GCA	TCG	CAC	TCA	AAT	GAC	AGT	TCA	ATA	CAA	CAG	GCT	TAT	TAT	CTT	GAC	AGT	GCT	540
161	K	M	A	S	H	S	N	D	S	S	I	Q	Q	A	Y	Y	L	D	S	A	180
541	TCT	AGG	CGA	AGA	GGT	CTC	GAT	CTG	GTT	ATT	CTG	GTT	GAC	AGA	TCC	AGC	AGC	GTT	GAT	CCT	600
181	S	R	R	R	G	L	D	L	V	I	L	V	D	R	S	S	S	V	D	P	200
601	TTT	GAC	TTG	AAC	GTT	GCC	AAG	AAA	TTC	ATA	AAA	TAT	TTG	CTT	CAA	GAA	TTC	GGC	GTA	AAG	660
201	F	D	L	N	V	A	K	K	F	I	K	Y	L	L	Q	E	F	G	V	K	220
661	AAC	AGC	AAC	AGC	AAC	CAA	ATG	GGC	ACT	CGG	GCA	GCT	GTT	ATT	GCG	TTT	GGA	ACG	ACC	ATT	720
221	N	S	N	S	N	Q	M	G	T	R	A	A	V	I	A	F	G	T	T	I	240
721	GAT	ATC	ATT	TTC	AAC	CTG	AAT	GAC	ACC	CAC	ATT	GAT	TCT	CCT	GAC	GCA	GCC	GGA	GTA	GCT	780
241	D	I	I	F	N	L	N	D	T	H	I	D	S	P	D	A	A	G	V	A	260
781	CTT	GAC	GAA	TTA	CTG	AAG	CCT	AAT	GGT	GGG	GGA	ACA	GCA	ATG	GTT	CCG	GCA	CTA	ACC	TCA	840
261	L	D	E	L	L	K	P	N	G	G	G	T	A	M	V	P	A	L	T	S	280
841	GTG	TTT	TTG	AAA	GTT	CAG	CCC	TTC	CTT	CGT	GAA	AGG	TCC	AAG	CAT	GCC	TTG	TTT	CTC	ATT	900
281	V	F	L	K	V	Q	P	F	L	R	E	R	S	K	H	A	L	F	L	I	300
901	AGC	GAC	GGA	GAA	CCA	AAC	ATC	GGC	GGC	GAT	GAC	ATC	ACT	CCC	GAA	GAC	ATA	TCG	AGG	CAG	960
301	S	D	G	E	P	N	I	G	G	D	D	I	T	P	E	D	I	S	R	Q	320
961	CTC	AAG	GAC	CAC	GCT	GCT	TTT	GAA	ATC	TTT	GCA	GTA	GGC	ATT	GGT	CCC	GAT	GTG	CAG	CGA	1020
321	L	K	D	H	A	A	F	E	I	F	A	V	G	I	G	P	D	V	Q	R	340
1021	AAA	ACA	TTG	GCC	AGC	GTC	GCA	TCC	GAA	CCC	GTT	TTG	CAT	CAC	GTT	TTC	ATG	CTT	GAT	AAA	1080
341	K	T	L	A	S	V	A	S	E	P	V	L	H	H	V	F	M	L	D	K	360
1081	TTC	ACC	GAC	CTC	GAG	GAC	ATG	ATG	GCC	ATC	ATG	AAA	AAT	CGT	CCG	ACT	GAA	GAA	CCG	CCT	1140
361	F	T	D	L	E	D	M	M	A	I	M	K	N	R	P	T	E	E	P	P	380
1141	GTT	AGC	CTG	GAT	AGA	TGT	GGT	CGC	ATG	GAT	GGA	AGG	AAT	GAG	AGC	GAG	GGC	TGG	CCG	TGG	1170
381	V	S	L	D	R	C	G	R	M	D	G	R	N	E	S	E	G	W	P	W	400
1171	GTC	GCA	GCA	GTA	TAC	GTG	TCT	TAT	GAA	GCA	AAC	AAC	TCT	ACG	CAA	CTA	AGC	CTA	TGC	GCC	1260
401	V	A	A	V	Y	V	S	Y	E	A	N	N	S	T	Q	L	S	L	C	A	420
1261	GGC	GCC	CTT	TTC	TGT	AGC	CAA	TGG	GTG	TTG	ACG	ACT	GCC	TCT	TGC	TTC	TTT	TAT	ATT	GAG	1320
421	G	A	L	F	C	S	Q	W	V	L	T	T	A	S	C	F	F	Y	I	E	440
1321	GAA	AAC	AGA	GTT	AGT	TTT	AAT	ATA	GAG	GAA	GTG	TTT	GTT	GTT	ATG	GGA	GAA	TAT	GAC	TTG	1380
441	E	N	R	V	S	F	N	I	E	E	V	F	V	V	M	G	E	Y	D	L	460
1381	ATG	GGA	GAG	GAT	CCG	AAA	CAA	AGG	AAT	TTC	TAC	GCT	AGT	AAA	GTA	GTC	ATC	CAC	CCG	GAC	1440
461	M	G	E	D	P	K	Q	R	N	F	Y	A	S	K	V	V	I	H	P	D	480

FIGURA 10 - Sequência de nucleotídeos e deduzida de aminoácidos de Lox-FB (Parte 2)

1441	TAT	GAT	GAA	TCT	GAC	ACA	CTG	CTG	AAC	AAT	TTG	GCA	CTG	GTC	AAG	CTT	CAC	AAG	CCT	GTG	1500
481	Y	D	E	S	D	T	L	L	N	N	L	A	L	V	K	L	H	K	P	V	500
1501	CCA	CTT	GAG	ACG	TTC	AAA	CCC	ATT	TGT	TTA	CCT	CCG	ACC	GAT	AGG	ACA	GTT	CCA	TTG	TTT	1560
501	P	L	E	T	F	K	P	I	C	L	P	P	T	D	R	T	V	P	L	F	520
1561	TTG	GAT	CAT	AAC	ATA	AAC	GCT	TCC	GTT	GCC	GGA	TGG	GGG	ATT	TCC	GCG	CCG	AAA	AAT	GGC	1620
521	L	D	H	N	I	N	A	S	V	A	G	W	G	I	S	A	P	K	N	G	540
1621	TTT	GAA	ACA	ATC	AGC	CCG	AGT	GCC	ACG	AGG	TCC	CAA	TTG	ACT	GTG	CGC	AAC	GTT	ATA	CTG	1680
541	F	E	T	I	S	P	S	A	T	R	S	Q	L	T	V	R	N	V	I	L	560
1681	GCA	GAA	GAA	CTT	GAA	TGC	CGA	AGG	AGA	ATC	AAC	CGA	AGG	TTA	AAC	GAA	ACC	ATA	GTT	TTT	1720
561	A	E	E	L	E	C	R	R	R	I	N	R	R	L	N	E	T	I	V	F	580
1721	TGC	GCA	GAA	ATC	AGT	GGC	ACT	ACT	GAC	AAT	TGT	GTA	GGC	AAT	CCT	GGT	TCC	CCA	ATA	ATA	1800
581	C	A	E	I	S	G	T	T	D	N	C	V	G	N	P	G	S	P	I	I	600
1801	GCC	AAA	GAT	CCC	CCC	ACC	GGA	ATC	AAC	CAC	GTT	TTG	GGA	CTC	TTG	AGC	CAG	AGG	AAT	AAA	1860
601	A	K	D	P	P	T	G	I	N	H	V	L	G	L	L	S	Q	R	N	K	620
1861	TGC	GAC	AAA	GGT	TCC	AGC	CAA	TAT	TTT	GAG	CTC	ACC	AAG	TAC	ATT	TCT	TGG	ATC	AAC	AAA	1920
621	C	D	K	G	S	S	Q	Y	F	E	L	T	K	Y	I	S	W	I	N	K	640
1891	GAA	TCA	GAC	AAA	TGT	CAC	TTT	TCC	CAC	TGG	AAT	TAG	1956								
631	E	S	D	K	C	H	F	S	H	W	N	*	651								

A sequência do cDNA foi obtida a partir da junção dos dois clones FB 5' e FB 3' e a sequência de aminoácidos foi deduzida com auxílio do programa Bioedit versão 7.0.9.0.

FIGURA 11 – Sequência deduzida de aminoácidos de Lox-FB**A**

```

      10      20      30      40      50      60      70
MASVEWAILG FLFWISTVIR QAGSSVCRPL PYSSLRNGNI IAESGGTFRF ECKNGYFLGG PSKIRCYKGN
      80      90     100     110     120     130     140
WTSERAPVCI ESGEFCDAPO KIANAKLHGP TVYEIDSVIQ YMCPRPGYVLV GNGRRIQLRS GYWSGITPTC
     150     160     170     180     190     200     210
MDESEPLQNV AERLNKEFVT KMAHSNDSS IQQAYYLSA SRRRGLDLVI LVDRSSSVDP FDLNVAKKFI
     220     230     240     250     260     270     280
KYLLEQFVGK NSNSNQMGTR AAVIAFGTTI DIIFNLNDTH IDSPDAAGVA LDELLKPNGG GTAMVPALTS
     290     300     310     320     330     340     350
VFLKVQPFLLR ERSKHALFLI SDGEPNIGGD DITPEDISRQ LKDHAAFEIF AVGIGPDVQR KTLASVASEP
     360     370     380     390     400     410     420
VLHHVFMLDK FTDLEDMMAI MKNRPTEEP VSLDRCGRMD GRNESEGWPW VAAVYVSIEA NNSTQLSLCA
     430     440     450     460     470     480     490
GALFCSQWVL TTASCFEYIE ENRVSFNIEE VFVVMGEYDL MGEDPKQRNF YASKVVIHPD YDESDTLNN
     500     510     520     530     540     550     560
LALVKLHKPV PLETFKPICL PPTDRTVPLF LDHNI NASVA GWGISAPKNG FETISPSATR SOLTVRNVIL
     570     580     590     600     610     620     630
AEELECRRI NRRLNETIVF CAEISGTTDN CVGNPGSPII AKDPPTGINH VLGLLSQRNK CDKGSSQYFE
     640     650
LTKYISWINK ESDKCHFSHW N

```

Ponto isoeletrico 5,73/MW:72,38 KDa**B**

[A] O peptídeo sinal está sublinhado e compreende os primeiros 25 aminoácidos. Os sítios de glicosilação e as cisteínas estão marcados em cinza e preto, respectivamente. [B] Arquitetura de domínios da sequência da proteína Lox-FB determinada com o uso da ferramenta BLAST (*Blast Local Alignment Tool*). As análises de predição para detectar o peptídeo sinal e os sítios de glicosilação foram realizadas com a ferramenta GPMW, disponível no endereço eletrônico <http://www.expasy.org/proteomics>.

4.3 Alinhamento múltiplo de Lox-FB com outras proteínas FB/C2

Embora as proteínas C2/FB de outros organismos, descritas na literatura até o momento, apresentem o mesmo padrão de composição de domínios, existem certas particularidades que as tornam diferentes como a quantidade de domínios CCPs, presentes na porção N-terminal. Por exemplo, o fator B encontrado em algumas espécies pertencentes ao Filo Echinodermata (SMITH; SHIH; DACHENHAUSEN, 2001), límulos (ZHU et al., 2005) e uma das isoformas presentes na anêmona-do-mar (KIMURA; SAKAGUCHI; NONAKA, 2009) contêm cinco domínios CCP ao invés de três como no fator B de humanos, o primeiro a ser caracterizado. Lox-FB apresenta somente dois domínios CCPs, assim como foi observado em Rd-FB, fator B proveniente do molusco bivalve *Ruditapes decussatus* (PADRO-ALVAREZ et al., 2009). Devido a essas diferenças, as sequências com mais de dois domínios CCPs tenderam a ficar fora do alinhamento. Cada módulo CCP tem cerca de 60 aminoácidos com alguns resíduos conservados, incluindo quatro cisteínas que se alinham perfeitamente (**FIGURA 12**).

FIGURA 12 – Alinhamento dos domínios CCPs do fator B humano (Hu-FB) e Lox-FB

CCP 1	Hu-FB	CSLE-GVEIK-----GGSFRLQEGQALEYVCPSGFYF	YPVQTRTCRSTGSWSTLKTQDQKTVRKAEC
CCP 2	Hu-FB	CPRPHDFENGEYWPRSPYYNVSDE---ISFHCYDGYTL	RGSANRTCQVNGRWSGQTA-----IC
CCP 3	Hu-FB	CSNP-GIPIGTRKV-GSQYRLDS---VTYHCSRGLTL	RGSQRRTCQEGGSWSGTEP-----SC
CCP 1	LoxFB	C-RP--LPYSSLRNGNIIAESGGT---FREECNKGYFL	GGPSKIRCYK-GNWTSERAP-----VC
CCP 2	LoxFB	CDAPQKIANAKLHG-PTVYEIDSV---IQYMCRFGYVL	VGNRRICLRSGYWSGITP-----TC
		C P	F C G G C G W C
			Y

Os domínios CCPs estão alinhados entre si e os resíduos de aminoácidos equivalentes estão assinalados em negrito e se encontram na base do alinhamento. As pontes dissulfeto sustentam o domínio CCP e são formadas entre a primeira e a terceira cisteína e entre a segunda e a quarta cisteína. O alinhamento foi feito com o algoritmo CLUSTAL W pelo programa Bioedit v.7.0.9.0.

Considerando os outros dois domínios, foi possível detectar muitos sítios conservados, tais como os cinco aminoácidos envolvidos na ligação ao fragmento C3b dependente de íons Mg^{2+} , presente no domínio vWFA; sete cisteínas e os aminoácidos, localizados próximos ao sítio ativo, também se dispuseram alinhados no domínio de serino protease (SP). No entanto, nenhum dos três resíduos de aminoácidos importantes para a atividade de serino protease (triade catalítica) foi identificado (**Figura 13**). Aparentemente, a triade catalítica clássica composta pelos aminoácidos histidina (H), ácido aspártico (D) e

serina (S) foram substituídos por outros resíduos: serina (S), asparagina (N) e prolina (P), respectivamente. Interessantemente, Lox-FB, assim como a terceira isoforma do componente FB-like da aranha *Hasarius adansoni* e do bivalve *Ruditapes decussatus* são similares, já que todos possuem somente dois domínios CCPs, dois resíduos de cisteína adicionais e nenhum deles preservaram a tríade catalítica (**FIGURA 13**). Considerando somente as espécies de aranhas, todas elas compartilham os mesmos aminoácidos na primeira e segunda posição, com exceção do último aminoácido. Dessa forma, é possível que estas proteínas não apresentem atividade proteolítica, mas, ainda assim, Lox-FB apresenta resíduos de aminoácidos conservados que circundam o sítio ativo como os aminoácidos treonina (T⁵⁴), alanina (A⁵⁶) e serina (S²¹⁴) (numeração da quimiotripsina) que desempenham um importante papel na estabilização da tríade catalítica (PAGE; CERA, 2008).

A estrutura resultante de Lox-FB após ser submetida à modelagem computacional, utilizando como base a estrutura cristalina do fator B humano (PDB 2ok5.1), apresentou uma identidade de 23,02 %. Apesar das diferenças em relação ao número de CCPs, foi possível observar uma notável similaridade estrutural entre os domínios CCP2-CCP3 do fator B humano e CCP1-CCP2 de Lox-FB (**FIGURA 14**). Aparentemente, o domínio vWFA de Lox-FB se sobrepõe perfeitamente ao do fator B humano, já que as principais α -hélices que circundam a região das folhas β estão presentes e conservadas nas mesmas posições com pequenas diferenças na conformação dos *loops*. Assim como foi observado para o alinhamento da estrutura primária, os aminoácidos que representam o sítio de ligação dependente de íons Mg^{2+} (MIDAS) estão conservados e ocupam as mesmas posições na estrutura tridimensional (**FIGURA 15A**). Com relação ao domínio serino protease (SP), a sobreposição não foi tão perfeita quanto àquela observada para o domínio vWFA, mas as regiões que constituem as estruturas secundárias de folha β se encontram sobrepostas (**FIGURA 15B**). Os três aminoácidos que compõem a tríade catalítica do fator B humano (H, D, S) se alinharam com os mesmos três aminoácidos (S, N, P) de Lox-FB observado no alinhamento das sequências primárias.

FIGURA 13 – Alinhamento múltiplo de Lox-FB (Parte 1)

FIGURA 13 – Alinhamento múltiplo de Lox-FB (Parte 2)

		*		*		*		#										
Lox-FB	RSQLTVRNVILA---	EELBC	-----R-RRINRRRLNETIV	CAE	---ISG-TTDNCVGNP	SGS	PII---AKDPPTG	INHVLLSQRN--	619									
Hd-FB	AFDMSSSSVSLs---	EEREC	-----PIRRRVTPPL	CAG	---QGSKTCFALVSG	SPLM---	AEDSSTGFQ	HILGLVDRR--	600									
Hd-FB	SPNLKQLRIPIQ---	SNERC	-----SKS-LEK-KLDPTFF	TDRMF	CAG	---TGETGVDTCLGD	SGGFLMQSMN	SEGFLYTQV	GIVSWG	819								
Hd-FB	SQILKELLPIQ---	SKQRC	-----TQSLLDN-KVATD	HFTDRMF	CAG	---DGKRGNDTC	KGDSGGFLMQ	SQSLNSEG	FLWTFQV	GIVSWG	973							
Tt-FB	VDQLKQIHLPIQ---	SRETC	-----VQSLNTK	DPMDFMI	CAG	---DGRGVADTC	QGDSSGGFLA	QSLDSEGM	NYWTFQV	GIVSWG	862							
Tt-FB	IDHLKQIRLPIQ---	NHETC	-----LKSLEKTK	KEMDFMI	CAG	---DSEGVDTCK	KGDSGGFMA	QSLVDDA	EMNYWTFQV	GIVSWG	945							
Rd-FB	EEKLMQNRQSVV---	KRKQC	-----R-KRYGDI	DMERFF	CVG	---EEFQESNS	CMGDI	GGPYM---	VQMSN-KRY	YSLGLAK	TR	653						
Nv_FB	YPVLLKQVLPFV---	SHRVC	-----Q-VNHTNAI	TKRMR	CAG	---DVMGERD	CKGDSGS	PIV---	VKRTD-GS	WAVGLSS	WGE	674						
Bb-FB	TLVPQALCLPIV---	SQPVC	-----K-KSTVVP	ITERQL	CAG	---YRFQEKD	ACRGDS	SGGPLV---	VRHRD-NL	WGIVGIMS	WGE	724						
Sp_FB	SPALMEIMI PPV---	LDSSC	-----SIAMSAH	IAVDITTE	CAG	---IERKDS	QCGDS	SGGPLV---	VQRNN-KY	RQI	GIVSYGI	808						
Aj_FB	EPHLQELILQLR---	RRREC	-----EAS-LDG-N	EEDWFDE	TMT	CAG	---YLGREVG	PCAGDS	SGGPLVK-	YAKRGN	SRRWTFV	GIVSWG	877					
Hu-FB	EKKLTRKEVYIK	NGDKK	GSC-ERDAQY	APGDKVD	ISEVVT	PRFL	CTGGV	SPYADP	NTCRGDS	SGGELI-	-VHKRS-	RFIQV	GIVSWG	VVD	723			
Hu-C2	GSKLINILNKM	GV---	EWTS	CAEVVS	QEKTMF	PNLTDV	REVVT	DQFL	CSG	---TQE-DE	SPCKG	SGGAVF	---	LERRF-R	FFQVGLV	SWG	LYN	703

	*										
Lox-FB	HC	---DKGS	-----SQMFEL	--TKYIS	WINKESDK	CHF	SHWN	-----	651		
Hd-FB3	HC	---SKKGQ	-----NQMIEL	--TRHIG	WINQATS	NCQL	KHWGVAK	-----	636		
Hd-FB2	GC	---GMEDM	-----YGFYTHV	--QKLRS	WIDELI	ITYQ			847		
Hd-FB1	GC	---GKENT	-----YGYVTHV	--QKFRS	WIDSTIE	AAMAAIQ			1006		
Tt-FB1	GC	---KNRGQ	-----YGFYTHV	--AKLRP	WIDKVMNS				889		
Tt-FB2	GC	---KLRGH	-----YGFYTHV	--AKLRP	WIDKVLNS				972		
Rd-FB	GC	---TKRNN	-----YSIFL	NLLHPTVSD	LSQVLDG	GNRVK	DKDAKKTDEQ		697		
Nv_FB	GC	---AQK GK	-----FGV	YADLLSAEYD	LWITRT	AGLVR	PGS		708		
Bb-FB	GC	---AQSNK	-----YGINTRV	--DTFLD	WITDTT	GLN			752		
Sp_FB	GC	---GVT	-----YGVVTRV	--PHYVD	WINGI	INK			833		
Aj_FB	GC	---ALQTE	-----LDF	FANV	--PHFV	PEINENIN	PEARLRNT		913		
Hu-FB	VC	---KNQKRQ	KQVP	-----AHARD	FHINL	--FQVLP	WLKEKLQ	DEDLGFL	764		
Hu-C2	BC	GLSG	LNRRKR	APRSKVP	PPPRDF	HINL	--FRMQP	WLRLQHLG	LDVNLN	FL	752

Alinhamento múltiplo dos domínios vWFA e serino protease de Lox-FB com sequências proteicas de FB/C2 de outros organismos. O alinhamento foi produzido utilizando o algoritmo MUSCLE disponível no programa MEGA. Aminoácidos destacados em preto indicam regiões idênticas. O sítio putativo de clivagem pelos components Fator D ou C1s é demonstrado pela seta roxa; os aminoácidos que compõem o sítio de ligação dependente de íons Mg^{2+} (MIDAS) e a tríade catalítica do domínio de serino protease estão indicados por setas azuis e vermelhas, respectivamente. Os três resíduos de aminoácidos (T, A, S) que são importantes para estabilização da tríade catalítica estão indicadas com o sinal # e os resíduos de cisteína conservados estão indicados em asterisco e as duas cisteínas adicionais presentes em Lox-FB, Hd-FB3 e Rd-FB estão destacados em cinza. *Loxosceles laeta* (Lox-FB), *Hasarius adansonii* (Hd-FB1; HD-FB2; Hd-FB3), *Branchiostoma belcheri* (Bb-FB; ABY28382.1), *Nematostella vectensis* (Nv-FB; BAH22726.1), *Ruditapes decussatus* (Rd-Fb; ACQ91095.1), *Apostichopus japonicus* (Aj-FB; ADX36428.1), *Strongylocentrotus purpuratus* (Sp-FB; AAC79682.1), *Tachypleus tridentatus* (Tt-FB-2; BAM15263.1), *Homo sapiens* (Hu-C2; P06681) (Hu-FB; P00751).

FIGURA 14- Modelo tridimensional de Lox-FB**FB Humano** **Lox-FB**

Construção do modelo molecular de Lox-FB baseado na estrutura do fator B humano (PDB 2ok5.1), utilizando a ferramenta SWISS-Model Workspace disponível no endereço eletrônico <http://swissmodel.expasy.org/workspace/>.

FIGURA 15 – Comparação da estrutura terciária do domínio vWFA e SP do FB humano

A

vWFA

MIDAS

B

SP

Tríade Catalítica

Sobreposição dos domínios vWFA de FB humano (rosa) e Lox-FB (cinza) **[A]** e dos domínios SP de FB humano (rosa) e Lox-FB (cinza) **[B]** e os resíduos de aminoácidos que compõem a tríade catalítica do FB humano (H,D,S assinalados em rosa) e Lox-FB (S,N,P assinalados em preto). As análises posteriores foram efetuadas utilizando o programa SwissPDBViewer.

4.4 Análises de similaridade e filogenia de Lox-FB

Lox-FB apresentou similaridade com as proteínas do complemento, apresentando identidade de 25% e 26% com FB e C2 de humanos, respectivamente (**TABELA 1**). Em relação ao FB e C2 de invertebrados, os valores variaram entre 22 e 43% de identidade. Lox-FB apresentou maior identidade (43%) e maior similaridade (62%) com a isoforma 3 do componente FB-like da aranha *Hasarius adansoni*.

