



Universidade de Brasília

Faculdade de Medicina

Programa de Pós-Graduação em Ciências Médicas

**Análise da expressão heteróloga de uma serino
protease (RnTrypsin) e de um inibidor de serino
protease (RnKazalPI) salivares de *Rhodnius
neglectus*, vetor da Doença de Chagas**

CAROLINE BARRETO VIEIRA

Brasília - DF
2018

CAROLINE BARRETO VIEIRA

Análise da expressão heteróloga de uma serino protease (RnTrypsin) e de um inibidor de serino protease (RnKazalPI) salivares de *Rhodnius neglectus*, vetor da Doença de Chagas

Orientadora: Dra. Carla Nunes de Araújo

Dissertação de mestrado apresentada ao Programa de Pós-graduação em Ciências Médicas da Universidade de Brasília como requisito parcial à obtenção do grau de Mestre.

Brasília - DF

2018

Trabalho desenvolvido no Laboratório de Interação Patógeno-Hospedeiro, Instituto de Ciências Biológicas da Universidade de Brasília, com o apoio financeiro da Fundação de Apoio à Pesquisa do Distrito Federal (FAPDF, Processo nº 193.001053/2015) e Coordenação de Aperfeiçoamento de Pessoal Nível Superior (CAPES).

“O Senhor é o meu pastor, nada me faltará.”

Salmos 23:1.

AGRADECIMENTOS

À minha orientadora, professora Dra. Carla Nunes de Araújo, uma pessoa brilhante, carinhosa e amiga. Muito obrigada pelo incentivo, paciência e ensinamentos que levarei para sempre comigo. Serei eternamente grata por ter me recebido e me conduzido na vida acadêmica e científica durante esses dois anos.

Aos professores Dr. Jaime Santana, Dra. Izabela Bastos e Dra. Flávia Motta, pelo apoio diário, incentivo e valiosas sugestões.

À todos os colegas do Laboratório de Interação Patógeno-Hospedeiro por serem acolhedores e prestativos, obrigada pelos momentos de aprendizado, reflexão e diversão. E em especial à Yanna Reis, Deborah Martins, Gabriel Santos, Camila Lasse, Marcelle Caminha e Kaio Bentes pelo apoio e ajuda durante as atividades desenvolvidas neste trabalho.

Ao meu marido Wilton Wallace, por ser a pessoa mais incrível que eu já conheci e que sempre acreditou em mim, as vezes até mais do que eu mesma. Pelo amor, compreensão, carinho, incentivo e muitas risadas. Eu amo você!

Aos meus familiares, meu pai Rui Vieira, minha mãe Marly Vieira, minha irmã Ana Sophia Vieira e meu cunhado Reinhold Bezerra, pelo amor, incentivo, risadas e tudo mais que fizeram para tornar esse caminho mais leve e alegre. E ao meu avô Ruy Carlos pela carinhosa e cuidadosa revisão desse trabalho. Amo vocês!

Ao Weverton Carlos e Juliana Batista pelo auxílio na manutenção das colônias de barbeiros no Biotério do IB.

Aos membros da banca examinadora, prof. Dra. Tatiana Amabile de Campos, Dra. Milene Aparecida Andrade e a prof. Dra. Flávia da Silva Nader Motta, por participarem da defesa da dissertação.

Ao CNPq, CAPES e FAPDF pelo financiamento deste trabalho.

Ao Programa de Pós-Graduação em Ciências Médicas da Faculdade de Medicina - UnB pela oportunidade de realização do mestrado.

SUMÁRIO

Agradecimentos	5
Sumário	6
Lista de Abreviaturas e Siglas	8
Lista de Figuras	9
Lista de Tabelas	11
Lista de Anexos	12
Resumo	13
Abstract	14
1. Introdução	15
1.1. Doença de Chagas	15
1.2. Fases da doença de Chagas	16
1.3. Tratamento da doença de Chagas	17
1.4. <i>Trypanosoma cruzi</i> e seu ciclo de vida	18
1.5. Triatomíneos	21
1.5.1. <i>Rhodnius neglectus</i>	23
1.6. Importância das proteínas salivares de artrópodes hematófagos ...	24
1.6.1. Vasodilatadores	27
1.6.2. Inibidores da agregação plaquetária	27
1.6.3. Anticoagulantes	28
1.6.4. Imunosupressores	28
1.7. Estudos das proteínas salivares de animais hematófagos	29
1.7.1. Serino proteases	31
1.7.2. Inibidores de serino proteases do tipo Kazal	32
2. Justificativa	33
3. Objetivos	34
3.1. Objetivo geral	34
3.2. Objetivos específicos	34
4. Metodologia	35
4.1. Seleção das sequências de proteínas salivares de <i>R. neglectus</i> ...	36
4.2. Análise <i>in silico</i>	36
4.3. Clonagem molecular	36
4.4. Transformação de <i>Escherichia coli</i> com pET100/D-TOPO_ RnTrypsin e pET100/D-TOPO_ RnKazalPI	39
4.5. Identificação de colônias de <i>E. coli</i> transformadas com pET100/D- TOPO_ RnKazalPI e análise por eletroforese em gel de agarose ...	40
4.6. Expressão de RnTrypsin e RnKazalPI recombinantes	41
4.7. Purificação da RnTrypsin recombinante	42