Para investigar a história evolutiva dos componentes C2 e fator B e determinar como a sequência das aranhas *Loxosceles* se inclui neste panorama, 56 sequências foram submetidas à análise filogenética. Uma árvore filogenética não enraizada foi construída, resultando em dois principais grupos, (98% de *Bootstrap*), um constituído pelas proteínas C2/FB de vertebrados e pelas ascídias, pertencentes ao Filo Urochordata e, o outro, pelas proteínas C2/FB-like encontradas nos demais organismos invertebrados (**FIGURA 16**). Considerando os componentes de vertebrados (cinza), existe uma clara separação entre os componentes C2 e FB, exceto para alguns componentes C2/FB dos peixes ósseos. O Fator B de lampreias, animais representantes dos vertebrados não mandibulados (Agnatha), está posicionado fora do grupo pertencente aos componentes C2 e FB de vertebrados mandibulados (Gnathostomata), sugerindo que o Fator B de lampreia represente um ancestral das proteínas FB e C2 de vertebrados superiores.

Analisando somente o grupo dos invertebrados, há três principais grupos: um deles é representado pelas sequências FB-like de ascídias (clado A) provavelmente o grupo invertebrado ancestral das sequências FB/C2 de vertebrados; o segundo grupo foi denominado clado B que compreende Lox-FB e sequências FB-like de cnidários, anfioxos, artrópodes e bivalves. Lox-FB está localizado no mesmo ramo das proteínas FB (isoforma 2) da lacraia *S. subspinipes* e da “aranha-do-mar” *Ammothea sp.* e no mesmo sub-ramo no qual está também presente a isoforma 3 das aranhas *H. adansoni* (*Bootstrap* de 100%). Finalmente, o terceiro grupo denominado clado C, consiste de dois subgrupos, um com sequências FB-like de equinodermos e o outro com oito isoformas encontradas em cinco espécies pertencentes ao Filo Arthropoda.

TABELA 1 - Análise comparativa da sequência Lox-FB com outras sequências similares

Organismo	Nº de Acesso	I (%)*	S (%)**	Score	E value
FB-3 <i>Hasarius adansoni</i>	dbj BAR45592.1	43	62	483	1e-158
FB-2 <i>Ammothoa sp</i>	dbj BAR45607.1	35	52	349	1e-106
FB/C2 <i>B. belcheri tsingtauense</i>	gb ABY28382.1	30	48	264	3e-74
FB precursor <i>Nematostella vectensis</i>	dbj BAH22727.1	30	46	237	1e-64
FB precursor <i>Nematostella vectensis</i>	dbj BAH22726.1	30	46	237	1e-64
FB-2 <i>S. subspinipes japonica</i>	dbj BAR45617.1	28	44	231	9e-63
FB precursor <i>Nematostella vectensis</i>	dbj BAH22728.1	28	46	224	2e-59
FB-like <i>Ruditapes decussatus</i>	gb ACQ91095.1	27	47	206	7e-54
FB-1 <i>Ammothoa sp</i>	dbj BAR45606.1	26	41	185	8e-46
FB-2 <i>Apostichopus japonicus</i>	gb AEP68015.1	27	42	180	1e-44
FB <i>Apostichopus japonicus</i>	gb ADX36428.1	27	44	180	1e-44
FB-1 <i>Hasarius adansoni</i>	dbj BAR45590.1	25	45	171	2e-41
FB <i>Strongylocentrotus purpuratus</i>	gb AAC79682.1	28	42	169	5e-41
FB <i>Lethenteron japonicum</i>	dbj BAA02763.1	23	40	154	3e-36
C2/FB-2 <i>Tachypleus tridentatus</i>	gb BAM15263.1	24	42	149	3e-34
FB-1 <i>S. subspinipes japonica</i>	dbj BAR45616.1	24	42	142	1e-31
FB <i>Danio rerio</i>	gb AAH97100.1	25	43	140	2e-31
C2/FB-1 <i>Tachypleus tridentatus</i>	gb BAM15262.1	22	41	139	5e-31
FB/C2B <i>Ginglymostoma cirratum</i>	gb AAY56127.1	25	41	135	5e-30
C2/FB-2 <i>C. rotundicauda</i>	gb ABK30937.1	23	42	131	1e-28
C2/FB-1 <i>C. rotundicauda</i>	gb AAV65032.2	23	42	131	2e-28
FB/C2A <i>Cyprinus carpio</i>	gb BAA34706.1	25	41	126	5e-27
FB/C2A-3 <i>Cyprinus carpio</i>	gb BAB32650.1	25	44	125	8e-27
FB-2 precursor <i>Mus musculus</i>	gb NP_001136178.1	24	42	125	8e-27
FB-1 precursor <i>Mus musculus</i>	gb NP_032224.2	24	42	125	1e-26
C2 <i>Homo sapiens</i>	gb AAB67975.1	26	41	119	4e-32
FB <i>Homo sapiens</i>	gb CAA51389.1	25	41	121	8e-33

*Identidade (I) foi calculada com base na porcentagem de aminoácidos idênticos por coluna/posição no alinhamento.

**Similaridade (S) foi calculada com base na porcentagem de resíduos não idênticos, mas similares, que são substituições conservativas (mantendo as propriedades físico-químicas dos aminoácidos)

O valor do E value de um hit, para um dado score corresponde a probabilidade de se obter, com outra sequência aleatória de mesmo tamanho e composição de letras, outro alinhamento com score igual ou superior. Quanto mais próximo de zero for o E value, mais confiável é a consulta.

FIGURA 16 – Árvore filogenética das sequências similares a Lox-FB

Relações filogenéticas entre Lox-FB a proteínas C2 e FB identificadas. A árvore foi construída com base no alinhamento feito com algoritmo MUSCLE e Máxima Verossimilhança (Modelo LG +G +I) utilizando o programa MEGA 6, *Bootstrap* de 1000 repetições. As sequências foram obtidas do banco de dados do *GenBank* Gorilla FB (Q864V9), Chimpanzee FB (Q864W0), Human FB (P00751), Orangutan FB (Q864W1), Marmoset FB (JAB17376.1), Mouse FB-1 (P04186), Mouse FB-2 (B8JIM5), Ox-FB (P81187), Boar FB (ABX82825.1), Snake FB (AAR21601.1), Frog FB (BAA06179.1), Frog FB-2 (AAI70192.1), Trout C2/FB (ADY68777.1), Trout FB 1 (AAC83699.1), Pufferfish FB/C2 OS (CAD21938.1), Catfish FB/C2 A (AEW10545.1), zebrafish FB (AAB19093.1), Carp FB/C2 A3 (BAB32650.1), Carp FB/C2 A2 (BAA78416..1), Carp FB/C2 A (BAA34706.1), Houndshark FB (BAB63203.1), Houndshark FB 2 (BAF62177.1), Catfish FB (AEW38662.1), Carp FB/C2 B (BAA34707.1), Trout C2/FB 2 (BAB19788.1), Mouse C2 (P21180), Ox C2 (Q3SYW2), Orangutan C2 (Q8SQ74), Gorilla C2 (Q863A0), Human C2 (P06681), Chimpanzee C2 (Q8SQ74), Lamprey FB (BAA02763.1), Lamprey FB 2 (BAG66069.1), Ascidia FB (AAK00631.1), Ciona FB-3 (BAD89301.1), Ciona FB-1 (BAD89299.1), Ciona FB-2 (BAD89300.1), Amphioxus FB/C2 (ABY28382.1), Anemone FB (BAH22726.1), Anemone FB 2 (BAH22728.1), Bivalve FB (ACQ91095.1), Sea cucumber FB 2 (AEP68015.1), Sea cucumber FB (ADX36428.1), Sea urchin FB (AAC79682.1), Horseshoe crab 1 C2/FB 1 (BAM15262.1), Horseshoe crab 1 C2/FB 2 (BAM15263.1), Horseshoe crab 2 C2/FB 1 (AAV65032), Horseshoe crab 2 C2/FB 2 (ABK30937.1), Jumping spider FB1 (BAR45590.1), Jumping spider FB2 (BAR45591.1), Jumping spider FB3 (BAR45592.1), sea spider FB 1 (BAR45606.1), sea spider FB 2 (BAR45607.1), centipede FB 1 (BAR45616.1), centipede FB 2 (BAR45617.1).

4.5 Clonagem do componente C3-like de *L. laeta* (Lox-C3)

4.5.1 Amplificação da região que contém o motivo tioéster da sequência Lox-C3

As aranhas *L. laeta* pertencem ao Filo Arthropoda e estão classificadas dentro do subfilo Chelicerata que inclui, também, os límulos *Tachypleus tridentatus* e *Carcinoscorpius rotundicauda*, e a aranha *Hasarius adansoni*. Este último foi o que apresentou maior identidade (49%) com a EST correspondente ao C3-like de *L. laeta* (EY188746) dentre todas as sequências C3-like identificadas.

Uma vez que a EST C3-like de *L. laeta* continha a sequência de nucleotídeos que codifica para parte do domínio A2M_receptor e que este domínio estava também presente na sequência C3-like da aranha *H. adansoni* (Haad-C3-1), foi determinado que este domínio se encontrava na região que codifica para a porção C-terminal da proteína, mais precisamente, na posição 1341 a 1519 da sequência de aminoácidos (**FIGURA 17A**). Utilizando os *primers* degenerado C3_TEP_2 *sense* (região tioéster; posição 1095) e o *antisense* (região A2M-receptor, posição 1517) e comparando com a sequência Haad-C3-1, foi estimado o tamanho de pares de bases esperado, após a reação de PCR, que seria de cerca de 1265 pb (**FIGURA 17B**). Após a reação de PCR, utilizando esses *primers*, foi observada a amplificação de um produto com cerca de 1200 pb (**FIGURA 18A**). Este produto de PCR foi purificado e clonado no vetor *pGEM-Teasy*. Após a etapa de transformação, somente seis colônias cresceram e, a partir destas, os cDNAs foram extraídos e submetidos à digestão com a enzima EcoRI (**FIGURA 18B**). Ao final desta reação, foi observado que somente dois clones apresentavam o inserto do tamanho esperado, sendo os cDNAs destes purificados e sequenciados.

As sequências de DNA dos clones 6 e 7 foram determinadas e foi observado que estes possuíam 1241 nucleotídeos, mas apresentavam diferenças na composição em 16 destes e em 9 aminoácidos. As sequências de nucleotídeos dos clones 6 e 7, assim como as regiões de cada um dos *primers sense* e *antisense*, estão demonstradas na **Figura 19**. À esta região do gene C3 foi atribuído o nome de Lox-C3_TEP e após deduzir a sequência proteica, composta por 413 aminoácidos, e analisá-la por bioinformática, observou-se a presença de parte de um domínio contendo o motivo tioéster (**FIGURA 20**).

Figura 17 - Análise da sequência EST Lox-C3

A

Score: 143 bits (360), Expect: 2e-35,					
Identidade: 90/184(49%) Similaridade 109/184(59%) Gaps: 13/184(7%)					
EST Lox-C3	1	GNGVGLLTVKLRYNVNPPEKYCKFDVKVNVTEEEKENHG-----VFNLPPKEKKFVLL	53		
		G+G+G L+VKLRYN+ P E CKFDVKVNVTE K V N K+ F			
Haad-C3-1	1341	GHGMGSLSVKLRYNLNPVEDICKFDVKVNVTESKPEQQEKDPWDEVLNQKGKDM-FERF	1399		
EST Lox-C3	54	QDDLIRGVKGALDAPPDEIIAEVDPRNRLQ-RSSRIGFHSAAVLSSRVKREDKMDNSNLV	112		
		+DLIRG+K LD P E+ +DPR R R+S I S + R KR+ +NS +			
Haad-C3-1	1400	PEDLIRGIKKELDEVPPEL--AIDPRRRGHVRASLIA--SPLIKKDRQKRQKPNENSKVK	1455		
EST Lox-C3	113	LAIKICARYLGDHENGMTIVDAGIFSGFKVKEDDLKKLVKSAEHFIDRYELSKRGIVFYFYL	172		
		L I IC RYLG+ + M+IVDAGIFSGF+ EDDL L H I RYE S RG+VFYFYL			
Haad-C3-1	1456	LNITICVRYLGNKDTEMSIVDAGIFSGFQPVEDDLIWLQDDHSHLIQRYEKSSRGVVFYFYL	1515		
EST Lox-C3	173	GKVP	176		
		KVP			
Haad-C3-1	1516	QKVP	1519		

B

Alinhamento de aminoácidos das sequências EST Lox-C3 (EY188746) e C3-like da aranha *Hasarius adansoni* (Haad-C3-1) [A] e arquitetura dos domínios de ambas as sequências com localização dos *primers* para amplificação da região que contém o domínio tioéster (Lox-C3_TEP) [B] utilizando como base a ferramenta Blastp (<http://blast.ncbi.nlm.gov/Blast.cgi>).

FIGURA 18 - Amplificação da região que contém o motivo tioéster de Lox-C3

Análise do produto amplificado após reação de PCR utilizando os *primers* degenerado *sense* C3 TEP_2 (região tioéster) e C3 *antisense* (região A2M_receptor) [A] e do tamanho dos fragmentos de cDNA, inseridos no vetor *pGEM-Teasy* após digestão com *EcoRI* [B], em géis de agarose a 1%.

FIGURA 19 – Componente C3-like – Lox-C3_TEP

A

Primer sense C3_TEP_2

AGCACATGGCTGACGGCTTTTATTATCAAAGTGTCTGTCAAGCAAGCAAAATGGTTCACATAGATGAAGACGTGATGTGCAG
 TGGTGTGAAATGGCTTGTGAGTAATCAGCAAGCAGATGGCTCATTTCATTGAACATAATCCAGTGTATCACACGAAATGATGG
 GTGGTGTAAAGGACGAACCCATTGACGCATTTGTTCTGATTACATTACAAGAATGCAGTTGTGAAATTCCAAATATTCAC
 ATTGTAAAAGCAAGGGCTTTGTCTTATCTGGAAGGACAATTGGGAGATATAAGTGAGCCTTTACCTCTAGCAATAGCTGCTTA
 TGC GTTAGCACTGTCAGAGAGTGAAGTGAAGCAAGCAATCAATAACAAGCTTCTGGCCCTTTCCAAATACGATGCAGATACAA
 ACACGATGTATTGGGAAACCGAAAGTAGTTCATACAATGTGGAAGCCACTGGGTATGGACTTCTCGCGCAGCTTACTTTGTTT
 GAAATGGCAAGGAGCAACTCTATTGTGAAGTGGTGGACGGACAGAGAATGGCACTGGTGCATTCAAATCTACGCAGGATAC
 TGTAATTGCACTCAAGCACTTTCAGAGTATGCCATTAGAGCTAAGACTCCAGCCTTAAACCTCGTATCTAATCTCACGTGCGA
 GTAACGATAGAAACTTCATGCAAACTAATGTTTGATCAAAGTAATGCGCTAATCCTGCAGGAGCACAAGCTTGATAAACCT
 GGAGGAACATTATTTGTCAATACAGCTGGAAATGGTGTGGACTTTTGACGGTGAAAGTTAAGATATAATGTCCAAATCCACC
 AGAAAAGTATTGCAAGTTTGATGTTAAAGTGAACGTTACTGAAGAAAAAGAAAACCATGGTGTTTTAACTTACCACCAAAGG
 AAAAAAATTCGTCTTGCTTCAAGATGACCTTATTAGAGGATAAAAGGTGCTTTAGATGCCCTCCTGATGAGATAATTGTT
 GAAGTCGATCCTCGTAATAGGCTCCAACGTTTCATCGCGAATGGGTTCCATTACGAGCAGTTTATCATCACGTGTTAAGAG
 GGAGGACAAAATGGATAATAGCAACTTAGTGTGGCAATAAAGATTTGTGCAAGATATTTAGGCGACCATGAAAACGGCATGA
 CGATTGTTGATGCTGGTATCTTCTCTGGTTTCAAAGTGAAGGAAGATGATTGAAAAAGCTTGTCAAGAGTGCAGGAACA
Primer C3 antisense

B

Primer sense C3_TEP_2

AGCACATGGCTGACGGCTTTTATTATCAAAGTGTCTGTCAAGCAAGCAAAATGGTTCACATAGATGAAGACGTGATGTGCAG
 TGGTGTGAAATGGCTTGTGAGTAATCAGCAAGCAGATGGCTCATTTCATTGAACATAATCCAGTGTATCACACGAAATGATGG
 GTGGTGTAAAGGACGAACCCATTGACGCATTTGTTCTGATTACATTACAAGAATGCAGTTGTGAAATTCCAAATATTCAC
 ATTGTAAAAGCAAGGGCTTTGTCTTATCTGGAAGGACAATTGGGAGATATAAGTGAGCCTTTACCTCTAGCAATAGCTGCTTA
 TGC GTTAGCACTGTCAGAGAGTGAAGTGAAGCAAGCAATCAATAACAAGCTTCTGGCCCTTTCCAAATACGATGCAGATACAA
 ACACGATGTATTGGGAAACCGAAAGTAGTTCATACAATGTGGAAGCCACTGGGTATGGACTTCTCGCGCAGCTTACTTTGTTT
 GAAATGGCAAGGAGCAACTCTATTGTGAAGTGGTGGACGGACAGAGAATGGCACTGGTGCATTCAAATCTACGCAGGATAC
 TGTAATTGCACTCAAGCACTTTCAGAGTATGCCATTAGAGCTAAGACTCCAGCCTTAAACCTCGTATCTAATCTCACGTGCGA
 GTAACGATAGAAACTTCATGCAAACTAATGTTTGATCAAAGTAATGCGCTAATCCTGCAGGAGCACAAGCTTGATAAACCT
 GGAGGAACATTATTTGTCAATACAGCTGGAAATGGTGTGGACTTTTGACGGTGAAAGTTAAGATATAATGTCCAAATCCACC
 AGAAAAGTATTGCAAGTTTGATGTTAAAGTGAACGTTACTGAAGAAAAAGAAAACCATGGTGTTTTAACTTACCACCAAAGG
 AAAAAAATTCGTCTTGCTTCAAGATGACCTTATTAGAGGATAAAAGGTGCTTTAGATGCCCTCCTGATGAGATAATTGTT
 GAAGTCGATCCTCGTAATAGGCTCCAACGTTTCATCGCGAATGGGTTCCATTACGAGCAGTTTATCATCACGTGTTAAGAG
 GGAGGACAAAATGGATAATAGCAACTTAGTGTGGCAATAAAGATTTGTGCAAGATATTTAGGCGACCATGAAAACGGCATGA
 CGATTGTTGATGCTGGTATCTTCTCTGGTTTCAAAGAGAAAGGAAGATGATTGAAAAAGCTTGTCAAGAGTGCAGGAACA
Primer C3 antisense

Sequências de nucleotídeos dos clones 6 [A] e clone 7 [B], produtos resultantes da amplificação da reação de PCR com os *primers* C3 *sense* degenerado C3 TEP_2 (região tioéster) e C3 *antisense*. Os nucleotídeos que se diferem entre os clones 6 e 7 estão destacados.

FIGURA 20 – Análise da sequência Lox-C3_TEP

A

Score = 814 bits (2102), Expect = 0.0, Method: Compositional matrix adjust. Identities = 404/413 (98%), Positives = 407/413 (99%), Gaps = 0/413 (0%)					
C3tep2_6	1	STWLTAFIIKVFCQASKMVHIDEDVMCSGVKWLVSNNQADGSFIEHNPVYHTEMMGGVKG	60		
C3tep2_7	1	STWLTAFIIKVFCQASKMVHIDEDVMCSGVKWLVSNNQADGSFIEHNPVYHTEMMGGVKG	60		
C3tep2_6	61	RTPLTAFVLITLQECSCIEPNIHIVKARALSYLEGHLGDI SEPLPLAIAAYALALSESEL	120		
C3tep2_7	61	QTPLTAFVLITLQECSCIEPNIHIVKARALSYLEGHLGDI SEPLPLAIAAYALALSESEL	120		
C3tep2_6	121	KQAINNKLLALS KYDADTNTMYWETESSYNVEATGYLLAQLTLFEMARSNSIVNWL DG	180		
C3tep2_7	121	KQAINNKLLALS KYDADTNTMYWETESSYNVEAT Y LLAQLTLFEMARSNSIVNWL+G	180		
C3tep2_6	181	QRMATGAFKSTQDTVIALQALSEYAIRAKTPALNLVSNLTSSNDRNFMQTL MFDQSNALI	240		
C3tep2_7	181	QRMATGAFKSTQDTVIALQALSEYAIRAKTPALNLVSNLTSSNDRNFMQTL MFDQSNALI	240		
C3tep2_6	241	LQEHKLDKPGGTLFVNTAGNGVGLLTVKLRYNVPNPPEKYCKFDVKVNVTEEEKENHGVFN	300		
C3tep2_7	241	LQEHKLDK GGTLFVNTAGNGVGLLTVKLRYNVPNPPEKYCKFDVKVNVTEEEKENHGVFN	300		
C3tep2_6	301	LPPKEKKFVLLQDDLIRGVKGALDAPPDEIIVEVDPRNRLQRSSRIGFHSAAVLSSRVKR	360		
C3tep2_7	301	LPPKEKKFVLLQDDLIRG+KGALDAPPDEII EVDPRNRLQRSSR GFHSAAVLSSRVKR	360		
C3tep2_6	361	EDKMDNSNLVLAIKICARYLGDHENGMTIVDAGIFSGFKVKEDDLKKLVKSAE	413		
C3tep2_7	361	EDKMDNSNLVLAIKICARYLGDHENGMTIVDAGIFSGFK KEDDLKKLVKSAE	413		

B

Alinhamento de aminoácidos das sequências Lox-C3_TEP clone 6 e Lox-C3_TEP clone 7 [B] e arquitetura dos domínios de ambas as sequências [B] determinada utilizando como base a ferramenta Blastp (<http://blast.ncbi.nlm.gov/Blast.cgi>).

4.5.2 Amplificação da região que contém o domínio ANATO da sequência Lox-C3

Após amplificação parcial do domínio contendo a região tioéster, foram desenhados novos *primers*: 5' CACACCACTGCACATCACGTCTTC 3' (*primer antisense* as1_TEP2) e 5' GAAGACGTGATGTGCAGTGGTGTG 3' (*primer sense* tioéster). Por meio de análise comparativa com a sequência de C3 da aranha *H. adansoni*, foi notado que faltava, ainda, identificar cerca de 3000 nucleotídeos da porção 5' do gene C3 de *L. laeta* e, utilizando o segundo *primer sense* degenerado C3 A2M_2 (posição 625; região ANATO) e o *primer antisense* as1_TEP2 (posição 1118; região tioéster), foi estimado o tamanho de pares de bases esperado após a reação de PCR que seria de, aproximadamente, 1479 pb (**FIGURA 21**).

FIGURA 21 – Sequência parcial Lox-C3

Comparação da sequência parcial C3-like de *L. laeta* (Lox-C3_TEP) e C3-like da aranha *H. adansoni* (Haad-C3-1) e localização dos *primers* para amplificação da região que compreende o domínio anato (Lox-C3_ANATO), utilizando como base a ferramenta Blastp (<http://blast.ncbi.nlm.gov/Blast.cgi>).

Após a reação de PCR, foi possível observar a amplificação de um único produto de cerca de 1500 pb (**FIGURA 22A**). Para aumentar as chances de clonar este produto, foi feita uma segunda reação de PCR (2X PCR) utilizando a enzima *Taq polymerase* (New England Biolabs, MA, EUA), resultando na amplificação do mesmo DNA nesta segunda reação (**FIGURA 22B**). Este foi, então, purificado e clonado no vetor *pGEM-Teasy*. Após transformação das bactérias, foram selecionadas 16 colônias das quais somente quatro clones apresentavam o inserto de tamanho com cerca de 1500 pb (**FIGURA 22C**).

A sequência de DNA composta por 1497 nucleotídeos, assim como as regiões de cada um dos *primers sense* e *antisense* estão demonstradas na **FIGURA 23A**. A esta região do gene

C3 foi atribuído o nome de Lox-C3_ANATO e após deduzir a sequência proteica, composta por 499 aminoácidos, e analisá-la por bioinformática, observou-se a presença de três domínios: o domínio ANATO, um domínio A2M e a porção 5' do domínio que contém o motivo tioéster (**FIGURA 23B**).

FIGURA 22 - Amplificação da região que contém o domínio ANATO de Lox-C3

Análise do produto amplificado após reação de PCR **[A]** e 2X PCR **[B]**, utilizando os *primers* degenerado *sense* C3 A2M_2 (região ANATO) e C3 *antisense* as1-TEP2 (região tioéster), e do tamanho dos fragmentos de cDNA, inseridos no vetor *pGEM-Teasy* após digestão com EcoRI **[C]**.

FIGURA 23 – Componente C3-like – Lox-C3_ANATO**A****Primer sense C3 A2M_2**

GACTTAGGCTGCGGTCCTGGTGGAGGTTTGACAACAGCATCTGTTTTACGAAATGCAGGTGTCGTGATG
 ATGACAAGGAATTATTCTCCAACAAATGATGCAAACAGTAGAGTTTCCGCGTGTGAAGAAAAAGTACGA
 AAGAAACGGGATGTATTAGACACGAAATTAAAGGAGTATGATGGAATTCTGCGACAATGTTGTGCTCTG
 GGAATGAATCACGATCGTTACTCCCGAACATGTCAACAAAGATTTGAAATTGTTAAGATGCACCTGAAA
 GATGAAGACAGGAATTGTTCTCAAACATTCTTGGACTGTTGTATGATTGCCGAAAAACATGGAGCGACT
 GGACTGACATTAATCCAGCCACAAAAATTTCAATCTGCCAGAATGGGTGCTTTTGGAGAAGACATTCAA
 TTCATTACCGAGGAAGAAGAAGAGGCATTTGAATCAGTAGTTTCTATTAGAAGTGATTTTCAGAGAACT
 TGGCTGTTTGTATGATGTTATTATTGGGCTGACGGTAAACACGAATTTCCAGTCAGTTTGCCCCAGAGT
 ATCACAACCTGGATTATTCAAGCAGTGTCCATATCTCAAGAACATGGAATGTGTGTGGCAACACCTGAA
 AGTGTAGAATCCTTCAAGAAGGTGTTTGTCCAATTAAATTTACCGTATTCTGTTGTGAAGAACGAGCAA
 ATCGAAGTTCGAGCAACAGTGTTCGAATTATGGAGACAAAAATATTTTAGTAGTTGTCTACATGTATGGT
 GTGAAAGACTTGTGCATTGGAAGTGAAGGCAAGAAGAGTGAAAGAAAGAGAGTGTTAGTGCATTCC
 AATTCTGCGCAGAGCGTTTCATTCCCTATCATTCCACTTGCGGAAGGAGTTTTTGACATCCAAGTAACT
 GCACTCAGTCCTCATGGAAGTGATGTGATAATACGAAAGCTAAATGTTGTGGCGGAAGGTAAGACTGTT
 GAAGAAGACATCACCATCAAATTAGATCCTTCAAATCAACAAAGGCGCAAACGGCGGCATGTAACAACT
 AATGACTATTGAGATTCTATTGATAACGTGCAGAAGGTACAAATTACCGCTATAAAGCTCAAGGCTCCA
 GAGGATTTTATTCCGGGAACAGAAAGTTGTACTATTACTGCTATAGGTGATCAGTTTGGTCCAACAGTC
 CAACTTCTCTTGAAAATCCAGAAAATCTCATAGACAGTCCAAGAGGTGTGCTGAGCAAGCCATGATT
 TATTTGGCACCAACCCTGTATACTATGAAGTACTTGAAAGCTACTAATAAAATCACTGCAGACACAGAA
 ACGAGAGGTTATGATTTTATTAGGAAAGGATACGACTTAATATTGACATATAGGAAAAATGACGGCTCC
 TATGGAGCTTGGGCGCACACACCACAAAGTACATGGCTGACTGCTTTTATTATCAAAGTGTTCTGTCAA
 GCAAGCAAAATGGTTCACATAGATGAAGACGTGATGTGCAGTGGTGTG

Primer antisense C3 as1_TEP2**B**

DLGCGPGGGLTTASVLRNAGVMMTRNYSPTNDANSRVSACEEKVRKKRDVLDTKLKEYDGILRQCCA
 LGMNHDRYSRTCQQRFEIVKMLKDEDRNCSQTFLDCCMIAEKHGATGLTLIQPQKFQSARMGAFGED
 IQFITEEEEEAFESVVSIRSDFRETWLFDDVIIGPDGKHEFPVSLPQSITTWIIQAVSISQEHGMCVA
 TPESVESFKKVFVQLNLPYSVVKNEQIEVRATVFNYGDKNILVVVYMGVKDLICIGTEEGKKSERKRV
 LVHSNSAQSVSFPIIPLAEGVFQVTDALSPHGSVDVIRKLNVAEGKTVEEDITIKLDPSNQRRKR
 RHVTTNDYSDSIDNVQKVQITAIKLKAPEDFIPGTESCTITAIGDQFGPTVQTSLENPENLIDSPRGC
 AEQAMIYLAPTLYTMKYLKATNKITADTETRGYDFIRKGYDLILTYRKNDGSYGAWAHTPQSTWLTA
 FIIKVFCQASKMVHIDEDVMCSGV

Sequências de nucleotídeos [A] e deduzida de aminoácidos com seus respectivos domínios [B] do produto resultante da amplificação da reação de PCR com os *primers* C3 sense degenerado C3 A2M_2 (região ANATO) e C3 *antisense* as1_TEP_2 (região tioéster).

4.5.3 Amplificação da região 5' do cDNA de Lox-C3

Com a identificação do domínio ANATO, foram desenhados novos *primers*: 5' GGTCCTGGTGGAGGTTTGACAAC 3' (*anato sense*) e 5' CGATCGTGATTTCATCCAGAGCAC 3' (*anato antisense*). Comparando à sequência de C3 da aranha *H. adansoni* (Haad-C3-1), foi observado que restava ainda identificar cerca de 1875 nucleotídeos para a porção 5' do gene C3 de *L. laeta* (**Figura 24**).

Após a reação de PCR, na qual foram utilizados os *primer anato antisense* (posição 700; ANATO de Haad-C3-1) e universal comercial UPM (Clontech, CA, EUA) (**Figura 24**), foi possível observar a amplificação de um produto de, aproximadamente, 2500 pb (**FIGURA 25A**). Este produto de PCR foi purificado, clonado no vetor *pGEM-Teasy*. Após a etapa de transformação, cDNA dos clones positivos foi extraído e submetido à digestão com a enzima EcoRI (**Figura 25B**) e, a partir da análise do gel de agarose 1%, foi observado que os clones 5, 7 e 9 apresentavam o tamanho esperado, sendo estes selecionados para a etapa de sequenciamento.

Com o auxílio dos *primers* universais T7, e Sp6 utilizados para sequenciar o vetor *pGEM-Teasy* e outros *primers* especificamente desenhados para resolver as regiões intermediárias do gene C3 5', foi possível determinar a sequência do clone 9 Lox-C3_5', composta por 2169 nucleotídeos.

FIGURA 24 – Amplificação da região terminal de Lox-C3 5'

Comparação das sequências parciais C3-like de *L. laeta* (Lox-C3_TEP e Lox-C3_ANATO) e C3-like da aranha *H. adansoni* (Haad-C3-1) e localização dos *primers* para amplificação da região terminal 5' utilizando como base a ferramenta Blastp (<http://blast.ncbi.nlm.gov/Blast.cgi>).

FIGURA 25 – Amplificação e clonagem do produto Lox-C3 da reação de 5' RACE-PCR

Análise dos produtos amplificados após reação de PCR [A] e do tamanho dos fragmentos de cDNA, inseridos no vetor *pGEM-Teasy* após digestão com EcoRI [B], em géis de agarose a 1%.

4.5.4 Amplificação da região 3' do cDNA de Lox-C3

Após determinar a sequência 5' do gene Lox-C3, o próximo passo seria juntar as sequências dos 2 clones obtidos anteriormente (Lox-C3_TEP e Lox-C3_ANATO), mas como observamos uma variação entre as duas sequências obtidas para a região Lox-C3_TEP (clones 6 e 7), a junção dos três clones (Lox-C3 5', Lox-C3_TEP e Lox-C3_ANATO) não foi realizada. Foi, então, amplificada a região 3' com a mesma técnica utilizada para determinar a porção 5', ou seja, o RACE-PCR. Para isso, foram utilizados três *primers sense* posicionados em três regiões diferentes em combinação com o *primer* universal UPM (**FIGURA 26**).

Após a reação 3' RACE-PCR, os produtos mais abundantes resultantes da reação contendo os *primers* tioéster *sense* e ANATO *sense* foram de, aproximadamente, 2000 e 3000 pares de bases, respectivamente (**FIGURA 27**). Não foi possível observar um produto mais definido na reação contendo o *primer* C3_ sense, posicionado mais próximo da região da extremidade 3' (C345c), mas, independente disso, todos os três produtos PCR totais foram purificados e clonados no vetor *pGEM-Teasy*. Após a digestão dos clones positivos com a enzima *EcoRI*, cinco clones dos seis selecionados para os produtos C345c (Clone Lox-C3 3'_A) (clones 1-4 e 6) e Tioéster (Clone Lox-C3 3'_B) (clones 7, 9, 10 e 12) apresentaram insertos com o tamanho esperado, sendo que os clones 2 e 10 foram sequenciados. Com relação ao produto ANATO (Clone Lox-C3 3'_C), os clones 16 e 17 foram escolhidos para a etapa de sequenciamento.

FIGURA 26 – Amplificação da região Lox-C3 3'

Comparação da sequência 5' C3-like de *L. laeta* (Lox-C3_5') e C3-like da aranha *H. adansonii* (Haad-C3-1) e localização dos *primers* para amplificação da região terminal 3' utilizando como base a ferramenta Blastp (<http://blast.ncbi.nlm.gov/Blast.cgi>).

FIGURA 27 - Amplificação e clonagem do produto Lox-C3 resultante da reação de 3' RACE-PCR

Análise dos produtos amplificados após reação de PCR [A] e do tamanho dos fragmentos de cDNA, inseridos no vetor *pGEM-Teasy* após digestão com EcoRI [B], em géis de agarose a 1%.

Após análise das sequências, foi verificado que os cDNAs dos clones 2 (Lox-C3 3'_2A), 10 (Lox-C3 3'_10B) e 17 (Lox-C3 3'_17C) faziam parte do mesmo gene. As sequências Lox-C3 3'_10B e Lox-C3 3'_17C apresentavam 1977 e 2330 nucleotídeos, respectivamente. O clone Lox-C3 3'_17C continha a porção 3' do gene C3 incompleta, não apresentando a porção A2M_receptor e o domínio C345c presente na porção C-terminal. Esse resultado pode ser explicado por uma falha na ligação do *primer* universal UPM na região da cauda de poli-A do RNA mensageiro. Sendo assim, a sequência total Lox-C3_3' foi resolvida a partir da junção das sequências do clone Lox-C3 3'_10B e Lox-C3 3'_17C, gerando uma sequência de 3438 nucleotídeos; em relação ao clone 16 (Lox-C3 3'_16C), foi possível identificar a porção 3' completa que continha, no total, 3434 nucleotídeos (**FIGURA 28**). Os dois clones Lox-C3 3' resultantes, quando alinhados entre si, apresentavam um valor de identidade de 88%, diferindo em 403 nucleotídeos, o que demonstra que a aranha *Loxosceles* possa apresentar duas isoformas C3-like (**TABELA 2**).

Para verificar se as sequências dos clones 10 e 17 (Lox-C3 3'_10B+17C) e clone 16 (Lox-C3 3'_16) correspondiam aos clones Lox-C3_TEP ou Lox-C3_ANATO, anteriormente gerados a partir de dois primers degenerados, os clones foram alinhados entre si. O alinhamento par a par do clone Lox-C3 3'_10B+17C com os clones 6/7 (TEP) e clone ANATO, resultou em uma identidade de 99%, variando em número de 15, 13 e 8 nucleotídeos, respectivamente. Os valores de identidade encontrados para Lox-C3 3'_16 foram menores variando em torno de 86 a 90%. (**TABELA 2**).

O clone C3 3' total correspondente a esses dois clones (Lox-C3 3'_10B+17C) foi alinhado com a sequência C3 5' (Lox-C3 5') obtida anteriormente por 5' RACE-PCR. O alinhamento produzido por esses dois clones resultou em uma sobreposição de 100%, resolvendo, portanto, a sequência completa de Lox-C3 (**FIGURA 29A**).

O clone 16 (Lox-C3 3'_16C) continha a porção 3' completa, totalizando 3434 nucleotídeos, mas quando este foi comparado com as sequências C3 5', não houve sobreposição de 100% para aquelas regiões comuns (**FIGURA 29B**). Como houve essa diferença, não foi possível identificar a sequência completa do gene C3 correspondente ao clone 16C, faltando assim, ainda, identificar a sua porção 5'.

FIGURA 28 – Sequência Lox-C3 3'

A porção 3' do gene que codifica para o componente C3 de *L. laeta* (Lox-C3 3') foi resolvida a partir do clone Lox-C3 3' 16C e por meio da junção da sequência do clone 17C e 10B (Lox-C3 3' 17C+10B) resultando em uma sequência total de 3434 e 3438 pares de bases, respectivamente.

TABELA 2 – Valores de identidade (%) com base no alinhamento par a par de diferentes clones Lox-C3

CLONES	Clone 6_TEP	Clone_7_TEP	Clone_ANATO	Clone (10B + 17C) 3'	Clone 16C 3'
Clone 6_TEP	-	99% (16 nt)	99% (2 nt)	99% (15 nt)	89% (132 nt)
Clone 7_TEP	99% (16 nt)	-	99% (2 nt)	99% (13 nt)	90% (130 nt)
Clone_ANATO	98% (2 nt)	98% (2nt)	-	99% (8 nt)	86% (203 nt)
Clone 10B + 17C 3'	99% (15 nt)	99% (13 nt)	99% (8nt)	-	88% (403 nt)
Clone_5'	X*	X*	100%	100%	89% (32 nt)

*Não houve sobreposição dos clones

Os números entre parênteses indicam a quantidade de nucleotídeos diferentes entre dois clones

FIGURA 29 – Alinhamento de Lox-C3 5' e Lox-C3 3'**A****Identidade:** 212/212 (100%)

```

Lox-C3 (9) 5'   1958  GGCCCTGGTGGAGGTTTGACAACAGCATCTGTTTTACGAAATGCAGGTGTCGTGATGATG  2017
                      |||
Lox-C3 (17+10) 3'   1   GGCCCTGGTGGAGGTTTGACAACAGCATCTGTTTTACGAAATGCAGGTGTCGTGATGATG  60

Lox-C3 (9) 5'   2018  ACAAGGAATTATTCTCCAACAAATGATGCAAACAGTAGAGTTCCGCGTGTGAGGAAAAA  2077
                      |||
Lox-C3 (17+10) 3'  61  ACAAGGAATTATTCTCCAACAAATGATGCAAACAGTAGAGTTCCGCGTGTGAGGAAAAA  120

Lox-C3 (9) 5'   2078  GTACGAAAGAAACGGGATGTATTAGACACGAAATTAAAGGAGTATGATGGAATTCTGCGA  2137
                      |||
Lox-C3 (17+10) 3' 121  GTACGAAAGAAACGGGATGTATTAGACACGAAATTAAAGGAGTATGATGGAATTCTGCGA  180

Lox-C3 (9) 5'   2138  CAATGTTGTGCTCTGGGAATGAATCACGATCG  2169
                      |||
Lox-C3 (17+10) 3' 181  CAATGTTGTGCTTTGGGAATGAATCACGATCG  212

```

B**Identidade:** 189/212 (89%)

```

Lox-C3 (9) 5'   1958  GGCCCTGGTGGAGGTTTGACAACAGCATCTGTTTTACGAAATGCAGGTGTCGTGATGATG  2017
                      || |||
Lox-C3 (16)3'   1   GGTCTCTGGTGGAGGTTTGACAACAGCATCCGTTTTACGAAATGCAGGTGTTGTGATGTTA  60

Lox-C3 (9) 5'   2018  ACAAGGAATTATTCTCCAACAAATGATGCAAACAGTAGAGTTCCGCGTGTGAGGAAAAA  2077
                      ||| |||
Lox-C3 (16)3'   61  ACACGGAATTATTCTCCATCAGATAGTGCAAACAACAGAGTTCCGCGTGTGAGGAAAAA  120

Lox-C3 (9) 5'   2078  GTACGAAAGAAACGGGATGTATTAGACACGAAATTAAAGGAGTATGATGGAATTCTGCGA  2137
                      ||| ||| |||
Lox-C3 (16) 3'  121  GTACGAAAGAAACGTAGTGTAATAAAAACGAAATTAAAGGAGTACGATGGAATTCTGCGT  180

Lox-C3 (9) 5'   2138  CAATGTTGTGCTCTGGGAATGAATCACGATCG  2169
                      || |||
Lox-C3 (16) 3'  181  CAGTGTGTGCTTTGGGAATGAATCGCGATCG  212

```

Alinhamento de nucleotídeos das sequências Lox-C3 5' (clone 9) e Lox-C3 3' (clone 17 +10) [A] e Lox-C3 5' (clone 9) e Lox-C3 3' (clone 16) [B] utilizando a ferramenta blastn disponível no endereço eletrônico (<http://blast.ncbi.nlm.gov/Blast.cgi>).

4.6 Caracterização do C3-like de *L. laeta*

Os clones provenientes dos produtos C3 5' (Lox-C3 5'_clone9) e C3 3' (Lox-C3 3'_10B e Lox-C3 3'_17C) geraram um cDNA composto por 5178 pares de bases que codifica para uma proteína madura de 1726 aminoácidos (**FIGURA 30**). A massa molecular de Lox-C3 foi mensurada por métodos bioinformáticos resultando num valor de aproximadamente 193 kDa e ponto isoelétrico de 8, sem considerar os oito sítios de glicosilação putativos. O peptídeo sinal é composto por 20 aminoácidos e a proteína madura por 1706 aminoácidos (**FIGURA 31A**). Lox-C3 apresenta uma organização de domínios similar àquela observada para os componentes C3/C4/C5 de vertebrados, ou seja, que apresenta um domínio de anafilatoxina central (ANATO), um domínio C345c na extremidade da porção C-terminal, além dos outros domínios A2M_N presentes tanto na porção N-terminal como na região C-terminal (**FIGURA 31B**).

FIGURA 30 – Sequências de cDNA e de aminoácidos deduzida de Lox-C3 (Parte 1)

1	ATG	CGT	GCA	TCT	GTT	TTA	ACG	CTA	CTT	GTA	GCT	ACG	ATC	TTG	TCT	TCA	GCT	AAA	GCC	TGG	60
1	M	R	A	S	V	L	T	L	L	V	A	T	I	L	S	S	A	K	A	W	20
61	AAT	GTT	TAT	GTT	GCT	GCA	CCT	AGT	CAC	ATA	CGC	GTG	GGT	GTG	GAA	GAA	AAA	GTA	GCC	ATT	120
21	N	V	Y	V	A	A	P	S	H	I	R	V	G	V	E	E	K	V	A	I	40
121	GTG	GTG	GAA	GGA	AAT	GCT	GCA	CCT	CGG	AAC	TTC	GAG	GTC	TAC	GTG	CAA	GAC	CAT	CCT	GGG	180
41	V	V	E	G	N	A	A	P	R	N	F	E	V	Y	V	Q	D	H	P	G	60
181	AAA	ATA	AAG	AAT	CTT	TCC	TCG	ACT	TTC	GGC	ACT	GTC	TTT	TCA	AAT	AGT	CCT	AAA	GTA	GTT	240
61	K	I	K	N	L	S	S	T	F	G	T	V	F	S	N	S	P	K	V	V	80
241	TCA	TTG	TCG	ATT	AAT	CCA	TCA	GAC	GTT	CCA	GAT	TTC	ATA	CCT	GGT	GAC	CCT	TCG	ACA	TTG	300
81	S	L	S	I	N	P	S	D	V	P	D	F	I	P	G	D	P	S	T	L	100
301	AGC	TTC	GTT	AGC	CTT	ATA	GTG	CAT	GGC	GGA	TCC	TTT	CAT	AAG	GAA	GTT	GTC	TTG	CCC	CTT	360
101	S	F	V	S	L	I	V	H	G	G	S	F	H	K	E	V	V	L	P	L	120
361	TCT	TAC	AAA	ACT	GGA	TAT	GTG	TTT	ATT	CAA	ACT	GAC	AAG	CCT	ATT	TAT	ACT	AAA	AAA	GAA	420
121	S	Y	K	T	G	Y	V	F	I	Q	T	D	K	P	I	Y	T	K	K	E	140
421	CAA	GTG	AAG	TTC	CGA	ATC	ATA	CCA	TTG	AAT	GAA	GAT	GCT	AGA	CCT	TCT	GCT	GAG	AAA	ATT	480
141	Q	V	K	F	R	I	I	P	L	N	E	D	A	R	P	S	A	E	K	I	160
481	AAA	TTG	CGT	ATC	AAC	AAT	CCC	AGA	GGT	CTT	ACC	GTG	CAA	GAC	ATG	GAG	TTC	AGT	GGA	AAA	540
161	K	L	R	I	N	N	P	R	G	L	T	V	Q	D	M	E	F	S	G	K	180
541	TCT	TTC	TTG	ACG	CCT	GAG	CCA	TAT	ACA	TTT	CCA	TCC	TAT	CCA	ATT	CTG	GGA	GTC	TGG	TCA	600
181	S	F	L	T	P	E	P	Y	T	F	P	S	Y	P	I	L	G	V	W	S	200
601	GCT	ACG	GTT	CTG	CAT	GGG	TAC	AAG	TTT	CAG	CAG	AAT	ACT	ACG	GTT	CAG	TTT	GAA	TTG	CAA	660
201	A	T	V	L	H	G	Y	K	F	Q	Q	N	T	T	V	Q	F	E	L	Q	220
661	GAA	TAT	GTG	CTG	CCA	ACT	TTC	ACC	GTG	AAA	CTA	AAG	GTT	CAA	GAA	GTT	ATT	TTG	CCT	TCA	720
221	E	Y	V	L	P	T	F	T	V	K	L	K	V	Q	E	V	I	L	P	S	240
721	GCT	ACA	ACC	ATT	GAC	ATT	TCA	GTT	GAA	TCA	CAA	TTC	GTG	TAT	GGA	AAG	CCA	GTT	AGT	GGC	780
241	A	T	T	I	D	I	S	V	E	S	Q	F	V	Y	G	K	P	V	S	G	260
781	ACT	GTA	ATA	TTC	AGG	ATA	GGC	ATA	AAA	GGT	GAA	ACA	CCA	GAA	GTT	ACG	TAT	TTT	GCA	ATT	840
261	T	V	I	F	R	I	G	I	K	G	E	T	P	E	V	T	Y	F	A	I	280
841	CTT	GAT	TCA	AAG	AAG	CTG	GTG	GAC	GGA	TCA	TAC	AAA	ACA	AGA	ATC	AAC	ACA	GAA	CAG	CTT	900
281	L	D	S	K	K	L	V	D	G	S	Y	K	T	R	I	N	T	E	Q	L	300
901	AAA	CGT	GAC	AAG	TCG	ATT	GGC	TGG	TTT	CCT	GAG	ATT	GAA	GGT	TCG	CGC	TTT	GTC	GTG	GAG	960
301	K	R	D	K	S	I	G	W	F	P	E	I	E	G	S	R	F	V	V	E	320
961	GCT	ACT	GTA	ACA	GAT	GAT	GCA	ACA	GGA	CAA	AAG	GAA	GTT	GTA	ATA	GAT	TCC	AAT	AGT	TAC	1020
321	A	T	V	T	D	D	A	T	G	Q	K	E	V	V	I	D	S	N	S	Y	340
1021	TTT	TCA	ACG	TCG	CCA	TTT	GTT	ATT	TCT	TTC	AAG	AGA	TGT	CTG	AGT	GAC	TTC	AAA	CCT	GGG	1080
341	F	S	T	S	P	F	V	I	S	F	K	R	C	L	S	D	F	K	P	G	360
1081	TTA	ATC	TCA	ATT	ATA	GAG	GCT	GAT	GTC	AAT	TAC	GTT	GAC	GGT	CGA	CCT	GCA	TCT	AAG	ATC	1140
361	L	I	S	I	I	E	A	D	V	N	Y	V	D	G	R	P	A	S	K	I	380
1141	AAG	ACA	ACT	ATC	TCT	GCA	GTA	GCT	GAC	AAT	GGC	GTT	CCC	TTG	AAA	ATA	ACG	AAC	CCC	GTT	1200
381	K	T	T	I	S	A	V	A	D	N	G	V	P	L	K	I	T	N	P	V	400
1201	TCT	GTA	AGT	AAT	GAG	GAA	GGA	AAA	ATA	TTG	TTC	GAA	ATA	CAA	CCA	GAA	TTG	GGA	CAT	TCA	1260
401	S	V	S	N	E	E	G	K	I	L	F	E	I	Q	P	E	L	G	H	S	420
1261	GTT	GTT	ACA	GTA	ACA	GTG	ACA	ACA	AGT	GAT	CCA	AAG	TAT	GCT	GGT	AGT	CAA	GCA	GTT	GGT	1320
421	V	V	T	V	T	V	T	T	S	D	P	K	Y	A	G	S	Q	A	V	G	440
1321	CGC	CAC	ATT	CAA	ACC	AAG	TTT	GAA	TCC	AAA	AAT	GCT	GCC	TAT	ATT	GCA	ATT	GCC	AGG	GAC	1380
421	R	H	I	Q	T	K	F	E	S	K	N	A	A	Y	I	A	I	A	R	D	460