4.8. Eletroforese em SDS-PAGE	43
4.9. <i>Western blot</i>	44
4.10. Produção de anticorpos policlonais contra a RnTrypsin	45
4.11. Colônia de triatomíneos	46
4.12. Obtenção da saliva	46
4.13. Teste de reatividade dos anticorpos com a saliva de <i>R. neglectus</i>	47
5. Resultados	48
5.1. Serino protease tipo tripsina de <i>R. neglectus</i> (RnTrypsin)	48
5.1.1. Análise <i>in silico</i>	48
5.1.2. Expressão da RnTrypsin	50
5.1.3. Purificação da RnTrypsin	60
5.1.4. Produção de anticorpos	61
5.1.5. Perfil da saliva de <i>R. neglectus</i>	62
5.2. Inibidor de serino protease tipo Kazal de <i>R. neglectus</i>	63
5.2.1. Análise <i>in silico</i>	63
5.2.2. Expressão da RnKazalPI	65
6. Discussão	72
7. Conclusão	76
8. Perspectivas	77
9. Referências bibliográficas	78
10. Anexos	94

LISTA DE ABREVIATURAS E SIGLAS

BLAST - *Basic Local Alignment Search Tool* (ferramenta básica de pesquisa de alinhamento local)

cDNA - DNA complementar

DNA - *deoxyribonucleic acid* (ácido desoxirribonucleico)

h - hora

IPTG - Isopropyl β -D-1- thiogalactopyranoside

kDa - Kilodalton

LB - meio de cultura Luria Bertani

M - molar

min - minuto

μ g - micrograma

μ L - microlitro

mM - milimolar

nm - nanômetro

OBP - *odorant-binding protein* (proteína ligante de odorante)

OD 600 - densidade óptica a 600 nm

pb - pares de bases

PBS - *Phosphate-Buffered Saline* (solução salina tamponada com fosfato)

PCR - *Polimerase Chain Reaction* (reação em cadeia da polimerase)

RNA_m - *Ribonucleic Acid messenger* (ácido ribonucleico mensageiro)

rpm - rotações por minuto

SDS-PAGE - *Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis* (gel de poliacrilamida contendo duodecil sulfato de sódio)

V - volts

LISTA DE FIGURAS

Figura 1. Tripomastigota metacíclico de <i>Trypanosoma cruzi</i>	18
Figura 2. Ciclo de vida do <i>Trypanosoma cruzi</i> no vetor triatomíneo e no hospedeiro mamífero.	19
Figura 3. Estágios de desenvolvimento do triatomíneo <i>Rhodnius neglectus</i> , mostrando os cinco estágios ninfais e o adulto.	21
Figura 4. Ninfa de quinto estágio de <i>Rhodnius neglectus</i> alimentada.	22
Figura 5. Glândulas salivares de <i>R. neglectus</i> . Apresentam coloração avermelhada devido à presença de hemeproteínas.	23
Figura 6. Representação esquemática da metodologia empregada para realizar e analisar a expressão da e RnTrypsin recombinante.	35
Figura 7. Representação esquemática da metodologia empregada para realizar e analisar a expressão da RnKazalPI recombinante.	35
Figura 8. Representações esquemáticas do vetor de expressão pET100/D-TOPO.	38
Figura 9. Alinhamento da sequência de aminoácidos da RnTrypsin com serino proteases putativas.	50
Figura 10. Expressão da proteína recombinante RnTrypsin nas frações solúvel e insolúvel de BL21(DE3) induzidas à 37°C por 3 h com 1 mM IPTG.	51
Figura 11. Expressão da proteína recombinante RnTrypsin nas frações solúvel e insolúvel de BL21(DE3) induzidas à 25°C por 22 h com 0,4 mM IPTG.	52
Figura 12. Análise da expressão da proteína recombinante RnTrypsin nas frações solúvel e insolúvel de BL21(DE3) induzidas à 20°C por 3 à 5 h com 0,5 mM IPTG.	53
Figura 13. Expressão da proteína recombinante RnTrypsin nas frações solúvel e insolúvel de BL21(DE3) induzidas à 20°C por 16 h com 0,5 mM IPTG.	54
Figura 14. Expressão da proteína recombinante RnTrypsin nas frações solúvel e insolúvel de BL21(DE3) induzidas à 20°C por 5 h com 0,5 mM IPTG e à 37°C por 5 h com 1 mM IPTG.	55
Figura 15. Expressão da proteína recombinante RnTrypsin nas frações solúvel e insolúvel de BL21(DE3) induzidas à 18°C por 24 e 40 h com 0,4 mM IPTG.	56
Figura 16. Expressão da proteína recombinante RnTrypsin nas frações solúvel e insolúvel de BL21(DE3) induzidas à 15°C por 20 h com 0,2 e 0,4 mM IPTG.	56