FIGURA 30 – Sequências de cDNA e de aminoácidos deduzida de Lox-C3 (Parte 2)

1381	AGT	AAG	GAT	AAA	CCC	AAG	CCT	GGA	ACA	ATG	TTT	ACT	GAA	ATA	GTT	TAT	ATG	GAG	TCT	GAA	1440
461	S	K	D	K	P	K	P	G	T	M	F	T	E	I	V	Y	M	E	S	E	480
1441	GAC	ATG	ACT	GCT	GTT	TAT	TAT	ATG	GTT	GTA	TCT	AGA	GGA	CGA	ATC	GCT	CTT	ATA	GGA	AAG	1500
481	D	M	T	A	V	Y	Y	M	V	V	S	R	G	R	I	A	L	I	G	K	500
1501	CTT	CAG	GGG	CGG	TTT	AAA	ACG	TAC	AGA	GTC	GAG	TTC	ATG	GTG	ACA	CAC	GAG	ATG	GTG	CCA	1560
501	L	Q	G	R	F	K	T	Y	R	V	E	F	M	V	T	H	E	M	V	P	520
1561	AAC	TTT	CGT	CTT	GTG	GTT	TTT	TCT	GAG	TAT	AAA	AAT	GAG	CTG	ATT	GCT	GAT	TCT	CTC	CGT	1620
521	N	F	R	L	V	V	F	S	E	Y	K	N	E	L	I	A	D	S	L	R	540
1621	ATT	GAT	GTC	GCA	AAT	GAA	TGT	GAT	CCA	AAA	GCC	GAT	GTT	AAA	TTG	ACG	CCA	AGC	TTC	AAT	1680
541	I	D	V	A	N	E	C	D	P	K	A	D	V	K	L	T	P	S	F	N	560
1681	ATC	AAA	GAG	CCT	GGA	AGC	AAC	GGG	AAG	ATA	CTC	ATT	GAA	GGA	AGC	AAA	TCC	ACT	GTC	GTC	1740
580	I	K	E	P	G	S	N	G	K	I	L	I	E	G	S	K	S	T	V	V	580
1741	GGT	TTG	TTA	GCT	GTT	GAT	TCT	GCT	GTC	TAT	GCA	TTA	TCT	AAG	AGA	GAT	AAG	CTC	ACA	AGA	1800
581	G	L	L	A	V	D	S	A	V	Y	A	L	S	K	R	D	K	L	T	R	600
1801	GCC	AAG	ATG	TTT	AAG	ACT	CTG	GAA	AAA	CAC	GAT	TTG	GGA	TGT	GGC	CCT	GGT	GGA	GGT	TTG	1860
601	A	K	M	F	K	T	L	E	K	H	D	L	G	C	G	P	G	G	G	L	620
1861	ACA	ACA	GCA	TCT	GTT	TTA	CGA	AAT	GCA	GGT	GTC	GTG	ATG	ATG	ACA	AGG	AAT	TAT	TCT	CCA	1920
621	T	T	A	S	V	L	R	N	A	G	V	V	M	M	T	R	N	Y	S	P	640
1921	ACA	AAT	GAT	GCA	AAC	AGT	AGA	GTT	TCC	GCG	TGT	GAG	GAA	AAA	GTA	CGA	AAG	AAA	CGG	GAT	1980
641	T	N	D	A	N	S	R	V	S	A	C	E	E	K	V	R	K	K	R	D	660
1981	GTA	TTA	GAC	ACG	AAA	TTA	AAG	GAG	TAT	GAT	GGA	ATT	CTG	CGA	CAA	TGT	TGT	GCT	CTG	GGA	2040
661	V	L	D	T	K	L	K	E	Y	D	G	I	L	R	Q	C	C	A	L	G	680
2041	ATG	AAT	CAC	GAT	CGT	TAC	TCC	CGA	ACA	TGT	CAA	CAA	AGA	TTT	GAA	ATT	GTT	AAG	ATG	CAC	2100
681	M	N	H	D	R	Y	S	R	T	C	Q	Q	R	F	E	I	V	K	M	H	700
2101	CTG	AAA	GAT	GAA	GAC	AGG	AAT	TGT	TCT	CAA	ACA	TTC	TTG	GAC	TGT	TGT	ATG	ATT	GCC	GAA	2160
701	L	K	D	E	D	R	N	C	S	Q	T	F	L	D	C	C	M	I	A	E	720
2161	AAA	CAT	GGA	GCG	ACT	GGA	CTG	ACA	TTA	ATC	CAG	CCA	CAA	AAA	TTT	CAA	TCT	GCC	AGA	ATG	2220
721	K	H	G	A	T	G	L	T	L	I	Q	P	Q	K	F	Q	S	A	R	M	740
2221	GGT	GCT	TTT	GGA	GAA	GAC	ATT	CAA	TTC	ATT	ACC	GAG	GAA	GAA	GAA	GAG	GCA	TTT	GAA	TCA	2280
741	G	A	F	G	E	D	I	Q	F	I	T	E	E	E	E	E	A	F	E	S	760
2281	GTA	GTT	TCT	ATT	AGA	AGT	GAT	TTC	AGA	GAA	ACT	TGG	CTG	TTT	GAT	GAT	GTT	ATT	ATT	GGG	2340
761	V	V	S	I	R	S	D	F	R	E	T	W	L	F	D	D	V	I	I	G	780
2341	CCT	GAC	GGT	AAA	CAC	GAA	TTT	CCA	GTC	AGT	TTG	CCC	CAG	AGT	ATC	ACA	ACC	TGG	ATT	ATT	2400
781	P	D	G	K	H	E	F	P	V	S	L	P	Q	S	I	T	T	W	I	I	800
2401	CAA	GCA	GTG	TCC	ATA	TCT	CAA	GAA	CAT	GGA	ATG	TGC	GTG	GCA	ACA	CCT	GAA	AGT	GTA	GAA	2460
801	Q	A	V	S	I	S	Q	E	H	G	M	C	V	A	T	P	E	S	V	E	820
2461	TCC	TTC	AAG	AAG	GTG	TTT	GTC	CAA	TTA	AAT	TTA	CCG	TAT	TCT	GTT	GTG	AAG	AAC	GAG	CAA	2520
821	S	F	K	K	V	F	V	Q	L	N	L	P	Y	S	V	V	K	N	E	Q	840
2521	ATC	GAA	GTT	CGA	GCA	ACA	GTG	TTC	AAT	TAT	GGA	GAC	AAA	AAT	ATT	TTA	GTA	GTT	GTC	TAC	2580
841	I	E	V	R	A	T	V	F	N	Y	G	D	K	N	I	L	V	V	V	Y	860
2522	ATG	TAT	GGT	GTG	AAA	GAC	TTG	TGC	ACT	GGA	ACT	GAG	GAA	GGC	AAG	AAG	AGT	GAA	AGA	AAG	2640
861	M	Y	G	V	K	D	L	C	T	G	T	E	E	G	K	K	S	E	R	K	880
2641	AGA	GTG	TTA	GTG	CAT	TCC	AAT	TCT	GCG	CAG	AGC	GTT	TCA	TTC	CCT	ATC	ATT	CCA	CTT	GCG	2700
881	R	V	L	V	H	S	N	S	A	Q	S	V	S	F	P	I	I	P	L	A	900
2701	GAA	GGA	GTT	TTT	GAC	ATC	CAA	GTA	ACT	GCA	CTC	AGT	CCT	CAT	GGA	AGT	GAT	GTG	ATA	ATA	2760
901	E	G	V	F	D	I	Q	V	T	A	L	S	P	H	G	S	D	V	I	I	920
2761	CGA	AAG	CTA	AAT	GTT	GTG	GCG	GAA	GGT	AAG	ACT	GTT	GAA	GAA	GAC	ATC	ACC	ATC	AAA	TTA	2820
921	R	K	L	N	V	V	A	E	G	K	T	V	E	E	D	I	T	I	K	L	940

FIGURA 30 – Sequências de cDNA e de aminoácidos deduzida de Lox-C3 (Parte 3)

2821	GAT	CCT	TCA	AAT	CAA	CAA	AGG	CGC	AAA	CGG	CGG	CAT	GTA	ACA	ACT	AAT	GAC	TAT	TCA	GAT	2880
941	D	P	S	N	Q	Q	R	R	K	R	R	H	V	T	T	N	D	Y	S	D	960
2881	TCT	ATT	GAT	AAC	GTG	CAG	AAG	GTA	CAA	ATT	ACC	GCT	ATA	AAG	CTC	AAG	GCT	CCA	GAG	GAT	2940
961	S	I	D	N	V	Q	K	V	Q	I	T	A	I	K	L	K	A	P	E	D	980
2941	TTT	ATT	CCG	GGA	ACA	GAA	AGT	TGT	ACT	ATT	ACT	GCT	ATA	GGT	GAT	CAG	TTT	GGT	CCA	ACA	3000
981	F	I	P	G	T	E	S	C	T	I	T	A	I	G	D	Q	F	G	P	T	1000
3001	GTC	CAA	ACT	TCT	CTT	GAA	AAT	CCA	GAA	AAT	CTC	ATA	GAC	AGT	CCA	AGA	GGT	TGT	GCT	GAG	3060
1001	V	Q	T	S	L	E	N	P	E	N	L	I	D	S	P	R	G	C	A	E	1020
3061	CAA	GCC	ATG	ATT	TAT	TTG	GCA	CCA	ACC	TTG	TAT	ACT	ATG	AAG	TAC	TTG	AAG	GCT	ACT	AAT	3120
1021	Q	A	M	I	Y	L	A	P	T	L	Y	T	M	K	Y	L	K	A	T	N	1040
3121	AAA	ATC	ACT	GCA	GAC	ACA	GAA	ACG	AGA	GGT	TAT	GAT	TTT	ATT	AGG	AAA	GGA	TAC	GAC	TTA	3180
1041	K	I	T	A	D	T	E	T	R	G	Y	D	F	I	R	K	G	Y	D	L	1060
3181	ATA	TTG	ACA	TAT	AGA	AAA	AAT	GAC	GGC	TCC	TAT	GGA	GCT	TGG	GCG	CAC	ACA	CCA	CAA	AGT	3240
1061	I	L	T	Y	R	K	N	D	G	S	Y	G	A	W	A	H	T	P	Q	S	1080
3241	ACA	TGG	CTG	ACT	GCT	TTT	ATT	ATC	AAA	GTG	TTC	TGT	CAA	GCA	AGC	AAA	ATG	GTT	CAC	ATA	3300
1081	T	W	L	T	A	F	I	I	K	V	F	C	Q	A	S	K	M	V	H	I	1100
3301	GAT	GAA	GAC	GTG	ATG	TGC	AGT	GGT	GTG	AAA	TGG	CTT	GTG	AGT	AAT	CAG	CAA	GCA	GAT	GGC	3360
1101	D	E	D	V	M	C	S	G	V	K	W	L	V	S	N	Q	Q	A	D	G	1120
3361	TCA	TTC	ATT	GAA	CAT	AAT	CCA	GTG	TAT	CAC	ACA	GAA	ATG	ATG	GGT	GGT	GTT	AAA	GGA	CAA	3420
1121	S	F	I	E	H	N	P	V	Y	H	T	E	M	M	G	G	V	K	G	Q	1140
3421	ACC	CCA	TTG	ACT	GCA	TTT	GTT	CTG	ATT	ACA	TTA	CAA	GAA	TGC	AGT	TGT	GAA	ATT	CCA	AAT	3480
1141	T	P	L	T	A	F	V	L	I	T	L	Q	E	C	S	C	E	I	P	N	1160
3481	ATT	CAC	ATT	GTA	AAA	GCA	AGG	GCT	TTG	TCT	TAT	CTG	GAA	GGA	CAT	TTG	GGA	GAT	ATA	AGT	3540
1161	I	H	I	V	K	A	R	A	L	S	Y	L	E	G	H	L	G	D	I	S	1180
3541	GAG	CCT	TTA	CCT	CTA	GCA	ATA	GCT	GCT	TAT	GCG	TTA	GCA	CTG	TCA	GAG	AGT	GAA	CTG	AAG	3600
1181	E	P	L	P	L	A	I	A	A	Y	A	L	A	L	S	E	S	E	L	K	1200
3601	CAA	GCA	ATC	AAT	AAC	AAG	CTT	CTG	GCC	CTT	TCC	AAA	TAC	GAT	GCA	GAT	ACG	AAC	ACG	ATG	3660
1201	Q	A	I	N	N	K	L	L	A	L	S	K	Y	D	A	D	T	N	T	M	1220
3661	TAT	TGG	GAA	ACC	GAA	AGT	AGT	TCA	TAC	AAT	GTG	GAA	GCC	ACT	GGG	TAT	GGA	CTT	CTC	GCA	3720
1221	Y	W	E	T	E	S	S	S	Y	N	V	E	A	T	G	Y	G	L	L	A	1240
3721	CAG	CTT	ACT	TTG	TTT	GAA	ATG	GCA	AGG	AGC	AAC	TCT	ATT	GTG	AAC	TGG	TTG	AAC	GGA	CAG	3780
1241	Q	L	T	L	F	E	M	A	R	S	N	S	I	V	N	W	L	N	G	Q	1260
3781	AGA	ATG	GCA	ACT	GGT	GCA	TTC	AAA	TCT	ACG	CAG	GAT	ACT	GTA	ATT	GCA	CTG	CAA	GCA	CTT	3840
1261	R	M	A	T	G	A	F	K	S	T	Q	D	T	V	I	A	L	Q	A	L	1280
3841	TCA	GAG	TAT	GCC	ATT	AGA	GCT	AAG	ACT	CCA	GCT	TTA	AAC	CTC	GTA	TCT	AAT	CTC	ACG	TCG	3900
1281	S	E	Y	A	I	R	A	K	T	P	A	L	N	L	V	S	N	L	T	S	1300
3901	AGT	AAC	GAT	AGA	AAC	TTC	ATG	CAA	ACA	CTA	ATG	TTT	GAT	CAA	AGT	AAT	GCG	CTA	ATC	CTG	3960
1301	S	N	D	R	N	F	M	Q	T	L	M	F	D	Q	S	N	A	L	I	L	1320
3961	CAG	GAG	CAC	AAG	CTT	GAT	AAA	CTT	GGA	GGA	ACA	TTA	TTT	GTC	AAT	ACA	GCT	GGA	AAT	GGT	4020
1321	Q	E	H	K	L	D	K	L	G	G	T	L	F	V	N	T	A	G	N	G	1340
4021	GTT	GGA	CTT	TTG	ACG	GTG	AAA	TTA	AGA	TAT	AAT	GTC	CCA	AAT	CCA	CCA	GAA	AAG	TAT	TGC	4080
1341	V	G	L	L	T	V	K	L	R	Y	N	V	P	N	P	P	E	K	Y	C	1360
4081	AAG	TTT	GAT	GTT	AAA	GTG	AAC	GTT	ACT	GAA	GAG	AAA	GAA	AAC	CAT	GGT	GTT	TTT	AAC	CTA	4140
1361	K	F	D	V	K	V	N	V	T	E	E	K	E	N	H	G	V	F	N	L	1380
4141	CCA	CCA	AAA	AAA	AAA	AAA	TTC	GTC	TTG	CTT	CAA	GAT	GAC	CTT	ATT	AGA	GGA	ATA	AAA	GGT	4200
1381	P	P	K	K	K	K	F	V	L	L	Q	D	D	L	I	R	G	I	K	G	1400
4201	GCT	TTA	GAT	GCC	CCT	CCT	GAT	GAG	ATA	ATT	GCT	GAA	GTC	GAT	CCT	CGT	AAT	AGG	CTC	CAA	4260
1401	A	L	D	A	P	P	D	E	I	I	A	E	V	D	P	R	N	R	L	Q	1420

FIGURA 30 – Sequências de cDNA e de aminoácidos deduzida de Lox-C3 (Parte 4)

```

4261 CGT TCA TCG CGA ATT GGG TTC CAT TCA GCA GCA GTT TTA TCA TCA CGT GTT AAG AGG GAG 4320
1421 R S S R I G F H S A A V L S S R V K R E 1440

4321 GAC AAA ATG GAT AAT AGC AAC TTA GTG TTG GCA ATA AAG ATT TGT GCA AGA TAT TTA GGC 4380
1441 D K M D N S N L V L A I K I C A R Y L G 1460

4381 GAC CAT GAA AAC GGC ATG ACG ATT GTT GAT GCT GGT ATC TTC TCT GGT TTC AAA GTG AAG 4440
1461 D H E N G M T I V D A G I F S G F K V K 1480

4441 GAA GAT GAT TTG AAA AAG CTT GTC AAG AGT GCG GGA CAT TTT ATT GAC AGG TAC GAG TTG 4500
1481 E D D L K K L V K S A G H F I D R Y E L 1500

4501 TCA AAA AGA GGA ATC GTG TTT TAT TTA GGA AAG GTG CCC AAG GAA GAG CCC TAC TGC TTT 4560
1501 S K R G I V F Y L G K V P K E E P Y C F 1520

4561 CAA TTT CGC ATC ATT AGA GAC TAT ATT GTT GGA AAC ACT CAA TCA ACT ATC ATC AAA GTA 4620
1521 Q F R I I R D Y I V G N T Q S T I I K V 1540

4621 TTT TCC TAC TAT GAT CCA GAT GAA TCT TGC ACA GTT TTC TAC AGT CCT GAT CAT AAC AGT 4680
1541 F S Y Y D P D E S C T V F Y S P D H N S 1560

4681 CCG ATG CTG CGC AAG ATA TGT AGT GGA GGT ATT TGT CAA TGC GCT GAA GGT GGT TGT CCC 4740
1561 P M L R K I C S G G I C Q C A E G G C P 1580

4741 CCC AAA AAA CCA TTT GCA GTT ATC CCT AAG GGG CGA CAA GGT TTA AAG CAC TCT GCT TGT 4800
1581 P K K P F A V I P K G R Q G L K H S A C 1600

4801 GAA GAA CAC AAC TAT GTA TGG CTT GTG AAA GTA GAA GGA CCA CTT GAA CGA AAG GAT GGA 4860
1601 E E H N Y V W L V K V E G P L E R K D G 1620

4861 TTT TTG AGG ATT AAA TTG CGA GTC CTG GAA GTA GTT AAA GAA GGC ATT GAG AAA GCA ACT 4920
1621 F L R I K L R V L E V V K E G I E K A T 1640

4921 CTT GTA GAA GAA AAA ATG CGC CTA TTT TTG GCC AGA GAC AAC TGC GAA AGC GCG GAA CTC 4980
1641 L V E E K M R L F L A R D N C E S A E L 1660

4981 TTT CCG GGT AAA GAG TAT ATG ATC ATG GGG AAA GAT GGT AAA AAG TAC GAA AAC GAA AAA 5040
1661 F P G K E Y M I M G K D G K K Y E N E K 1680

5041 GGA GAA CTA TGG CAC CGA TAT ATG TTG GAC AAG ACA TCA AAA ATA TAT TTA TGG ACT CCC 5100
1681 G E L W H R Y M L D K T S K I Y L W T P 1700

5101 ATT CGC GAG GCC AAG AAT CAG AAC ATG ATG AGA GCC TTG CAT TGG CTA AGA GCG GAA TTG 5160
1701 I R E A K N Q N M M R A L H W L R A E L 1720

5161 AAG AAA GGC TGT GAT CTG TAG 5175
1721 K K G C D L * 1726

```

A sequência do cDNA foi obtida a partir da junção dos clones C3 5' e C3 3' e a sequência de aminoácidos foi deduzida usando o programa Bioedit versão 7.0.9.0.