Figura 17. Expressão da proteína recombinante RnTrypsin nas frações solúvel e insolúvel de BL21(DE3)pLysS induzidas à 37°C por 3 h com 1 mM IPTG e à 20°C por 16 h com 0,5 mM IPTG.	57
Figura 18. Expressão da proteína recombinante RnTrypsin nas frações solúvel e insolúvel de Rosetta(DE3)pLysS induzidas à 37°C por 3 h com 1 mM IPTG, à 25°C por 22 h com 0,4 mM IPTG, à 20°C por 16 h com 0,5 mM IPTG e à 15°C por 22 h com 0,2 e 0,4 mM IPTG.	58
Figura 19. Purificação dos corpos de inclusão da fração insolúvel de RnTrypsin por cromatografia de afinidade em níquel, utilizando His-tag; e análise por <i>Western blot</i> , utilizando anticorpo primário anti-His, da proteína recombinante RnTrypsin eluída com 5 mM de imidazol.	60
Figura 20. <i>Western blot</i> dos soros pré-imune e imune anti-RnTrypsin produzidos por 5 camundongos, diluídos 1:400 e 1:100.	61
Figura 21. <i>Western blot</i> dos soros pré-imune e imune anti-RnTrypsin do camundongo nº 2, diluídos 1:200, e do anticorpo Anti-His, diluído 1:2000, utilizados para avaliar reatividade com a eluição na fração contendo 5 mM de Imidazol resultante da purificação.	62
Figura 22. Expressão das proteínas totais do extrato salivar de <i>R. neglectus</i> em gel de poliacrilamida 12%.	62
Figura 23. Alinhamento da sequência de aminoácidos da RnKazalPI com inibidores de serino proteases tipo Kazal putativos.	65
Figura 24. Expressão da proteína recombinante RnKazalPI nas frações solúvel e insolúvel de BL21(DE3) induzidas à 37°C por 3 h com 1 mM IPTG.	65
Figura 25. Expressão da proteína recombinante RnKazalPI nas frações solúvel e insolúvel de BL21(DE3) induzidas à 20°C por 5 h com 0,5 mM IPTG.	66
Figura 26. Expressão da proteína recombinante RnKazalPI nas frações solúvel e insolúvel de BL21(DE3) induzidas à 20°C por 16 h com 0,5 mM IPTG.	67
Figura 27. Expressão da proteína recombinante RnKazalPI nas frações solúvel e insolúvel de BL21(DE3)pLysS induzidas à 37°C por 3 h com 1 mM IPTG e à 20°C por 16 h com 0,5 mM IPTG.	68
Figura 28. Expressão da proteína recombinante RnKazalPI nas frações solúvel e insolúvel de Rosetta(DE3)pLysS induzidas à 37°C por 3 h com 1 mM IPTG e à 20°C por 16 h com 0,5 mM IPTG.	69
Figura 29. Identificação de colônias transformadas de <i>E. coli</i> Rosetta(DE3)pLysS transformada com o plasmídeo pET100/D-TOPO_RnKazalPI.	70
Figura 30. Expressão da proteína recombinante RnKazalPI nas frações solúvel e insolúvel de Rosetta(DE3)pLysS (mesmo clone da K3 - vide figura 25) induzidas à 18°C por 3 h com 1 mM IPTG.	70

LISTA DE TABELAS

Tabela 1. Características dos insertos utilizados.	39
Tabela 2. Características das linhagens de <i>E. coli</i> utilizadas nesse estudo.	48
Tabela 3. Resultados da análise <i>in silico</i> da serino protease tipo tripsina de <i>R. neglectus</i>	48
Tabela 4. Condições de expressão da proteína RnTrypsin recombinante em sistema de expressão bacteriano.	59
Tabela 5. Resultados da análise <i>in silico</i> do inibidor de serino protease tipo Kazal de <i>R. neglectus</i>	64
Tabela 6. Condições de expressão da proteína RnKazalPI recombinante em sistema de expressão bacteriano.	71