FIGURA 31 - Sequência deduzida de aminoácidos de Lox-C3

A

10	20	30	40	50	60	70	80
MRASVLTLLV	ATILSSAKAW	NVYVAAPSHI	RVGVEEKVAI	VVEGNAAPRN	FEVYVQDHPG	KIKLSSTFG	TVFSNPKVV
90	100	110	120	130	140	150	160
SLSINPSDVP	DFIPGDPSTL	SFVSLIVHGG	SFHKEVVLPL	SYKTGYVFIQ	TDKPIYTKKE	QVKFRIIPLN	EDARPSAEKI
170	180	190	200	210	220	230	240
KLRINNPRGL	TVQDMEFSGK	SFLTPEPYTF	PSYPILGVWS	ATVLHGYKFQ	QNTTVQFELQ	EYVLPTFTVK	LKVQEVILPS
250	260	270	280	290	300	310	320
ATTIDISVES	QFVYGKPVSG	TVIFRIGIKG	ETPEVTYFAI	LDSKKLVDGS	YKTRINTEQL	KRDKSIGWFP	EIEGSRFVVE
330	340	350	360	370	380	390	400
ATVTDDATGQ	KEVVIDSNSY	FSTSPFVISF	KRCLSDFKPG	LISIEADVNI	YVDGRPASKI	KTITISAVADN	GVPLKITNPV
410	420	430	440	450	460	470	480
SVSNEEGKIL	FEIQPELGHS	VVTVTVTTS	PKYAGSQAVG	RHIQTKFESK	NAAYIAIARD	SKDKPKPGTM	FTEIVYMESE
490	500	510	520	530	540	550	560
DMTAVYYMVV	SRGRIALIGK	LQGRFKTYRV	EFMVTHEMVP	NFRLVVFSEY	KNELIADSLR	IDVANECDPK	ADVKLTPSFN
570	580	590	600	610	620	630	640
IKEPGSNGKI	LIEGSKSTVV	GLLAVDSAVY	ALSKRDKLTR	AKMFKTLEKH	DLGCGPGGGL	TTASVLRNAG	VVMMTRNYSF
650	660	670	680	690	700	710	720
TNDANSRVSA	CEEKVRKKRD	VLDTKLKEYD	GILRQCCALG	MNHDYRSRT	QQRFEIVKMH	LKDEDRNCSQ	TFLDCCMIAE
730	740	750	760	770	780	790	800
KHGATGLTLI	QPQKFQSARM	GAFGEDIQFI	TEEEEEAFES	VVSIRSDFRE	TWLFDDVIIG	PDGKHEFPVS	LPQSITTWII
810	820	830	840	850	860	870	880
QAVSISQEHG	MCVATPESVE	SFKKVVFQNL	LPYSVVKNEQ	IEVRATVFN	GDKNILVVVY	MYGVKDLCTG	TEEGKKSERK
890	900	910	920	930	940	950	960
RVLVHSNSAQ	SVSFPIIPLA	EGVFDIQVTA	LSPHGSVDII	RKLNVAEAGK	TVEEDITIKL	DPSNQRRKR	RHVTNDYSD
970	980	990	1000	1010	1020	1030	1040
SIDNVQKVQI	TAIKLKAPED	FIPGTESC	TAIGDQFGPT	VQTSLENPEN	LIDSPRC	QAMIYLAAPT	YTMKYLKATN
1050	1060	1070	1080	1090	1100	1110	1120
KITADTETRG	YDFIRKGYDL	ILTYRKNDGS	YGAWAHTPQS	TWLTAFIIKV	FCQASKMVHI	DEDVMCSGVK	WLVSNQQADG
1130	1140	1150	1160	1170	1180	1190	1200
SFIEHNPVYH	TEMMGGVKQG	TPLTAFVLIT	LQECSC	IEIPN	IHIVKARALS	YLEGHLGDIS	EPLPLAIAAY
1210	1220	1230	1240	1250	1260	1270	1280
QAINNKLLAL	SKYDADTNTM	YWETESSSYN	VEATGYGLLA	QLTLFEMARS	NSIVNWLNGQ	RMATGAFKST	QDTVIALQAL
1290	1300	1310	1320	1330	1340	1350	1360
SEYAIRAKTP	ALNLVSNLTS	SNDRNFMQTL	MFDQSNALIL	QEHKLDKLG	TLFVNNTAGNG	VGLLTVKLRY	NVPNPPEKYC
1370	1380	1390	1400	1410	1420	1430	1440
KFDVKVIVTE	EKENHGVFNL	PPKKKKFVLL	QDDLIRGIKGA	LDAPPDEII	AEVDPRNRLQ	RSSRIGFHS	AVLSSRVKRE
1450	1460	1470	1480	1490	1500	1510	1520
DKMDSNVLV	AIKICARYLG	DHENGMTIVD	AGIFSGFKVK	EDDLKKLVKS	AGHFIDRYEL	SKRGIVFYL	KVPKEEPYCF
1530	1540	1550	1560	1570	1580	1590	1600
QFRIIRDYIV	GNTQSTIIKV	FSYDPDES	TVFYSPDHNS	PMLRKICSGG	ICQCAEGGCP	PKKPFVAVIPK	GRQGLKHSAC
1610	1620	1630	1640	1650	1660	1670	1680
EEHNYVWLK	VEGPLERKDG	FLRIKLRVLE	VVKEGIEKAT	LVEEKMRFL	ARDNCSAEL	FPGKEYMIMG	KDGKKYENEK
1690	1700	1710	1720				
GELWHRYMLD	KTSKIYLWTP	IREAKNQNM	RALHWLRAEL	KKGC	DL		

Ponto isoeletrico 7.99/ MW: 192990.55 Da

B

[A] O peptídeo sinal está sublinhado e compreende os primeiros 20 aminoácidos. Os sítios de glicosilação e as cisteínas estão marcados em azul e vermelho, respectivamente. As predições para detectar o peptídeo sinal e os sítios de glicosilação foram realizadas usando a ferramenta GPMW disponível no endereço eletrônico <http://www.expasy.org/proteomics>. [B] Arquitetura de domínios da proteína Lox-C3 foi determinada com a ferramenta BLAST (Blast Local Alignment Tool).

4.7 Alinhamento múltiplo de Lox-C3 com outras proteínas C3/C4/C5

O alinhamento produzido entre a sequência do C3-like de *L. laeta* com proteínas do complemento de invertebrados e vertebrados mostrou certa similaridade, com conservação dos seis resíduos de cisteína críticos para a estrutura peptídica do domínio ANATO. Além disso, o alinhamento revelou dois sítios de clivagem putativos, um presente entre as cadeias α e β e outro localizado entre as cadeias α e γ . Os aminoácidos que compõem a região de ligação das cadeias α e β , posicionadas anteriormente à região ANATO, se alinharam perfeitamente, sendo que os mesmos quatro aminoácidos RKKR (676-679), encontrados na sequência do C4 humano foram também observados na sequência Lox-C3 (656-659). Diferentemente de C4, o C3 humano é constituído somente por duas cadeias: β (1-645) e α (650-1641), não possuindo a região de junção, presente no C4 humano, entre a cadeia α e γ , cujo sítio de clivagem inclui quatro aminoácidos arginina (RRRR, posição 1451-1454). O C4 humano, assim como, provavelmente, os componentes C3-like de *H. adansoni*, *T. tridentatus* e *R. decussatus*, apresentam três cadeias, α , β e γ e o sítio de clivagem putativo entre as cadeias α e γ é representado pelos aminoácidos RQKR, RVKR e RKKR, respectivamente. A mesma sequência de aminoácidos encontrada em *T. tridentatus* RVKR foi detectada na sequência Lox-C3 (posição 1436-1439), embora não esteja perfeitamente alinhada, como é observado para a sequência RQKR da aranha *H. adansoni* (**FIGURA 32**).

Embora ainda não tenha sido mostrado na literatura que a properdina se ligue aos componentes C3-like de límulo, foi possível também identificar, nestas sequências, similaridade com os aminoácidos que compõem a região ligante à properdina do C3 humano. A região de ligação ao fator B, fator H, CR1 e CR3 se encontra alinhada com a presença de quatro aminoácidos idênticos.

Os motivos tioéster de todas as onze sequências se alinharam perfeitamente, mantendo-se conservados, com exceção do aminoácido glicina (G) que foi substituído pelo aminoácido alanina (A) na sequência de Lox-C3. Além disso, o aminoácido Histidina (H) que compõe o sítio catalítico se manteve conservado em todos eles, com exceção do componente C3-like isoforma 2 da aranha *H. adansoni* (Hd-C32).

Para verificar a disposição da estrutura putativa de Lox-C3, sua respectiva sequência de aminoácidos foi submetida à modelagem computacional tendo como base a estrutura

cristalina que do C3 bovino (PDB 2b392.A) que apresentou identidade de 27,7%. Como a estrutura do C3 humano é bem mais documentada na literatura e por também apresentar identidade em torno de 27%, as posteriores análises comparativas foram realizadas com base em sua estrutura (PDB 2473.A). A **FIGURA 33** mostra que, apesar do baixo valor de identidade entre as duas proteínas, é possível observar uma notável similaridade estrutural entre elas. As regiões de α -hélice e de folha β pregueadas estão dispostas de forma similar nas duas proteínas, sendo que diferenças podem ser observadas, principalmente, na região dos *loops*. Todos os oito domínios macroglobulina, que formam o arcabouço da estrutura do componente C3 humano, também são encontrados e dispostos nas mesmas posições no modelo molecular de Lox-C3. A região do domínio ANATO e do tioéster também se encontram nas mesmas posições, indicando que Lox-C3 possa desempenhar atividades biológicas semelhantes ao do C3 humano.

FIGURA 32 – Alinhamento Múltiplo de Lox-C3 (Parte 1)

	/MG1														
Lox-C3	---	MRASVLTLLVATILSSAK---	AWNVY	VAA	PSHIRVG	VEEKVAIVVE	GNA---	APRN	FE--	VYVQGH	P	GKIKNLSST	FG-TV-FSNS	76	
Hd-C31	--	MDMWSCAVFLLVGVFGHTC---	AQQIY	VVA	PNTRLRN	SDETI	IAVAID	GN---	IGAV	VS--	VFVQDH	P	GKVKNISQT	LV-AV-QPGQ	77
Hd-C32	---	MVLSLHLVFLIFVLIGKIT---	AQNY	VAV	NSLRLN	SDEFIGVAAD	GN---	QGGT	VN--	IYVQDY	P	GKIKNITIT	TD-QI-DPGT	76	
Tt-C3	--	MESRILLFTLTIVFLCIRIE---	AANIF	VAS	PNTRLRIG	EQETVSVLLE	GN---	KAET	VE--	VYLDQH	P	GKTKTFSRT	VG-RA-EPRI	77	
Bb-C3	--	MIRSLVCVCLLFAAVLQQ---	GPAYF	ISV	PNTLLKVG	TEETVSVNVF	NVA---	NPVR	VK--	VYLDQY	P	DRKTTFSEA	EV-DV-NQDE	77	
Nv-C3	----	MLRCVVLMLAIGCSYA---	VQRYM	VIS	PVDFVHG	EPERVSI	AVY	DVP---	SPVT	VK--	VYLDQY	P	NRRTKFSQA	QA-AF-AKDN	74
Aj-C3	---	MFLWICCLVLVRVQLTAAQ---	PSHIF	VSA	PNIRAG	NIETVLVSF	DT---	EEVP	VT--	VTLKN-	-	GDGSNICP	VLQV-NPDN	75	
Rd-C3	----	MLAILLLFGILVSEVQ---	CGYYF	VTAP	NVLRFD	QDETVVVSF	GL---	NNAA	VK--	VWLEH-	--	ENKQFSVK	NV-LVKDEKH	71	
Sp-C3	MG	SIILIFVLAIASLVPGGVTDPPATFF	VIS	PNVFRAG	VEEKVVVTLI	RSPNVQNI	E	VK--	VSLMR-	P	GSTVTFSED	TR-LV-SPGS	85		
Hu-C3	MG	TSGPSGLLLLLTHLPLALG---	SPMYS	IIT	PNILRL	SEETMVEAH	DAQ---	GDVP	VT--	VTVHDF	P	GKKVLVSSE	KT-VL-TPAT	80	
Hu-C4B	-	MRLWLGIWASSFFTL--	SLQ--	KPRLL	LFS	PSVHLG	VPLSVGVQLQ	DVP---	RGQV	VKGS	VFLRNP	SRNNVPCSPK	VDFTLSSERD	81	

															/MG2
Lox-C3	PKVVTLSINPSDVPDFIPGDPSTLSFVSLIVHG----	GS-	----	FHKEVV	LPLSYKT	GYV	FIQTDKPIYT	KKEQV	KFRII	PLNEDAR	PSA	157			
Hd-C31	PELFKIKQLNSQNFPPNFLSGPGFPKYVSLTAVF----	PD-	----	FRKELI	IPVSNQSGYI	FIQTDKPIYT	PKERASIRII	PLNEDSR	RE	158					
Hd-C32	PKLFKVRLDTENFPYGLSSPGFPKYVRLTAHF----	QG-	----	YKELL	IPINEKYGHM	FIQTDKPIYT	PKDRVLFRRM	HLNESAL	PTD	157					
Tt-C3	STEFKVQVNPEDLPDKDILAATAKHVSVLVAKA----	GN-	----	WFHKE	LLVNPRSGYV	FIQTDKPIYT	PKQTVHMRRI	TLNEDLIP	QN	159					
Bb-C3	PSLVTVRVNPDLNPE-----	SRA	T	KRYVYVAKS----	DDP	QLT-FQKEAQ	VLLSYQCGYV	FIQTDKPIYT	PNQKVVMRIM	PLDQDMT	PAS	158			
Nv-C3	PGYLVTKVDPADLDP-----	QGS	S	QDKQYVYLVAQC----	DDP	KL-KFKKKEV	VLLSYRDGVI	FIQTDKPIYT	PDQTVKFRVI	PLGFDLKE	PSA	156			
Aj-C3	PQIVSLYVDPENVPQ-VETEDDLKVYAKLEVTC-----	P	PL	NNFQ	EEKD	VLISYEE	CEV	FIQTDKPIYT	PDQRVNIRLL	VLDQY	LRPS	158			
Rd-C3	PENIFVRVTENDTVSTLTTEEKTRK--	VKL	CAEW----	NG-	----	QKQTR	IILSYHSGYL	FIQTDKPIYT	PKEKKVIRTL	ALDESLK	AVD	150			
Sp-C3	SQSFSVVIQAGDLMP-----	NEG	A	FQH-MVLKAES----	LNP	IY-PFEEQTD	ILVTLQSGYV	FIQTDKPIYT	PNQDVMIKVM	SLDQML	PSD	166			
Hu-C3	NHMGNVTFITIPANRE-FKSEKGRNKFVTVQATF----	GT-	----	QVVEK	VVLVSLQSGYL	FIQTDKPIYT	PGSTVLYRIF	TVNHKL	LPV	160					
Hu-C4B	FALLSLQVPLKDAKSCGLHQLLRGPEVQLVAHSPWLKDSL	SRT	TNI	QGIN	LLFSSRR	CHL	FIQTDQPIYN	PGQRVRYRVF	ALDQKMR	EST	171				

																	/MG3																	
Lox-C3	E-KIKLRINPRGLTVQDMEFS--GKSF-----LTPEPY	TFPSYPIIGV	WSATVLHG	YK	FQ	QNTTVQ	FE	LOEYVLPFT	T	VKLK--VQE-	235																							
Hd-C31	E-RFLQIRNPKNIIVEEKYFN--DKNR-KLDKAFASHVY	REPTYFVIGE	WSATVRYGHD	LEQ	NATVR	FE	LEBYVLPFT	T	VELK--VPD-	241																								
Hd-C32	E-TFQIQRNPNNSIVDQLFD--GKRKSELKKVSVTHVY	KIPTYFVIGE	WSATVQYGYN	LK	QNTTVH	FE	VOEYVLPMT	T	IQLE--TPE-	241																								
Tt-C3	K-MITLQIKNPQDITVEHQNWKV-GKVGRRIE--FYTYDY	SFPYPLFGE	WSAVVSYGYD	LL	YNTTVK	FE	VKEYVLPFS	S	VDVT--APE-	242																								
Bb-C3	Q-PVKLEILNPQGIIVERKTFFG-SATG-----FIATF	DEPAFLBGN	WTATAHGYPE	MQ	LNVS	TQ	FE	VKEYVLPFT	Y	VRII--PSNP	238																							
Nv-C3	K-KVNVVVKNPNIRVQWNSMS-TSTG-----FISETI	SLGSFAMIGN	WTVEAFYGHG	FV	HNTSTQ	FE	VREYVLPFY	S	VKIQRHPLKN	238																								
Aj-C3	E-PVQLDVLSPQGVIVHRSAIR-ARNG-----FKTRVF	SFPHBLFBN	WSVVAQYKKG	LK	SNRSV	Q	FE	VKEYVLPFE	S	VKLT--TPG-	237																							
Rd-C3	GWQVGMDIVSPSNKTLGRKLIG-SPSG-----FYQND	TIPPPYELQL	WSARAFYKQG	FET	ESHAL	FE	VREYVLPFT	S	VTID--VDVE	231																								
Sp-C3	R-ELHVEIMDPGTSVKRLRETH-ADVP-----ASIGRL	QVHNHEASVN	GPFLRSLMDPS	-	SLSSQ	VT	FE	VKEYVLPFS	S	VSIE--TPK-	244																							
Hu-C3	R-TVMVNIENPEGIFVKQDLSLSSQNLG-----VLPLSW	DIPELVNM	Q	WKIRAY	YENS	PQ	QVFST	FE	FE	VKEYVLPFS	S	VIVEP-TEK	242																					
Hu-C4B	D-TITVMVBNSHGLVRKKEVY--MPSS-----IFQDDF	VIPDISEPT	WKISARFSDG	LES	NSSTQ	FE	VKKYVLPNF	S	VKIT--PGKP	250																								

Lox-C3	VILP---	SATTIDISVESQFVY	GKPVSGTVIFRIGIKGET	-PEV	TYFAIL	DSKKLVD---	---	GSYKTRI	NT-EQLKR--	DKSIG-WFPE	311						
Hd-C31	VVLP---	KDETIQIEVKANYVY	GKVKGI	VTFRLGKVG	GDH	SPDAVFFFAVI	TPKELDD---	---	GSYVLRI	KT-DDLLR--	HKDIE-WFPE	318					
Hd-C32	IILP---	KDEEISIKVSARYVY	GKVKGMVSFRLC	IKKPT	DSETKFVAII	GSRELTD---	---	GFYNLRL	NV-QELLE--	KSNI	SFAE	318					
Tt-C3	IILE---	STKKITGSAHAKYVY	GEPVHG	TANFKFVGK	LMT	-	GDIIII	IGTL	YNKQLQD---	---	GKVDYQI	DP-NEFMK--	HGKIS-GFPD	318			
Bb-C3	YILP---	QDDVTSGEVEALTY	GKGV	VDGFLDKV	KFI	DL	-	GNRQLFAQL	PT-EVNG---	---	GFGEYEI	KT-ARIK--	DLDL--WFFP-	310			
Nv-C3	YLIR---	KTHYPI	SISARYTY	DEKPEV	PNVRLSVGRK	-	QESKLMDV	TTI-LND---	---	GCSLNF	DP-AIKID-	LPDKNLWFFP-	313				
Aj-C3	YILD---	GQEA	VDITVSARYVY	DEKPV	LVGFYAKV	VLG	FD	-	GEIDIKVK	SGQMLST---	---	GSVEFRI	NM-EDLGE--	---	N-WFAE	309	
Rd-C3	YILP---	QTKHITVTVAKYVY	GKPVQGNARLT	TLRLKGD	-	KDTY	YI	IDL	KKQL	LENSPD	VEGSATEFRL	DVKKDIL	LDSS	LLRDK-PFP-	315		
Sp-C3	YILA---	GADSV	VTIIA	EYV	GKPVIG	RYTTFKYGVVDN-	-	KTQ	DEL	Q	GLGDS---	---	GRAEVTL	---	DDLST--	DFGED-WFVR	317
Hu-C3	YYIY---	NEKGL	EVTTI	TARFLY	GKVEG	TAFVIF	GIQDGE	-	QRIS	LPESL	KRIPIED---	---	GSGE	VVL	SR-KVLLD--	GVQNP-RAED	317
Hu-C4B	YILT	VPGLHDEMQLDI	QARYIY	GKPVQGVAYV	RFL	LLDED	GKKT	TF	FRGLE	SQTKLVN---	---	GSHIS	L	SK-AEFQD-A	LEKLN	MGITD	332

/MG4																																																																																									
Lox-C3	E	G	S	R	F	V	V	B	A	T	V	D	D	A	T	G	Q	K	E	V	V	I	D	S	N	S	Y	F	S	M	S	P	F	V	I	S	F	K	R	L	S	D	F	K	P	G	L	I	S	I	E	A	D	V	N	Y	V	D	G	P	A	S	K	I	K	T	T	I	S	A	V	A	--	D	N	G	V	P	L	K	I	T	N	P	398				
Hd-C31	E	G	S	H	L	V	V	E	B	A	T	V	D	A	A	T	G	K	N	E	T	T	I	D	S	K	G	R	F	S	K	T	P	F	L	I	S	F	K	R	L	K	D	F	K	P	G	L	I	T	S	V	E	A	D	V	T	Y	V	D	G	S	P	A	G	G	V	R	T	K	I	Q	A	V	A	--	D	N	G	K	K	L	F	I	A	V	S	406	
Hd-C32	Q	N	S	H	L	I	V	E	B	A	T	V	D	I	A	T	G	N	E	T	T	I	D	S	K	R	F	S	M	S	P	F	S	I	S	F	K	R	L	S	D	F	K	P	G	L	I	T	S	V	E	A	D	V	N	Y	V	D	G	P	A	E	G	I	T	M	I	R	A	N	--	G	N	G	Q	A	I	E	I	K	D	406							
Tt-C3	E	G	N	H	L	F	V	E	V	S	L	E	Q	A	T	G	Q	R	E	T	G	I	N	E	N	G	I	F	T	L	S	P	Y	D	V	S	F	K	R	L	N	N	F	M	B	G	Y	T	T	L	P	V	D	I	N	F	V	S	K	P	A	G	I	P	A	V	I	T	V	D	--	E	K	G	R	A	V	V	Q	K	S	406							
Bb-C3	E	G	S	R	L	Y	L	E	B	A	T	V	E	E	A	T	G	A	G	L	R	E	M	A	V	L	T	S	V	R	E	T	S	P	F	R	I	G	Y	D	T	A	T	H	F	K	P	G	L	I	T	S	V	E	A	D	V	N	Y	V	D	G	P	A	S	K	I	K	T	T	I	S	A	V	--	D	N	G	V	P	L	K	I	T	N	P	399		
Nv-C3	D	G	R	F	L	Q	I	E	A	T	V	I	E	K	A	N	G	L	R	E	S	V	T	D	N	S	I	H	F	T	T	T	P	F	V	I	K	Y	Q	E	T	A	E	Y	F	R	P	G	L	I	T	S	V	E	A	D	V	N	Y	V	D	G	P	A	S	K	I	K	T	T	I	S	A	V	--	D	N	G	V	P	L	K	I	T	N	P	399		
Aj-C3	E	G	R	L	L	F	V	E	V	S	L	E	S	A	G	L	S	E	N	V	S	D	I	S	T	K	F	I	F	V	S	P	Y	Q	F	K	W	D	R	T	V	H	W	F	K	P	G	L	I	T	S	V	E	A	D	V	N	Y	V	D	G	P	A	S	K	I	K	T	T	I	S	A	V	--	D	N	G	V	P	L	K	I	T	N	P	399			
Rd-C3	N	G	K	R	L	E	V	I	A	T	V	Y	E	S	A	T	G	N	E	E	A	A	S	H	D	G	T	I	F	T	E	S	P	F	I	F	K	F	T	K	S	K	L	N	F	R	E	P	G	Y	K	Y	L	K	V	E	L	F	Y	V	N	G	K	P	A	K	D	T	D	V	E	V	S	M	F	E	--	D	G	V	L	K	E	E	K	A	G	--	401
Sp-C3	R	G	R	H	F	H	A	N	A	T	V	H	E	S	A	T	G	Q	F	S	E	A	S	I	N	T	K	A	L	F	V	D	S	P	Y	F	K	S	T	R	T	V	Q	H	F	K	P	G	L	I	T	S	V	E	A	D	V	N	Y	V	D	G	P	A	S	K	I	K	T	T	I	S	A	V	--	D	N	G	V	P	L	K	I	T	N	P	399		
Hu-C3	V	G	K	S	L	Y	V	S	A	T	V	I	L	H	S	G	S	D	M	V	Q	A	E	R	S	G	I	P	I	V	T	S	P	Y	Q	I	H	T	K	T	P	K	Y	F	K	P	G	L	I	T	S	V	E	A	D	V	N	Y	V	D	G	P	A	S	K	I	K	T	T	I	S	A	V	--	D	N	G	V	P	L	K	I	T	N	P	399			
Hu-C4B	Q	G	L	R	L	Y	V	A	A	I	E	S	P	G	C	E	M	E	A	E	L	T	S	W	Y	F	I	S	V	S	P	F	SL	DL	S	K	T	R	H	L	V	E	G	A	P	F	L	L	Q	A	L	V	R	E	M	S	S	P	A	S	G	I	P	V	K	S	A	T	V	S	S	P	G	S	V	P	E	V	Q	D	I	Q	422						

																/MG5
Lox-C3	VS-----	VSNEEGKILFEIQPELGH	SVTVTVTTS	DPKY-----	A	GSQAVGRHIQ	TKFESKNAAY	IAIAR-DSKD	----	KPKPG	468					
Hd-C31	SA-----	QSDENKGVSFIEIQPELHHK	SVTITLET	DPRY-----	E	GSQAKGIFEQ	HAYQSGQDAY	IAIAR-SSPQ	----	KLKPG	475					
Hd-C32	SA-----	ISDEKGVKSFEVRQLH	HNHEILVTLIV	DKEY-----	----	-----LSEASFQK	HAFQSSATNSY	IALDR-MSQ-	----	TLKAN	471					
Tt-C3	ED-----	TSDEYRCNFLVNP	SKDLKQIKVEVKTN	DGRG-----	A	QYQRKGQHRM	SEQSSEFGGV	IATDRGTTKQ	----	DLKVN	476					
Bb-C3	NNEHNSD----	TTNQYQASFTVDP	PPGTQTLTVTAKTE	QVGL-----	PV	AHQAEQNFEA	TPYQSPSGY	LLVRV-LQRG	----	PVPVD	474					
Nv-C3	ANAA	NMD-----	QTGPEGKAEFIV	IPDRVATL	RVKVRTE	EKDV----	QP	EENAVKELIA	RAYTSQSGDF	LMVRQ-LR--	----	KATVG	475			
Aj-C3	QRQNGENR	PTLSNTNNRQ	VNFRIYVPSDCQ	TIQIQLQ	TQ	DQRLINADGT	NNNAEQFVVV	QYISPTSDS	FLLIR-LTQE	----	LVNVD	483				
Rd-C3	-----	VTNDDGRI	IQQSTAAK	KAKKISFKVSTS	NNEF-----	-----	ESDEFV	TAYPGKNQM	VE-----	----	YVTTD	457				
Sp-C3	DIAGGQE-	EISNTDDQ	GVSMVLNVQSSITS	IHIVAATH	DPQY-----	P	NNQADISFEV	SPGVSANGHD	YLVR-PQNT	DDRN	LNVDVS	489				
Hu-C3	-----	LTQGDV	AKLSINTHP	SQKPLSITVR	T	KQEL-----	SE	AEQATRTMQA	LPYSTVGN	SN	LYLHLSVLR	T	-----	ELRPG	468	
Hu-C4B	Q-----	NTDGSQ	VSPII	PIPTQISELQSVS	AG	SPHP-----	----	AIARLT	AAPPSSGGPGF	LSIERPDSR-	----	PPRVG	486			

FIGURA 32 – Alinhamento Múltiplo de Lox-C3 (Parte 2)

																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					</
--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

Forma ponte dissulfeto com cadeia α

[illegible]

Sítio de processamento α/β

[illegible]

Sítio de Clivagem putativo para

		*	**	C3 convertase	↓	Sítio de ligação FH e FB, CR1 e CR3				
Iox-C3	EIVKMH-----KDEDRNC	SQTFDLC	MAIAEKHGA	TGLTLTLPQK	FQSARMGAFG	EDIQFI---	EEEEEAFAFVS	VSIRSFDFRET	771	
Hd-C31	NVVQKYL-----DGEYPTCA	KAFAMC	CCLYGLRNL	ADMVKVPGQ	ALGRMGGBNL	EEPEFI---	AEFEEDFEKQ	LTVRKDFRET	774	
Hd-C32	NAVEHYL-----IDEEYINC	STIFLD	CCLYKQRIEL	DSWRNEKMRS	YMTVGKAFED	DDIGLI-DID	PELEDDMERK	STRRTYFEET	776	
Tt-C3	AIFEKY-----DFDGSC	MAAFILGCC	-----N	EKHLYLRLNI	EKEGRGRFGF	DPADQLITVG	LEEEQELLKQ	LVNRKDFRET	777	
Bb-C3	KL-----NSTSQDBCY	FAFLTCC	-----NHA	RSLLRLRGR	GRMGGG-GLL	LDIDID---	-----EDESQ	LVARTEFEET	762	
Nv-C3	LQDFDLTDKV---SPWIRPTEL	KQCYKTFIYCC	-----	-QARYGKSDM	GLLARS-IDF	GDFGPS---	---DEELMEN	SQVRTYFEET	769	
Aj-C3	NRFVNLLRRKYE	EDGTIDERVELIA	FMICC	-----QGA	EBAR---NDR	GRSGVS-NPL	NDILDI---	-----EDRPD	AAVRSDFRET	788
Rd-C3	LILA-----EENRLSK	CVMAFFKSC	-----V	DESEYMITTE	TKIVSK-SLD	DESEYL-ED	IAKLADNGII	SKTRSDFRS	766	
Sp-C3	RVLE-----AECNLEPD	QTQVYRDCC	-----I	NSMNPVVTAT	SRSGGD-GGE	QNAA---	-----	VKVRDDFRET	784	
Hu-C3	RFIS-----LGEACKV	FLLDCC	-----NYI	TELRQRHARA	SHLGLA-RSN	LDEBII---	-----AEEN	IYVRSDFRET	770	
Hu-C4B	ARVQ-----QPDCE	REFPFLSCC	-----QFA	ESLRKRSQRK	QAGLIQ-LAL	EILQEE---	-----DLIDEDD	IPVRSDFRET	781	

Sítio de ligação FH e FB, CR1 e CR3

Sítio de ligação FH e FB, CR1 e CR3

Forma ponte dissulfeto com cadeia β

																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					</
--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

 [/CUB](#)

Lxx-C3	KNILVVVYMYGVKDL	IGTGE	GGKSERKRVLVHSNAQSV	SFPFI	ELAEG	VFDIQVTA--	--LSPHG-SD	VIIRKLNVA	EGKTVEEDIT	937
Hd-C3	TKIGAVVYMYGAGDL	SGTQAGE	SEKSRKLIVTEGQSAATV	TFPFI	ELKAE	DFVIKVA--	--LTPAG-SD	VVQRTLHVA	EGVTKIEDIP	940
Hd-C32	RP1RAVVYMYGTGL	SGSTPGQK	SERQYIVTEGQSAATV	TFPVM	LKAE	VIPKVA--	--LSSLG-SD	VIIVRELNVVP	EGITRFINIP	942
Tt-C3	QSVRASVYMYGVKGL	SGAAGEQ	RTERKVLIDIGSLAKSV	SFPVV	LKVGV	EFPVIRVV--	--FTVHG-SD	FIEKSLNVVP	EGAKDSKLSI	944
Bb-C3	QDLRVNVYMQGVEG	SGRAGER	SERKLTIFKGNDAASV	LFPII	PLEVG	TFPIRVVA--	--FSTAAGD	IIIEKLVQIP	EGVERRLVS	929
Nv-C3	QQLRMNLYLVGHEKF	CSKVKPRAR	QLIERFTMKAQSSQTV	SILII	EKEIG	LIPIQVFA--	--VSALE-SD	SESRNLNVVP	EGVGQIKTQS	935
Aj-C3	QDQQVRVFLRTKEV	SGENPNYS	AEAKRIEVTQNNAFSI	TYVVI	GLEVG	EHPIITVAA--	--VYSGT-SD	AVEKILRVVP	EGIERFSVSS	954
Rd-C3	AAKTANVYLTKGNDL	GYNSDPK	PSPRVQVLPNNSAKV	SFPMI	LKDGL	FLPVTVSAIV	TDLGIPPE-VD	VIIEKRLVVN	EGIEEKLTVI	939
Sp-C3	SDFEVSVNFTVDQGL	CTAENP---	AIRHV	FVESKRAAST	TFHCT	PCGSR	GIPYTVTA--	--GGSKK-RD	SVRNNLRVVS	947
Hu-C3	QELKVRLVELLHNPAF	SLATTKRR--	HAQTVI	LPKKSLS	PYVVI	LELKTG	LQVEVVKAA--	--VYHFI-SD	GVRSKLVQIP	943
Hu-C4B	KNLTVSVHVS	PVEGLAGGG--	LAQTVI	LPAGSARPV	AFSVV	ETAAI	AVSLKVVARG	SFEFPVG--D	AVSKVLQIEK	945

FIGURA 32 – Alinhamento Múltiplo de Lox-C3 (Parte 4)

Sítio de clivagem α

Lox-C3	LIRGVKGALDAPPDEIIVEVDPRNRLQRSSRIGFHSAAVLSSRVKREDKM	DNSNLVLAIK	*ICARVYL-GDH	ENGMTIVVAG	IFSGFKVKED	1482
Hd-C31	LIRGIKKLEDEVPPELAIDPRRRGHVRAS---LIASPLIKKDKRDKPN	ENSKVKLNIT	ICVRYL-GNK	DTEMSIVVAG	IFSGFQPVED	1488
Hd-C32	-----TKYEEIIDVRSHLNSGSD-----IPLTESR	PSSKLYRIS	ICVRYL-GDA	DAAMSIVVAG	IFSGFQPVKE	1442
Tt-C3	LLRNL-GFPKKRERSLDNIQDYARKVKNRGRVVDGG-----KRGND	GKSKLLLEIQ	ICVKYL-SHV	DSNMAIEAG	IFSGFKVLID	1484
Bb-C3	GQEGDEDYPDYDGLDARERSAFSSRSRIG--GLSSGSRSRRIARQAED	DQSQHLYLVR	VCTSRGQRG	ASNWALMIG	MFSGFEPVKQ	1486
Nv-C3	AMRRNGKKNNKKKHKHKKRNGKCRKPKKGGKGGKGGKKDKNEKQKQK	IQEVNSLRIN	TRTRYK-KAG	KTGMALIDVG	VFGFAVDER	1490
Aj-C3	---ALADNGDGPPEKPPNAGRGLDHARS-----RVGLRRDAQRDE	GRGAILNVR	VVAIVL-KSV	PTNMAIDVG	IYSGFKVLID	1494
Rd-C3	DICGKCEG-GDVNQYDDYDIDNFGNQPEIDPKMESNFGAKKFAANRRVRA	TATKKCIFRS	VSSK---NKK	KYGMSTVKFG	LETGVEVIKS	1482
Sp-C3	-----VPSDEVS	QGVNKGMLIT	VCTSRN-GDG	THHMLIDVG	LYSGFKAVEE	1447
Hu-C3	-----APETEKRP	QDAKNMILE	ICTRYR-GDQ	DATMSILDIS	MMTGFAPDTD	1416
Hu-C4B	TMEANEDYEDYEYDELPAKDDPDAPLQPVTLPLQFEGRRNRRRFBAPKV	EEQESRVHYT	VCIWRNGKVG	LSGMALIDVT	LLSGFHALRA	1500

Sítio de Ligação à Properdina MG/

Lox-C3	---DLKKLVQ-SAEHFIDRYEL---SKRG---IVFYLGKVPKEE---PY	CFQFRIIRDY	IVGNQOSTII	KVFSTYDPE	SCVIFYSP--	1556
Hd-C31	---DLIWLQD-DHSHLIQRYEK---SSRG---VVFYLGKVPVPIAG---DY	CFSFHVVRQY	IVGNQOSTVI	KVYDYNNPDA	TCTKFSYP--	1562
Hd-C32	---DLIKIED-EPSQLIKQIEL---SKRG---VIFYLTIKVPSTV---KY	CFHFVRLREY	IVGNQOMSYI	KVYDYNNPDA	TCTKFSYP--	1516
Tt-C3	---ELKQLVK-VKNSKIARFEA---SDKS---VVFYMDSAPHDK---PY	CFKFRIVRQF	IVGNQOSSVV	KVYDYNNPDA	SCVIFYSP--	1558
Bb-C3	---DLQQLNRAGKAVQRVET---NTRA---VLIWFDEITSE---EI	CVSIRMKRMV	IVGAPQVPVY	KVYDYNNPDA	ACTHIFYEP--	1560
Nv-C3	---SLKELQK-KMEGSIGKYEI---SERS---VVFYVINEIRNDR---DL	CVAFNVTRVF	EVGTQVQVPVY	KVYDYNNPDA	ACTHIFYEP--	1564
Aj-C3	---DLQGVVE---DNDSISREFI---SQRS---VIFYADFIPTQSQEPLE	EIQFNVKRRH	TVGNIKPVAV	KVYDYNNPDA	EOVKFYEPAS	1572
Rd-C3	---DLDKLIA-ENNTNRIARTEM---PSDGKGFIFYLHEITSET---NF	IFRLKDGFDG	EKNRQPSVY	KVYDYNNPDA	HCTQFYGT--	1560
Sp-C3	GLTGLKQVNVN---GSSLISSYEA---SSRS---VIFYLDITIPSE---DL	CFTFSAESDV	VYGNQAAAV	KVYDYNNPDA	SCVIFYEP--	1523
Hu-C3	---DLKQLAN-GVDRIYSKYEIDKAFSDRNT---LITVLDKVSHE---DD	CLAFKVHQYF	NVELQPGAV	KVYDYNNPDA	SCVIFYEP--	1495
Hu-C4B	---DLEKLTS-LSDRYVSHFET---EGPH---VLIWFDSVPTSR---E	CAGFEAVQEV	PVGLQVQASA	TLYDYNNPDA	RCSVIFYEP--	1572

/C345c/

Lox-C3	-D-HNSPMLRKIC--SGGICQCAEGG---C--PPKKPFVAI-----PK	GRQGLKHSAC	*EEH---NYVWL	VKVEGPLERK	DGFLRIKLR-	1627
Hd-C31	-G-SNSPMLRTIC--EGGICQCAEGG---C--PPNPFSEISSMS-TAQ	RREELKTIFC	ENY---DYVMK	GKLNGRLQKD	GGFLNISFT-	1638
Hd-C32	-I-SNSPVRTIC--EGGICQCAEGG---C--PPRTPFADLEGMQ-VSE	RNNLLVKVFC	EDF---DYVMK	GELH-YIRKE	GGFFNISFI-	1591
Tt-C3	-D-NQSPVIRTIC--EGSVQCAEGG---C--PPRHPFEGITQVHDISE	SRKLLLDRAV	VDH---DYVMK	GTVE-SKRKE	NGFRYISFER	1634
Bb-C3	-G-QGSPLLATLC--DGSQCVCAEGK---C--PKETSPKKL-----REL	TQRDLLEKAC	NDL---DYAFR	VIVM-EVRQE	GSFDRISMI-	1631
Nv-C3	-N-ENSALRMGIC--NEEECRQIQDK---C--TPRDPDTPG-----	---DLRSLVC	KNN-F-YAIA	AKVL-LVDDE	RSMLTYYTVQ-	1628
Aj-C3	VD-SRSLLDYIC--EGSVCLCAEGD---C--AATDSKPKD-----	---ELRLSCE	ANA-ATAYAQ	ISIM-NSQND	NGFKVYATAT	1639
Rd-C3	-GLNRGKSVSYQCGKDKKEVCFQSM---C-VQPVVEEIKKL-AKK-KKK	KALKLYKYAC	NTDKANYAVV	VTID-EVSYN	PAVQEKNARG	1642
Sp-C3	-G-DGSAALLSTLC--SENECICGSGSLEYCNIDPCPGPYTRI-----	---DLGEGTA	ASH-SSYALK	IRID-EVEIK	EGFRICKFT-	1595
Hu-C3	-E-KEDGKLNKLC--RDELICRAEEN---C--FIQKSDDKV-----	TLERLDKAC	EPG-VDYVYK	TRIV-KVQLS	NDFDEYIMA-	1564
Hu-C4B	-P-SKSRLLATLC--SAEVQCAEGG---C--PRQRALERG---LQD	EDGYRMKFAV	YYPVEYGFQ	VKVL-REDSR	AAFRLFETK-	1646

C345c /

Lox-C3	-VLEVVKIEGIEKATLVEEKMRLLFLARDNCSA-ELF---PGKBYMIMGKD	GKKYENEEK-G	ELWPRY-MID	KTSKIYLTTP	----IREAKN	1706
Hd-C31	-IADVIKAGIEKAEDIVDETRTLLARDNCPA-NLF---PGVTYLIIGKD	GQNYMNNEN-G	ETWYRY-LLD	KTSVIHMTTS	----AKEAKN	1717
Hd-C32	-VNEVIKAGIERKDIESDRTLLARDNCLAA-ALR---YGTTYLIIGKD	GEGYMNNEH-G	HIWYRY-VLD	QTSIIQQMAS	YLGDLKDAGS	1674
Tt-C3	-VTSVFKIEGIEQKQNLHTTKVLMVRDSCSA-DLD---IQQBYVIMGRD	GAQFKDEDTG	ILLYRY-LID	QTSIFKTR	----ISVAEN	1714
Bb-C3	-VESPIKKGLD--DVPFDGEERVFWKRKTCQSL-QLV---EGTTYLLMGKD	GTKYTDQEG-G	FDSFRY-VIT	EQSFVAWPT	-GAKAERKKF	1711
Nv-C3	-VLEVYNAGRG--KVNGKNIIELKQSGCREP-NMN--KMEBYFIAGSG	KGGV-----	---Y-SID	SSVYAARTR	-----QNDKKL	1695
Aj-C3	-ITQVLKTAND--EVFGEEVRQFVWSSSCSAV-RMQ---NQRSYLVIGQT	VLPYADTD-G	NDSFRY-IID	AATTIEIPL	----RNKRRV	1716
Rd-C3	VVVTSINKGIE--DLNEDEAIAFFWFKSCYP-P-ELE---EGKTYLLIGRD	GNKFRIGE--	EEKYRYPLMG	TALVIENITP	QQARQNKKKY	1724
Sp-C3	-VLNPKITGDE--DVPHQARQLFINEGDCD-CVKVGRNIGGTFLLVGQK	SLKYTTQE-G	ERRYRY-VYG	PTSKLEPWA	----TRKARN	1675
Hu-C3	-IEQTIKSGSD--EVQVGQQTFTFISPIKCREALKLE---EKKHYLLMWGLS	SDFWGEKP--	--NLSSY-IIG	KDTWEHNP	-EDBCEQDEEN	1642
Hu-C4B	-ITQVLHFTKD-VKAAANQMRFNLVRASCLL--RLE---PGKBYLLIMGD	GATYDLEG--	--HPSY-LLD	SNSWIEEMPS	-ERLCRSTRQ	1723

Lox-C3	QNMMLRALHWL RA--ELKK-G	CDL--	1726	
Hd-C31	KNLQRDLNSV TK--ALKENG	CEL--	1738	
Hd-C32	KLLQRDLNQV TK--ILKEDG	CDM--	1695	
Tt-C3	KQLTKAFRWL EKHHMMGGGG	CPQ--	1737	
Bb-C3	KRVADKFQQL AD--KIFLQG	CTT--	1732	
Nv-C3	KKLASRID--	-----GS	CAK--	1708
Aj-C3	REQFLKMQUEF KEDMLHGNNN	Q--	1738	
Rd-C3	KLLHTVLTNF KR--LMLGNG	QON--	1745	
Sp-C3	AAVFDKLVAE EAQ-MAPASA	CSNEN	1699	
Hu-C3	KQKQCQDLGAF TE--SMVVG	CPN--	1663	
Hu-C4B	RAACAQLNDF LQ--EYGTQG	QOV--	1744	

Alinhamento múltiplo da sequência Lox-C3 com outras sequências proteicas similares ao C3 de outros organismos. O alinhamento foi produzido utilizando o algoritmo MUSCLE disponível no programa MEGA 6. Aminoácidos destacados em preto indicam regiões idênticas. A região de cada domínio está indicada acima da sequência, assim como as regiões que conferem alguma atividade funcional. Os resíduos de cisteína estão indicados com asteriscos. *L. laeta* (Lox-C3), *H. adansonii* (Hd-C31; BAK64109.1, Hd-C32 BAK64110.1), *T. tridentatus* (Tt-C3; BAH02276.1), *B. belcheri* (BbC3; ABK78771.2), *N. vectensis* (NvC3; BAH22723.1), *A. japonicus* (ADN97000.1), *R. decussatus* (RdC3; ACN37845.1). *Homo sapiens* (Hu-C3; P01024) (C4B; P0C0L5).

FIGURA 33 - Modelo tridimensional de Lox-C3

Construção do modelo molecular de Lox-C3 com base na estrutura do C3 bovino (PDB 2b39.2A), utilizando a ferramenta SWISS-Model Workspace disponível no endereço eletrônico <http://swissmodel.expasy.org/workspace/>. As análises posteriores foram efetuadas utilizando o programa SwissPDBViewer e utilizando como base comparativa a estrutura do C3 humano (PDB 2A73).

4.8 Análises de similaridade e filogenia da sequência Lox-C3

A sequência deduzida de aminoácidos Lox-C3 (1726 aa) foi submetida à análise de homologia e exibiu uma similaridade significativa com proteínas do complemento, apresentando valores de identidade de 29%, 28% e 27% com C3, C4 e C5 humanos, respectivamente. Os valores de identidade obtidos com C3 de invertebrados foram maiores, variando entre 30% e 54%, sendo este último representado pelo C3 da aranha da família Salcitidae, *Hasarius adansoni* (TABELA 3).

Para investigar a história evolutiva das proteínas da superfamília TEP, além de Lox-C3, 97 sequências derivadas de vertebrados e invertebrados foram utilizadas e submetidas à análise filogenética. A árvore filogenética não enraizada, construída com base no método de Máxima Verossimilhança (FIGURA 34), demonstra a divergência de dois grupos (*Bootstrap* de 100%), um deles representado pelos componentes da família α -2 macroglobulina e o segundo representado pelas proteínas pertencentes à família C3/C4/C5 e, considerando este último, pode se observar uma clara separação (*Bootstrap* de 95%) entre as proteínas provenientes de vertebrados (azul) e de invertebrados (rosa). Com relação ao grupo C3/C4/C5 de invertebrados, o clado representado pelos cnidários, animais diblásticos (grupo A) parece ter divergido primeiro, apesar de obter um valor de confiança baixo (58%). Em seguida, ocorre a divergência (34%) entre o grupo B, composto pelos equinodermos e anfioxos (Cephalochordata), e o restante dos componentes de invertebrados (C, D, E, F) e vertebrados. O clado C (artrópodes) e D (moluscos) aparecem como grupos irmãos em 36% do total das árvores amostradas e o grupo E, representado pelo molusco cefalópode (*Squid*), aparece como grupo ancestral mais recente (31%) dos componentes C3-like das ascídias (F) e dos componentes C3/C4/C5 de vertebrados.

Com relação ao clado C, no qual estão presentes todos os componentes C3 de artrópodes, ocorre uma divergência de dois ramos (98%), um representado pelo Subfilo Myriapoda (centípede) e outro pelo Subfilo Chelicerata. A sequência C3-like das aranhas *Loxosceles* está posicionada dentro do grupo do Subfilo Chelicerata e mais, especificamente, no mesmo subramo composto pelas duas isoformas de C3 encontradas na aranha *Hasarius adansoni* (C3-1 e C3-2 *Jumping Spider*). A configuração do clado C demonstrou ser bem confiável, já que obteve valores altos de *Bootstrap*, entre 95 e 100%.

Tabela 3 - Análise comparativa da sequência Lox-C3 com outras sequências similares

Organismo	Nº de Acesso	I (%) [*]	S (%) ^{**}	Score	E-value
C3-1 <i>Hasarius adansoni</i>	BAK64109.1	53	70	1921	0.0
C3-2 <i>Hasarius adansoni</i>	BAK64110.1	45	64	1552	0.0
C3 <i>T. tridentatus</i>	BAH02276.1	44	62	1456	0.0
C3 <i>C. rotundicauda</i>	AAQ08323.1	43	62	1449	0.0
C3 <i>Ammothoa sp.</i>	BAR45600.1	36	56	1113	0.0
C3 <i>B. belcheri</i>	BAB47146.1	36	54	1020	0.0
C3-1 <i>Scolopendra s. japonicus</i>	BAR45608.1	34	53	1016	0.0
C3-2 <i>Scolopendra s. japonicus</i>	BAR45609.1	34	53	1003	0.0
C3-3 <i>Scolopendra s. japonicus</i>	BAR45610.1	33	53	918	0.0
C3 <i>N. vectensis</i>	BAH22724.1	32	51	828	0.0
C3-2 <i>N. vectensis</i>	BAH22725.1	31	50	824	0.0
C3 <i>A. japonicas</i>	ADN97000.1	30	50	796	0.0
C3 <i>Swiftia exserta</i>	AAN86548.1	31	49	764	0.0
C3 <i>Acropora millepora</i>	ABK78771.2	34	52	736	0.0
C3 <i>Haliplanella lineata</i>	BAJ05269.1	31	51	671	0.0
C3 <i>Ruditapes decussatus</i>	ACN37845.1	31	48	704	0.0
C3 <i>Crassostrea gigas</i>	W8JCY2	30	49	691	0.0
C4 <i>Egretta garzeta</i>	A0A091NJ3	30	48	664	0.0
C4-A <i>Pelodiscus sinensis</i>	XP_006125003.1	30	48	658	0.0
C4 <i>Acanthisitta chloris</i>	XP_009072684.1	29	47	655	0.0
C4A <i>Chelonia mydas</i>	XP_007070975.1	30	47	655	0.0
C4-1 <i>Anser anser</i>	ADY68766.1	30	48	651	0.0
C4 <i>Xenopus laevis</i>	AAI70422.1	32	47	637	0.0
C4-1 <i>Cyprinus carpio</i>	BAB03284.1	29	47	630	0.0
C3 <i>Euprymna scolopes</i>	ACF04700.1	28	45	620	0.0
C3 <i>Homo sapiens</i>	P01024.2	27	45	516	4e-157
C4-B <i>Homo sapiens</i>	NP_001002029.3	28	45	558	7e-165

*Identidade (I) foi calculada com base na porcentagem de aminoácidos idênticos por coluna/posição no alinhamento.

**Similaridade (S) foi calculada com base na porcentagem de resíduos não idênticos, mas similares, que são substituições conservativas (mantendo as propriedades físico-químicas de aminoácidos). O valor do E value de um hit, para um dado score, corresponde a probabilidade de se obter com outra sequência aleatória de mesmo tamanho e composição de letras, outro alinhamento com score igual ou superior. Quanto mais próximo de zero for o E value, mais confiável é a consulta.

5 DISCUSSÃO

As pesquisas dirigidas a evolução do sistema complemento têm avançado muito nos últimos anos e, devido às inovações tecnológicas, incluindo as análises genômicas e transcriptômicas, a possibilidade de se descobrir genes relacionados ao sistema complemento, em diferentes organismos da linhagem dos invertebrados, tem crescido. Em 1996, Smith e colaboradores identificaram genes *C3-like* expressos nos coelomócitos do ouriço-do-mar *S. purpuratus* e, a partir daí, foi sugerido que o sistema complemento teria se originado na linhagem dos deuterostômios, uma vez que este componente não havia sido identificado no inseto *Drosophila melanogaster* e no nematoda *Caenorhabditis elegans*. Esta hipótese foi reformulada pela identificação de genes *C3-like* em vários organismos da linhagem dos protostômios, incluindo a aranha *Hasarius adansoni* (SEKIGUCHI; FUJITO; NONAKA, 2011), os límulos *C. rotundicauda* (ZHU et al., 2005) e *T. tridentatus* (ARIKI et al., 2008), representantes do Filo Arthropoda; dos moluscos *R. decussatus* (Bivalvia) (PADRO-ALVAREZ et al., 2009) e *E. scolopes* (Cephalopoda) (CASTILLO; GOODSON; MCFALL-NGAI, 2009) e da linhagem anterior à divergência dos deuterostômios e protostômios, o coral *Swiftia exserta* e anêmona *N. vectensis* (Cnidaria) (DISHAW; SMITH; BIGGER, 2005; KIMURA; SAKAGUCHI, NONAKA, 2009). Estas últimas descobertas sugerem que o surgimento dos primeiros genes de componentes do complemento ocorreu anteriormente à divergência dos Cnidaria e Bilateria, há mais de 1 bilhão e 300 milhões de anos, e que a ausência destes em *Drosophila melanogaster* deve ser devida a uma perda secundária (NONAKA; KIMURA, 2006). O presente trabalho confirma esta hipótese, uma vez que genes para FB e *C3-like* foram encontrados expressos na glândula de veneno de *L. laeta*.

Inúmeros trabalhos têm indicado a existência de um sistema complemento pouco complexo, denominado como “arqueo-sistema complemento” envolvendo proteínas *C3-like* associadas ao fator B (DOODS, 2002), como observado em lampréias (NONAKA; TAKAHASHI, SASAKI, 1994), ouriços-do-mar (SMITH; SHIH; DACHENHAUSEN, 2001), límulos (ZHU et al., 2005), urocordados (YOSHIZAKI et al., 2005), anfioxos (HE et al., 2008), anêmonas-do-mar (KIMURA; SAKAGUCHI; NONAKA, 2009) e pepinos-do-mar (ZHONG; ZHANG; CHANG, 2012). O fator B é um componente da via alternativa e como sua presença foi observada em organismos da linhagem primitiva (cnidários e protostômios) é possível que esses organismos sejam dotados da via alternativa.

Todas as sequências das proteínas C2 e fator B caracterizadas, até o momento, preservam a mesma composição de domínios clássica encontrada em proteínas de vertebrados relacionadas. Apesar de algumas particularidades existentes na porção C-terminal, tais como dois domínios extras como observado no ouriço do mar e nos límulo ou domínios adicionais, como o domínio similar ao receptor de lipoproteína de baixa densidade (LDLR) encontrado nas ascídias ou o domínio similar ao fator de crescimento (EGF_CA) no anfioxo, as proteínas C2/FB apresentam regiões conservadas que desempenham papéis importantes para a ativação e, por isso, podem ser consideradas ortólogas dos componentes C2 e fator B de mamíferos.

O presente trabalho representa a primeira identificação de um homólogo de fator B em aranhas *Loxosceles* (Arachnida) e, assim como as proteínas FB/C2 de vertebrados, Lox-FB é uma proteína em mosaico composto por domínios CCP, vWFA e serino protease (SP). Regiões altamente conservadas foram identificadas em Lox-FB tais como as posições *consensus* presente no domínio CCP, o sítio MIDAS encontrado no domínio vWFA e as posições conservadas próximas ao centro catalítico do domínio SP. Contudo, considerando a arquitetura de domínios, Lox-FB apresenta somente dois CCPs, um a menos do que é visto em vertebrados. Esses resultados levantam dúvidas sobre as propriedades funcionais de Lox-FB, como também foi observado no fator B-like presente na aranha *Hasarius adansoni* (isoforma 3), na lacraia *Scolopendra subspinipes japonica* (isoforma 2), na “aranha-do-mar” *Ammothea sp* e no bivalve *Ruditapes decussatus* (PADRO-ALVAREZ et al., 2009; SEKIGUCHI ; NONAKA, 2015). O fator B pertence à classe de peptidases do tipo A, uma vez que apresenta um domínio SP com uma conformação característica da família das quimiotripsinas, sendo que quase todos os representantes desta classe utilizam a tríade catalítica composta pelos seguintes aminoácidos: ácido aspártico (D), histidina (H) e serina (S) (PAGE; CERA, 2008). Apesar de algumas similaridades nas estruturas secundárias entre o fator B humano e Lox-FB, a tríade catalítica não se manteve conservada e, portanto, a atividade de uma C3 convertase putativa de *Loxosceles* é, ainda, uma questão a ser esclarecida.

Até o momento, não há estudos suficientes sobre a existência de uma atividade proteolítica específica de uma C3 convertase C3bBb-like da via alternativa e de como esta é produzida em invertebrados. Zhu et al. (2008) demonstraram que a atividade de serino protease presente no plasma do límulo *C. rotundicauda* era induzida por alguns padrões moleculares associados aos patógenos (PAMPs), como o LPS e o LTA de uma maneira

dependente de íons Ca^{2+} e Mg^{2+} . Contudo, não está claro se CrC2/Bf participa diretamente na ativação de CrC3 ou se precisa ser ativado previamente por uma outra serino protease, similar ao fator D dos vertebrados. Le saux et al. (2008) identificaram que CrC2/Bf era capaz de se ligar a três membros de receptores de reconhecimento padrão (PRRs): proteína ligante à galactose (GBP), carcinolectina-5 (CL-5) e proteína C-reativa (CRP), promovendo, então, sua ligação aos patógenos e, consequentemente, ativando o sistema complemento. Esses estudos sugerem que CrC2/Bf poderia desempenhar um papel equivalente à MASP, participando, talvez, de uma via similar a das lectinas. Embora o sistema complemento dos límulo seja o mais bem estudado, considerando os ensaios de atividade proteolítica, a maior parte dos dados derivada de FB/C2-like de invertebrados é representada por características baseadas em suas estruturas putativas. Em outras palavras, não há evidência experimental de que aquelas isoformas que não possuíam a tríade catalítica clássica tenham, de fato, perdido sua atividade proteolítica.

Além disso, é possível que no sistema complemento das aranhas *Loxosceles* exista um mecanismo de ativação diferente daquele apresentado pelos mamíferos, assim como já foi relatado para o límulo *Tachypleus tridentatus*. Esses organismos apresentam uma serino protease denominada Fator C, que além de funcionar como um fator iniciador da coagulação da hemolinfa, estocado dentro dos hemócitos, atua como uma C3 convertase na superfície de bactérias Gram-negativas, na fase inicial da ativação do sistema complemento. Talvez, um componente similar ao fator C esteja presente na hemolinfa ou expresso na glândula de veneno das aranhas *Loxosceles*, no entanto, a possível atividade catalítica de Lox-FB deve ser ainda avaliada. Recentemente, Tagawa et al. (2012) caracterizaram duas isoformas de fator B do límulo *T. tridentatus* (TtC2/FB-1 e TtC2/FB-2) sendo que ambas são indispensáveis para a deposição do fragmento TtC3b em bactéria gram-positiva e fungos. Mesmo não tendo identificado uma serino protease similar ao fator D nos límulos, foi concluído que existem outros componentes, como as lectinas do plasma importantes para o recrutamento de C3bBb-like na superfície de bactérias gram positivas ou de fungos, mas não nas bactérias gram negativas. Portanto, parece que o mecanismo de ativação do sistema complemento de límulo é diferente quando comparado aos de mamíferos e, talvez, as aranhas *Loxosceles* apresentem o mesmo padrão de ativação.

Na análise filogenética, produzida com base na sequência de Lox-FB e de outras sequências similares (**Figura 16**), foi detectada uma clara divergência entre as proteínas

presentes em vertebrados e em invertebrados. No clado representado pelos componentes C2 ou fator B de vertebrados, o FB da lampreia *L. camtschaticum* aparece como grupo irmão dos vertebrados, indicando também que os eventos de duplicação gênica que levaram à divergência dos genes C2 e fator B, aconteceram antes do surgimento da linhagem dos vertebrados mandibulados. Essa configuração da árvore filogenética está de acordo com a ausência da via clássica no grupo dos vertebrados não mandibulados (Agnatha) visto que estes não apresentam imunoglobulinas (SUNYER; LAMBRIS, 1998). Quase todas as sequências de invertebrados apresentam mais de uma isoforma e, alguns deles, agruparam em um mesmo clado (grupo C) como observado para as duas isoformas (1 e 2) dos límulos *Tachypleus tridentatus* (horseshoe crab 2 C2/FB), *Carcinoscorpius rotundicauda* (horseshoe crab 1 C2/FB), pepino-do-mar *Apostichopus japonicus* e a aranha *Hasarius adansoni*. Contudo, algumas isoformas não se agruparam em um mesmo clado; por exemplo, a segunda isoforma da lacraia *Scolopendra subspinipes* e da “aranha-do-mar” *Ammothea sp.* e a terceira isoforma da aranha *Hasarius adansoni* se agruparam no clado B no qual Lox-FB também está presente. Essas proteínas parecem ser diferentes daquelas isoformas presentes nos respectivos organismos encontrados no grupo C e, aparentemente, possui uma relação filogenética mais próxima com as sequências FB-like do bivalve *Ruditapes decussatus* e do anfioxo *B. belcheri*.

A história filogenética de uma proteína (microevolução), nem sempre acompanha a evolução das espécies (macroevolução), porque as pressões evolutivas que agem no nível de proteína e no organismo todo podem ser divergentes. Além da participação do tipo de patógeno que infectará o hospedeiro, outros fatores influenciarão na construção do repertório imunológico e seu respectivo modo de ação, como, por exemplo, a adoção de diferentes habitats, histórias de vida, diferenças na massa corpórea e na sua complexidade (LOKER, 2012). Considerando o organismo total, existem muitos tipos de pressões seletivas que influenciam a capacidade de sobrevivência, como as mudanças climáticas, predatismo, disponibilidade de comida e infecções. O fator B é uma proteína relacionada à resposta imunológica e as pressões seletivas direcionadas são representadas por infecções recorrentes que os organismos são expostos. Portanto, é possível que o bivalve *Ruditapes decussatus* e a aranha *Loxosceles laeta* tenham sido infectados por patógenos que expressam os mesmos padrões moleculares, explicando, assim, a distribuição dessas espécies em um mesmo grupo na árvore filogenética.

Para o componente Lox-C3, foram obtidos diferentes clones, 2 clones Lox-C3_TEP (clones 6 e 7), Lox-C3_ANATO, Lox-C3_5', Lox-C3 3'_10B+17C e Lox-C3 3'_16C e somente a partir das sequências Lox-C3 5' e Lox-C3 3'_10B+17C foi possível identificar a sequência completa de Lox-C3. Os demais clones apresentaram variações na sequência de DNA e de aminoácidos, sugerindo a presença de polimorfismos e isoformas de C3-like de aranhas *Loxosceles*. É necessário identificar a porção 5' da sequência C3 correspondente ao clone Lox-C3 3'_16C e novas análises devem ser realizadas para resolver essa sequência.

Os resultados referentes à amplificação do componente C3-like de *L. laeta* (Lox-C3) permitiram identificar domínios característicos do componente C3 de vertebrados, incluindo os domínios α 2-macroglobulina (A2M) presentes tanto na cadeia α como na cadeia β , um domínio anafilatoxina (ANATO), um domínio contendo tioéster (TED) e um domínio C345R. A presença dos domínios C345R e ANATO é comum a todas as proteínas relacionadas ao complemento como C3, C4 e C5 humano, diferentemente do que é observado nas proteínas pertencentes à família α 2-macroglobulina, como TEP, A2M e CD109, que não apresentam estes domínios. Além dos domínios identificados, no alinhamento de Lox-C3 com a sequência protéica de C3 humano, foram observadas regiões características correspondentes ao sítio de ligação aos fatores H e B do sistema complemento, à properdina e ao receptor CR3 presente em leucócitos. A presença de um sítio de ligação para CR3 sugere uma possível atividade opsonizante para o C3 nessas aranhas.

O alinhamento entre as sequências Lox-C3 e C3-like de outros organismos (**Figura 32**) revelaram várias regiões conservadas nos cinco domínios analisados, além da presença de uma combinação dos quatro aminoácidos RKKR e RVKR que constituem o sítio de clivagem $\alpha\beta$ e $\alpha\gamma$, respectivamente. A combinação dos quatro aminoácidos RKKR presentes no C4 humano foi predominante nos seis organismos analisados, com exceção do componente C3-like de *N. vectensis* que, provavelmente, mantém as duas subunidades α e β juntas somente por forças não covalentes (KIMURA; SAKAGUCHI; NONAKA, 2009). Embora as regiões RVKR, de Lox-C3 e de C3-like do límulo *C. rotundicauda*, não se alinhem completamente é possível que a presença desta sugira a existência de uma terceira cadeia no C3 de *Loxosceles*, assim como no C4 humano. Estes resultados suportam a hipótese de que a proteína C3 ancestral seja composta por três cadeias, como foi também observado no molusco *R. decussatus* (PRADO-ALVAREZ et al., 2009), no coral *S. exserta* (DISHAW; SMITH, BIGGER, 2005), em uma das isoformas do C3-like presentes na anêmona-do-mar *N. vectensis* (NvC3-2) e na lampréia

Lampetra japonica (NONAKA; TAKAHASHI, 1992). No entanto, a análise de um maior número de moléculas C3-like, que podem estar presentes em outros organismos da linhagem dos protostômios é, ainda, necessária para a identificação da estrutura do C3 original.

O motivo tioéster formado pelos cinco aminoácidos GCGEQ, encontrado no componente C3 de vertebrados e invertebrados até agora estudados, se manteve conservado, assim como observado para a sequência C3-like de *L. laeta*. Com exceção do aminoácido glicina (G) localizado na posição 1011 dos componentes C3 e C4 de humanos, que foi substituído por alanina (A) em Lox-C3, todos os outros quatro aminoácidos foram idênticos. Como a glicina e a alanina são aminoácidos apolares e se diferenciam apenas pela presença de um radical metil presente neste último, provavelmente, esta diferença não tenha comprometido a função exercida por este sítio, mesmo porque, os principais aminoácidos envolvidos na interação do tioéster ativo aos grupos amino ou hidroxila presentes na superfície de microrganismos são o ácido glutâmico (Q) e a cisteína (C). Após a clivagem de C3 e a consequente liberação do fragmento C3a, uma mudança conformacional ocorre no fragmento C3b que permite que a histidina (H), localizada na posição 1126 no C3 humano, se ligue ao tioéster para formar um intermediário acil-imidazol com o ácido glutâmico (Q) e libere a cisteína (C) ao mesmo tempo. O ânion tiolato livre da cisteína age como catalisador para a ligação dos substratos que apresentam grupos hidroxila ao intermediário (GADJEVA et al., 1998). Em quase todos os componentes C3-like de invertebrados estudados até o momento, inclusive Lox-C3, apresentam o resíduo catalítico histidina, indicando ainda mais, a presença de um motivo tioéster funcional. No entanto, foi observado que a isoforma 2, diferentemente da isoforma 1 de C3-like das aranhas *H. adansonii* não apresenta o resíduo de histidina, indicando uma possível perda de função.

A presença do domínio ANATO caracterizado pela presença dos seis resíduos de cisteína perfeitamente alinhados tanto para os componentes C3, C4 e C5 humano, como também para os componentes C3-like de invertebrados, sugere uma conformação similar observada na molécula do C3 humano, mantendo o sítio tioéster inativo sem se expor à superfície. Além disso, sua presença também sugere uma atividade pró-inflamatória caracterizada pela liberação da anafilatoxina como o fragmento C3a de vertebrados.

Com base na árvore filogenética de Lox-C3 demonstrada na **Figura 34**, é possível notar que os componentes C3 de invertebrados estão agrupados em um mesmo clado e parecem ter divergido anteriormente aos componentes C3, C4 e C5 presentes nos

representantes do Filo Chordata. A topologia da árvore, construída com base nas sequências utilizadas, indica que os componentes C3 de invertebrados representem a forma ancestral desta proteína e que foi a partir desta que se originaram os componentes da superfamília C3/C4/C5 presente em vertebrados. Os componentes C3 presentes nos dois representantes do Subfilo Urochordata (*C. intestinalis* e *H. roretzi*) ocupam a posição intermediária entre os vertebrados e os invertebrados e esta posição filogenética na árvore construída é coerente com algumas funções de C3 já caracterizadas nestes organismos. Tem sido relatado que os hemócitos provenientes de *H. roretzi* pré-incubados com seu respectivo plasma aumentaram a sua capacidade de fagocitar as células de levedura, em comparação aos hemócitos não incubados com o plasma e que esta propriedade de opsonização foi diminuída após tratamento com anticorpo anti-C3 de ascídias (NONAKA; AZUMI, 1998). Parece que essa propriedade de opsonização representa uma função primitiva do sistema complemento, também registrada em límulus *C. rotundicauda* (ZHU et al., 2005). Como a sequência Lox-C3 está agrupada no mesmo clado ocupado pelos límulus, é provável que essa mesma função de opsonização também tenha se mantido nas aranhas *Loxosceles*.

Outra observação interessante, encontrada na árvore filogenética dos componentes pertencentes à superfamília TEP é que os organismos da linhagem dos cnidários, representados pelas anêmonas e corais, formam um grupo monofilético separado das demais sequências C3 de invertebrados. Essa configuração pode ser explicada devido às características particulares do componente C3 das anêmonas, como por exemplo, uma região altamente rica de aminoácidos lisina e arginina inserida no domínio MG8 e o outro fator exclusivo é a ausência de uma ponte dissulfeto que liga as cadeias α e β , sugerindo que as duas cadeias sejam mantidas juntas somente por forças não covalentes (KIMURA, SAKAGUCHI; NONAKA, 2009). Curiosamente, o componente C3 do anfioxo (Cephalochordata) se encontra mais próximo aos componentes equivalentes dos equinodermos, ao invés de se agrupar com os outros representantes do mesmo filo Chordata como as ascídias. Nas árvores filogenéticas construídas para Lox-C3 e para Lox-FB, os dois componentes provenientes do anfioxo *B. belcheri* e das ascídias se encontram em clados separados. Como as topologias das árvores filogenéticas, relativas aos ramos onde estão presentes o fator B e C3 do anfioxo, mostraram-se similares, é provável que estas proteínas estejam mais relacionadas evolutivamente com os respectivos componentes considerados mais primitivos. No caso da filogenia das proteínas da Família C2/FB, o componente FB-like do anfioxo se agrupou no

clado B no qual estão presentes as isoformas FB possivelmente inativas. Contudo, as análises filogenéticas aqui apresentadas para ambos os componentes, ainda não são conclusivas, visto que, para algumas relações filogenéticas, não foi possível atingir valores de *Bootstrap* significativos.

Outros estudos serão necessários para entender a natureza da proteína Lox-FB e Lox-C3 e desvendar como estes dois componentes interagem com os demais componentes do sistema complemento, possivelmente, presentes na hemolinfa/glândula de veneno das aranhas *Loxosceles*.

6 CONCLUSÕES

- Este trabalho representa a primeira caracterização de dois genes que compõem o sistema complemento das aranhas *Loxosceles*, Lox-C3 e Lox-FB, similares aos componentes C3 e Fator B de outros organismos;
- Em contraste às proteínas C2/FB que contêm três módulos de CCP, Lox-FB contém somente dois CCPs.
- A maioria das regiões funcionais é conservada em Lox-FB comparado com outras proteínas FB-like, como por exemplo, o sítio MIDAS localizado no domínio vWFA. No entanto, aqueles aminoácidos que compõem a tríade catalítica, presentes no domínio serino protease, aparentemente, foram substituídos por aminoácidos diferentes.
- Na análise filogenética, Lox-FB foi incluído no grupo dos invertebrados, sendo a isoforma 3 do fator B-like da aranha *Hasarius adansoni*, a mais relacionada, com uma identidade de 43%.
- Uma estrutura de três cadeias para o Lox-C3 maduro é prevista, baseada na presença de sítios de clivagem entre as cadeias α e γ , assim como é observado nas proteínas C3-like de invertebrados.
- Todos os domínios funcionais conhecidos no C3 humano como o domínio ANATO e domínio que contém o motivo tioéster encontram-se conservados em Lox-C3.
- A análise filogenética do componente C3 do sistema complemento permitiu inferir que Lox-C3 está localizado dentro do grupo das proteínas existentes em invertebrados e possui uma maior similaridade (53%) com o C3 da aranha *H. adansoni* da Família Salcidae, cuja função primitiva de opsonização do componente C3 possa, talvez, ter sido mantida nas aranhas *Loxosceles*.
- As análises preditivas com relação à estrutura e filogenia de ambos os componentes, contribuirão para novos estudos direcionados a entender a natureza de Lox-C3 e Lox-FB e investigar como estes interagem entre si e com outros possíveis componentes presentes na glândula de veneno ou na hemolinfa.

REFERÊNCIAS*

ADAMS, M. D.; CELNIKER, S. E.; HOLT, R. A.; EVANUS, C. A.; GOCAYNEJ, D.; AMANATIDES, P. G.; SCHERER, S. E.; LI, P. W.; HOSKINS, R. A.; GALLE, R. F. The genome sequence of *Drosophila melanogaster*. **Science**, v. 287, p. 85-95, 2000

AL-SHARIF, W. Z.; SUNYER, J. O.; LAMBRIS, J. D.; SMITH, L. C. Sea urchin coelomocytes specifically express a homologue of the complement component C3. **Journal of Immunology**, v. 160, p. 2983-2997, 1998

ALTSCHUL, S. F.; MADDEN, T. L.; SCHÄFFER, A. A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Research**, v. 25, p. 3389-3402, 1997

ARIKI, S.; TAKAHARA, S.; SHIBATA, T.; FUKUOKA, T.; OZAKI, A.; ENDO, Y.; FUJITA, T.; KOSHIBA, T.; KAWABATA, S. Factor C acts a lipopolysaccharide-responsive C3 convertase in horseshoe crab complement activation. **Journal of Immunology**, v. 181, p. 7994-8001, 2008

ARLAUD, G. J.; GABORIAUD, C.; THIELENS, N. M.; ROSSI, M. Structural biology of C1. **Biochemical Society Transactions**, v. 30, p. 1001-1006, 2002

AUSUBEL, F. A; **Current protocols in molecular biology**. New York: John Wiley and Sons, 1995. v.2

BARBARO, K. C.; SOUSA, M. V.; MORHY, L.; EICKSTEDT, V. R.; MOTA, I. Compared chemical properties of dermonecrotic and lethal toxins from spiders of the genus *Loxosceles* (Araneae). **Journal of Protein Chemistry**, v. 15, p. 337-343, 1996

BARRETO, O. C.; DE CILLO, D. M.; NONOYAMA, K.; ANTONIO, L. C.; MORENA, P.; CARDOSO, J.L. Methemoglobinemia associated with loxoscelism. **Revista do Instituto de Medicina Tropical de São Paulo**, v. 32, p. 1-5, 1990

BAXTER, R. H. G.; CHANG, C.; CHELLIAH, Y.; BLANDIN, S.; LEVASHINA, E. A.; DEISENHOFER, J. Structural basis for conserved complement factor-like function in the antimalarial protein TEP1. **Proceedings of the National Academy of Sciences**, v. 104, p. 11615-11620, 2007

BENTLEY, D. R.; CAMPBELL, R. D. C2 and Factor B: structure and genetics. **Biochemical Society Symposium**, v. 51, p. 7-18, 1986

BERNSTEIN B.; EHRLICH, F. Brown recluse spider bites. **Journal of Emergency Medicine**, v.4, p. 457-462, 1986

BEXBORN, F.; ANDERSSON, P. O.; CHEN, H.; NILSSON, B.; EKDAHL, K. N. The tick-over theory revisited: formation and regulation of the soluble alternative complement C3 convertase (C3(H₂O)Bb). **Molecular Immunology**, v. 45, p. 2370–2379, 2008

* De acordo com: ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR6023**: informação e documentação: referências: elaboração. Rio de Janeiro, 2002

BORDOLI, L.; KIEFER, F.; ARNOLD, K.; BENKERT, P.; BATTEY, J.; SCHWEDE, T. Protein Structure homology modeling using SWISS-MODEL workspace. **Nature Protocols**, v. 4, p. 1-13.

BÜCHERL, W. Biology and venoms of the most important South American spiders of the genera *Phoneutria*, *Loxosceles*, *Lycosa*, and *Latrodectus*. **American Zoologist**, v. 9, p.157–159, 1969

CAROLL, M. C. The complement system in regulation of adaptive immunity. **Nature Immunology**, v. 5, p. 981-986, 2004

CASTILLO, M. G.; GOODSON, M. S.; MCFALL-NGAI, M. Identification an molecular characterization of a complement C3 molecule in a lophotrochozoan, the Hawaiian bobtail squid *Euprymna scolopes*. **Developmental & Comparative Immunology**, v. 33, p. 69-76, 2009

CHONN, A.; CULLIS, P. R.; DEVINE, D. V. The role of surface charge in the activation of the classical and alternative pathways of complement by liposomes. **Journal of Immunology**, v. 146, p. 4234-4241, 1991

DEMPSEY, P. W.; ALLISON, M. E.; AKKARAJU, S.; GOODNOW, C. C.; FEARON, D. T. C3d of complement as a molecular adjuvant: Bridging innate and acquired immunity. **Science**, v. 271, p. 348–350, 1996

DISHAW, L. J.; SMITH, S. L.; BIGGER, C. H. Characterization of a C3-like cDNA in a coral: phylogenetic implications. **Immunogenetics**, v. 57, p. 535-548, 2005

DZIK, J. M. The ancestry and cumulative evolution of immune reactions. **Acta Biochimica Polonica**, v. 57, p. 443-466, 2010

DOBÓ, J.; SHROEDER, V.; JENNY, L.; CERVENAK, L.; ZÁVODSZKY, P.; GÁL, P. Multiples roles of complement MASP-1 at the interface of innate immune response and coagulation. **Molecular Immunology**, v. 61, n.2, p.69-78, 2014

DOODS, A. W. Which came first, the lectin/classical pathway or the alternative pathway of complement?. **Immunobiology**, v. 205, p. 340-354, 2002

FEARON, D. T.; AUSTEN, K. F. Properdin: binding to C3b and stabilization of the C3b-dependent C3 convertase. **The Journal of Experimental Medicine**, v. 142, p. 856–863, 1975

FERNANDES-PEDROSA, M.; JUNQUEIRA DE AZEVEDO, I. L.; GONÇALVES-DE-ANDRADE, R. M.; VAN DER BERG, C. W.; RAMOS, C. R.; LEE HO, P.; TAMBOURGI, D. V. Molecular cloning and expression of a functional dermonecrotic and haemolytic factor from *Loxosceles laeta* venom. **Biochemical and Biophysical Research Communications**, v. 298, n. 5, p. 638–645, 2002

FERNANDES-PEDROSA, M.; JUNQUEIRA-DE-AZEVEDO, I. L. M.; GONÇALVES-DE-ANDRADE, R. M.; KOBASHI, L. S.; ALMEIDA, D. D.; LEE HO, P.; TAMBOURGI, D. V. Transcriptome analysis of

Loxosceles laeta (Araneae, Sicariidae) spider venomous gland using expressed sequence tags. **BMC Genomics**, v. 9, n. 279, 2008

FISCHER, M. L.; VASCONCELOS-NETO, J. Microhabitats occupied by *Loxosceles intermedia* and *Loxosceles laeta* (Araneae:Sicariidae) in Curitiba, Paraná, Brazil. **Journal of Medical Entomology**, v. 42, p. 756-765, 2005

FUJITA, T.; ENDO, Y.; NONAKA, M. Primitive complement system recognition and activation. **Molecular Immunology**, v. 41, p. 103-111, 2004

FUTRELL, J. Loxoscelism. **The American Journal of the Medical Sciences**, v. 304, p. 261-26, 1992

GADJEVA, M., DODDS. A. W., TANIGUCHI-SIDLE, A., WILLIS, A. C., ISENMAN, D. E, LAW, S. K. A. The covalent Binding Reaction of Complement C3. **Journal of Immunology**, v.161, p. 985-990, 1998

GALIANO, M. E; HALL, M. Datos adicionales sobre el ciclo vital de *Loxosceles laeta* (Nicolet, 1849) (Araneae). **Physis**, v. 32, p. 277-288, 1973

GONÇALVES-DE-ANDRADE, R. M.; DE OLIVEIRA, K. C.; GIUSTI, A. L.; DIAS DA SILVA, W.; TAMBOURGI, D. V. Ontogenetic development of *Loxosceles intermedia* spider venom. **Toxicon**, v. 37, p. 627–632, 1999

GONÇALVES-DE-ANDRADE, R. M.; LOURENÇO, W. R.; TAMBOURGI, D. V. Comparison of the fertility between *Loxosceles intermedia* and *Loxosceles laeta* spiders (Araneae Sicariidae). **Journal of Arachnology**, v. 28, p. 245-247, 2000

GONÇALVES-DE-ANDRADE, R. M.; TAMBOURGI, D. V. First record on *Loxosceles laeta* (Nicolet, 1849) (Araneae, Sicariidae) in the west zone of Sao Paulo City , Sao Paulo, Brazil, and considerations regarding its geographic distribution. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 36, p. 425-426, 2003

GROS, P.; MILDEN, F. J.; JANSSEN, B. J. C. Complement driven by conformational changes. **Nature Reviews Immunology**, v. 8, p. 48-58, 2008

HALL, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. **Nucleic Acids Symposium Series**, v. 41, p. 95-98, 1999

HARBOE, M.; MOLLNES, T. E. The alternative complement pathway revisited. **Journal of Cellular and Molecular Medicine**, v. 12, p. 1074-1084, 2008

HE, Y.; TANG, B.; ZHANG, S.; LIU, Z.; ZHAO, B.; CHEN, L. Molecular and immunochemical demonstration of a novel member of Bf/C2 homolog in amphioxus *Branchiostoma belcheri*: Implications for involvement of hepatic cecum in acute phase response. **Fish and Shellfish Immunology**, v.24, p. 768-778, 2008

HOLMES, D. S.; QUIGLEY, M. A rapid boiling method for the preparation of bacterial plasmids. **Analytical Biochemistry**, v. 114, p. 193-197, 1981

HOURCADE, D. E.; WAGNER, L. M.; OGLESBY, T. J. Analysis of the short consensus repeats of human complement factor B by site-directed mutagenesis. **The Journal of Biological Chemistry**, v. 270, p. 19716-19722, 1995

HOURCADE, D. E. The role of properdin in the assembly of the alternative pathway C3 convertase of complement. **The Journal of Biological Chemistry**, v. 281, p. 2128-2132, 2006

HOGAN, C. J. H.; BARBARO, K. C.; WINKEL, K. Loxoscelism: Old Obstacles, New Directions. **Annals of Emergency Medicine**, v. 44, p. 608-624, 2004

JANSSEN, B. J. C.; HUIZINGA, E. G.; RAAIJMAKERS, H. C. A.; ROOS, A.; DAHA, M. R.; NILSSON-EKDAHL, K.; NILSSON, B.; GROS, P. Structures of complement component C3 provides insights into the function and evolution of immunity. **Nature**, v. 437, p. 505-511, 2005

KIMURA, A.; SAKAGUCHI, E.; NONAKA, M. Multi-component complement system of Cnidaria: C3, Bf, and MASP genes expressed in the endodermal tissues of a sea anemone, *Nematostela vectensis*. **Immunobiology**, v. 214, p. 165-178, 2009

KRISHNAN, V.; PONNURAJ, K.; XU, Y.; MACON, K.; VOLANAKIS, J. E.; NARAYANA, S. V. L. The crystal structure of cobra venom factor, a cofactor for C3 and C5 convertase CVFBb. **Structure**, v. 17, p. 611-619, 2009

KISHORE, U.; GHAI, R.; GREENHOUH, T. J.; SHRIVE, A. K.; BONIFATI, D. M.; GADJEVA, M. G.; WATERS, P.; KOJOUHAROVA, M. S.; CHAKRABORTY, T.; AGRAWAL, A. Structural and functional anatomy of the globular domain of complement protein C1q. **Immunology Letters**, v. 95, p. 113-128, 2004

LACHMANN, P. J.; MILLER-EBERHARD, H. J. The demonstration in human serum of "conglutinin activating factor" and its effect on the third component of complement. **Journal of Immunology**, v. 100, p. 691-698, 1968

LAMBRIS, J. D. The multifunctional role of C3, the third component of complement. **Immunology Today**, v. 9, p. 387-393, 1988

LE SAUX, A.; NG, P. M.; KOH, J. J.; LOW, D. H.; LEONG, G.; HO, B.; DING, J. L. The macromolecular assembly of pathogen-recognition receptors is impelled by serine proteases, via their complement control protein modules. **Journal of Molecular Biology**, v. 377, p. 902-913, 2008

LOKER, E.; Macroevolutionary immunology: a role for immunity in the diversification of animal life. **Frontiers in Immunology**, v. 3, p.1-21, 2012

LOWRIE, D. C. Starvation longevity of *Loxosceles laeta* (Nicolet) (Araneae). **Entomological News**, v. 91, p. 130-132, 1980

MARINO, R.; KIMURA, Y.; DE SANTIS, R.; LAMBRIS, J. D.; PINTO, M. R. Complement in urochordates: cloning and characterization of two C3-like genes in the ascidian *Ciona intestinalis*. **Immunogenetics**, v. 53, p. 1055-1064, 2002

MARQUES-DA-SILVA, E.; FICHER, M. L. Distribuição das espécies do gênero *Loxosceles* Heineken & Lowe, 1835 (Araneae; Sicariidae) no Estado do Paraná. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 38, p. 331-335, 2005

MILDER, F. J.; GOMES, L.; SCHOUTEN, A.; JANSSEN, B. J. C.; HUIZINGA, E. G.; ROMIJIN, R. A.; HEMRIKA, W.; ROOS, A.; DAHA, M. R.; GROS, P. Factor B structure provides insights into activation of the central protease of the complement system. **Nature Structural and Molecular Biology**, v. 14, p. 224-228, 2007

MINISTÉRIO DA SAÚDE. **Acidentes por animais peçonhentos**. Disponível em: <http://dtr2004.saude.gov.br/sinanweb>, 2015

MITCHELL, T. J.; NAUGHTON, M.; NORSWORTHY, P.; DAVIES, K. A.; WALPORT, M. J.; MORLEY, B. J. IFN-gamma up-regulates expression of the complement components C3 e C4 by stabilization of mRNA. **Journal of Immunology**, v. 156, p. 4429-4434, 1996

NANCE, W. E. Hemolytic anemia of necrotic arachnidism. **The American Journal of Medicine**, v. 31, p. 801-7, 1961

NONAKA, M.; TAKAHASHI, M. Complete complementary DNA sequence of the third complement of lamprey. Implication for the evolution of thioester containing-proteins. **Journal of Immunology**, v. 148, p. 3290-3295, 1992

NONAKA, M.; TAKAHASHI, M.; SASAKI, M. Molecular cloning of a lamprey homologue of the mammalian MHC class III gene, complement factor B. **Journal of Immunology**, v. 152, p. 2263-2269, 1994

NONAKA, M.; AZUMI, K. Opsonic complement system of the solitary ascidian, *Halocynthia roretzi*. **Developmental and Comparative Immunology**, v. 23, p. 421-427, 1999

NONAKA, M.; KIMURA, A. Genomic view of the evolution of the complement system. **Immunogenetics**, v. 58, p. 701-713, 2006

NONAKA, M. The complement C3 protein family in invertebrates. **Invertebrate Survival Journal**, v. 8, p. 21-32, 2011

NONAKA, M. Evolution of the Complement System. **Subcell Biochemistry**, v. 80, p. 31-43, 2014

OLIVEIRA, K. C.; GONÇALVES-DE-ANDRADE, R. M.; GIUSTI, A. L.; DIAS DA SILVA, W.; TAMBOURGI, D. V. Sex-linked variation of *Loxosceles intermedia* spider venoms. **Toxicon**, v. 37, p. 217-221, 1999

OLIVEIRA, K. C.; GONÇALVES-DE-ANDRADE, R. M.; PIAZZA, R. M. F.; FERREIRA, J. M. C.; VAN DER BERG, C. W.; TAMBOURGI, D. V. Variations in *Loxosceles* spider venom composition and toxicity contribute to the severity of envenomation. **Toxicon**, v. 45, p. 421-429, 2005

PAGE, M. J.; CERA, E. D. Serine peptidases: Classification, structure and function. **Cellular and Molecular Life Sciences**, v. 65, p. 1220-1236, 2008

PAULI, I., PUKA, J., GUBERT, I. C., MINOZZO, J. C. The efficacy of antivenom in loxoscelism treatment. **Toxicon**, v. 48, p.123-137, 2006

PRADO-ALVAREZ, M.; ROTLANT J.; GESTAL, C.; NOVOA, B.; FIGUERAS A. Characterization of an C3 and a factor B-like in the carpet-shell clam, *Ruditapes decussates*. **Fish Shellfish Immunology**, v. 26, p. 305-315, 2009

PRETEL, F.; GONÇALVES-DE-ANDRADE, R. M.; MAGNOLI, F. C.; DA SILVA, M. E.; FERREIRA, J. M.; VAN DEN BERG, C. W.; TAMBOURGI, D. V. Analysis of the toxic potential of venom from *Loxosceles adelaida*, a Brazilian brown spider from karstic areas. **Toxicon**, v. 45, p. 449-58, 2005

RICKLIN, D.; HAJISHENGALLIS, G.; YANG, K.; LAMBRIS, J. D. Complement: a key system for immune surveillance and homeostasis. **Nature Immunology**, v. 11, p. 785-797, 2010

ROUMENINA, L. T.; POPOV, K. T.; BUREEVA, S. V.; KOJOUHAROVA, M.; GADJEVA, M.; RABHERU, S.; THAKRAR, R.; KAPLUN, A.; KISHORE, U. Interaction of the globular domain of human C1q with *Salmonella typhimurium* lipopolysaccharide. **Biochimica et Biophysica Acta**, v. 1784, p. 1271-1276, 2008

SANTI-FERRARA, G. I.; FERNANDES-PEDROSA, M.; JUNQUEIRA-DE-AZEVEDO, I.; GONÇALVES-DE-ANDRADE, R. M.; PORTARO, F. C. V.; MANZONI-DE-ALMEIDA, D.; MURAKAMI, M. T.; ARNI, R. K.; VAN DER BERG, C. W.; LEE HO, P.; TAMBOURGI, D. V. SMase II, a new sphingomyelinase D from *Loxosceles laeta* venom gland: Molecular cloning, expression, function and structural analysis. **Toxicon**, v. 53, p. 743-753, 2009

SEKIGUCHI, R.; FUJITO, N.; NONAKA, M. Evolution of the thioester-containing proteins (TEPs) of the arthropoda, revealed by molecular cloning of TEPs genes from a spider, *Hasarius adansoni*. **Developmental and Comparative Immunology**, v. 36, p. 483-489, 2012

SEKIGUCHI, R.; NONAKA, M. Evolution of the complement system in protostomes revealed by de novo transcriptome analysis of six species of Arthropoda. **Developmental and Comparative Immunology**, v.50. p. 58-67, 2015

SEZERINO, U. M.; ZANNIN, M.; COELHO, L. K.; GONÇALVES, J.; GRANDO, M.; MATTOSINHO, S. G.; CARDOSO, J. L. C.; EICKSTEDT, V. R.; FRANÇA, F. O. S.; BARBARO, K. C.; FAN, H. W. A clinical and epidemiological study of *Loxosceles* spider envenoming in Santa Catarina, Brazil. **Transactions of Royal Society of Tropical Medicine Hygiene**, v. 92, p. 546–548, 1998

SILVEIRA, A. L. Primeiro registro sinantrópico de *Loxosceles laeta* (Nicolet, 1849) (Aranae, Sicariidae) no município do Rio de Janeiro, estado do Rio de Janeiro. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 42, p. 723-726, 2009

SMITH, L. C.; SHIH, C.; DACHENHAUSEN, S. G. Coelomocytes express SpBf, a homologue of factor B, the second component in the sea urchin complement system. **Journal of Immunology**, v. 161, p. 6784-6793, 1998

SOLOMON, K. R.; SHARMA, P.; CHAN, M.; MORRISON, P. T.; FINBERG, R. W. CD109 represents a novel branch of the alpha2-macroglobulin complement gene family. **Gene**, v. 327, p. 171-183, 2004

SUNYER, J. O.; LAMBRIS, J. D. Evolution and diversity of the complement system of poikilothermic vertebrates. **Immunological Reviews**, v. 27, p. 549-564, 1998

SWANSON, D. V.; VETTER, R. S. Loxoscelism. **Clinics in Dermatology**, v. 24, p. 213-221, 2006

TAGAWA, K.; YOSHIHARA, T.; SHIBATA, T.; KITAZAKI, K.; ENDO, Y.; FUJITA, T.; KOSHIBA, T.; KAWABATA, S. Microbe-specific C3b deposition in the horseshoe crab complement system in a C2/Factor B-dependent or –independent manner. **Plos one**, v. 7, e36873, 2012

TAMBOURGI, D. V.; MAGNOLI, F. C.; VON EICKSTEDT, V. R.; BENEDETTI, Z. C.; PETRICEVICH, V. L.; DA SILVA, W. D. Incorporation of a 35-kilodalton purified protein from *Loxosceles intermedia* spider venom transforms human erythrocytes into activators of autologous complement alternative pathway. **Journal of Immunology**, v.155, p. 4459-4466, 1995

TAMBOURGI, D. V.; MAGNOLI, F. C.; VAN DER BERG, C. W.; MORGAN, B. P.; DE ARAUJO, P. S.; ALVES, E. W.; DA SILVA, W. D. Sphingomyelinases in the venom of the spider *Loxosceles intermedia* are responsible for both dermonecrosis and complement–dependent hemolysis. **Biochemical and Biophysical Research Communications**, v. 251, p. 366-373, 1998

TAMBOURGI, D. V.; MORGAN, B. P.; GONÇALVES-DE-ANDRADE, R. M.; MAGNOLI, F. C.; VAN DEN BERG, C. W. *Loxosceles intermedia* spider envenomation induces activation of an endogenous metalloproteinase, resulting in cleavage of glycophorins from erythrocyte surface and facilitating complement-mediated lysis. **Blood**, v. 95, p. 683-691, 2000

TAMBOURGI, D. V.; DE-SOUZA-DA-SILVA, M.; BILLINGTON, S. J.; GONÇALVES-DE-ANDRADE, R. M.; MAGNOLI, F. C.; SONGER, J. G.; VAN DER BERG, C. W. Mechanism of induction of complement susceptibility of erythrocytes by spider and bacterial sphingomyelinases. **Immunology**, v. 107, p. 93-101, 2002

TAMBOURGI, D. V.; PAIXÃO-CALVACANTE, D.; GONÇALVES-DE-ANDRADE, R. M.; FERNANDES-PEDROSA, M.; MAGNOLI, F. C.; MORGAN, B. P.; VAN DER BERG, C. W. *Loxosceles* sphingomyelinase induces complement-dependent dermonecrosis, neutrophil infiltration, and endogenous gelatinase expression. **Journal of Investigative Dermatology**, v. 124, p. 725-731, 2005

TAMBOURGI, D. V.; FERNANDES-PEDROSA, M.; GONÇALVES-DE-ANDRADE R. M.; BILLINGTON, S. J.; GRIFFITHS, M.; VAN DER BERG, C. W. Sphingomyelinases D induce direct association of C1q to the erythrocyte membrane causing complement mediated autologous haemolysis. **Molecular Immunology**, v. 44, p. 576-582, 2007

TAMURA, K.; STECHER, G.; PETERSON, D.; FILIPSKI, A.; KUMAR, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. **Molecular Biology and Evolution**, v.30, p. 2725-2729, 2013

The *C. elegans* Sequencing Consortium. Genome Sequence of the nematode *C. elegans*: a platform for investigating biology. **Science**, v. 282, p. 2012-2018, 1998

VAN MEETEREN, L. A.; FREDERIKS, F.; GIEPMANS, B. N.; FERNANDES-PEDROSA, M.; BILLINGTON, S. J.; JOST, B.H.; TAMBOURGI, D. V.; MOOLENAAR, W. H. Spider and bacterial sphingomyelinases D target cellular lysophosphatidic acid receptors by hydrolyzing lysophosphatidylcholine. **Journal of Biological Chemistry**, v. 279, p. 10833-10836, 2004

VOLANAKIS, J. E.; KAPLAN, M. H. Interaction of C-reactive protein complexes with the complement system. II. Consumption of the guinea pig complement by CRP complexes: requirement for human C1q. **Journal of Immunology**, v. 113, p. 9-17, 1974

WALPORT, M. J. Complement. First of two parts. **The New England Journal of Medicine**, v. 344, p. 1058–1066, 2001

YOSHIZAKI, F. Y.; IKAWA, S.; SATAKE, M.; SATOH, N.; NONAKA, M. Structure and the evolutionary implication of the triplicated complement factor B genes of a urochordate ascidian, *Ciona intestinalis*. **Immunogenetics**, v. 56, p. 930-942, 2005

ZARKADIS, I. K.; MASTELLOS, D.; LAMBRIS, J. Phylogenetic aspects of the complement system. **Developmental and Comparative Immunology**, v. 25, p. 745-762, 2001

ZHONG, L.; ZHANG, F.; CHANG, Y. Gene cloning and function analysis of complement B factor -2 of *Apostichopus japonicus*. **Fish and Shellfish Immunology**, v. 33, p. 504-513, 2012

ZHU, Y.; THANGAMANI, S.; HO, B.; DING, L.D. The ancient origin of the complement system. **The EMBO Journal**, v. 24, p. 382-394, 2005

ZOBEL-THROPP, P., KERINS, A. E., BINFORD, G. J., Sphingomyelinase D in sicariid spider venom is a potent insecticidal toxin. **Toxicon**, v. 60, p.265-271, 2012