LISTA DE ANEXOS

Anexo 1. Códon-otimização da RnTrypsin para expressão em sistema bacteriano.	95
Anexo 2. Códon-otimização da RnKazalPI para expressão em sistema bacteriano.	96
Anexo 3. Clonagem da RnTrypsin no vetor pET100/D-TOPO.	97
Anexo 4. Clonagem da RnKazalPI no vetor pET100/D-TOPO.	98
Anexo 5. Declaração do Comitê de Ética no Uso Animal da Universidade de Brasília - UnB.	99
Anexo 6. Resultados do software SignalP para a RnTrypsin.	100
Anexo 7. Resultados do software TargetP para a RnTrypsin.	101
Anexo 8. Resultados do software ProtParam para a RnTrypsin.	102
Anexo 9. Resultados do software NetPhos para a RnTrypsin.	103
Anexo 10. Resultados do software NetNGlyc para a RnTrypsin.	104
Anexo 11. Resultados do software NetOGlyc para a RnTrypsin.	105
Anexo 12. Resultados do software Protein Blast para a RnTrypsin.	106
Anexo 13. Resultados do software SignalP para a RnKazalPI.	107
Anexo 14. Resultados do software TargetP para a RnKazalPI.	108
Anexo 15. Resultados do software ProtParam para a RnKazalPI.	109
Anexo 16. Resultados do software NetPhos para a RnKazalPI.	110
Anexo 17. Resultados do software NetNGlyc para a RnKazalPI.	111
Anexo 18. Resultados do software NetOGlyc para a RnKazalPI.	112
Anexo 19. Resultados do software Protein Blast para a RnKazalPI	113

RESUMO

O triatomíneo *Rhodnius neglectus* (Hemiptera: Reduviidae) é um vetor potencial do protozoário *Trypanosoma cruzi*, agente etiológico da doença de Chagas. O *R. neglectus* apresenta em sua saliva moléculas importantes que neutralizam os mecanismos hemostáticos do hospedeiro, auxiliando na alimentação sanguínea e, de forma indireta, na transmissão do protozoário. Assim, a saliva desse inseto hematófago se torna um relevante alvo de pesquisas devido à presença de moléculas com elevado potencial farmacológico, como por exemplo serino proteases e inibidores de serino protease. Este estudo teve por objetivo a expressão heteróloga da serino protease do tipo tripsina (RnTrypsin) e do inibidor de serino protease tipo Kazal (RnKazalPI) salivares de *R. neglectus*. As sequências nucleotídicas das proteínas foram obtidas a partir do transcriptoma de *R. neglectus*, clonadas no vetor pET100/D-TOPO e expressas em bactérias *Escherichia coli* da linhagem BL21(DE3), BL21(DE3)pLysS e Rosetta(DE3)pLysS. A serino protease do tipo tripsina (RnTrypsin) recombinante foi observada somente na fração insolúvel após a lise das bactérias, sendo então purificada por cromatografia de afinidade e usada para a produção de anticorpos policlonais em camundongos. Após a otimização do protocolo de expressão, o inibidor de serino protease tipo Kazal (RnKazalPI) foi expresso na fração solúvel do lisado bacteriano. Como perspectiva, é pretendido utilizar o sistema de expressão de proteínas recombinantes em células de inseto/baculovírus para expressar a serino protease do tipo tripsina (RnTrypsin) recombinante na fração solúvel, e também determinar a caracterização do inibidor de serino protease tipo Kazal (RnKazalPI) recombinante solúvel, buscando contribuir na compreensão das interações entre vetor e hospedeiro e potenciais atividades farmacológicas.

Palavras-chaves: Triatomíneo, transcriptoma, proteínas salivares, proteína recombinante, *Escherichia coli*.

Abstract

The triatomine *Rhodnius neglectus* (Hemiptera: Reduviidae) is a potential vector of the protozoan *Trypanosoma cruzi*, etiologic agent of Chagas' disease. *R. neglectus* presents in its saliva important molecules that neutralize the haemostatic mechanisms of the host, aiding in the blood supply and, indirectly, in the transmission of the protozoan. Thus, the saliva of this hematophagous insect becomes a relevant research target due to the presence of molecules with high pharmacological potential, for example serine proteases and serine protease inhibitors. This study aimed at the heterologous expression of the trypsin-like serine protease (RnTrypsin) and the Kazal type serine protease inhibitor (RnKazalPI) of *R. neglectus*. The nucleotide sequences of the proteins were obtained from the transcriptome of *R. neglectus*, cloned in the vector pET100/D-TOPO and expressed in *Escherichia coli* bacteria of line BL21(DE3), BL(DE3)pLysS and Rosetta(DE3)pLysS. The recombinant trypsin-like serine protease (RnTrypsin) was observed only in the insoluble fraction after bacterial lysis, and was then purified by affinity chromatography and used for the production of polyclonal antibodies in mice. After optimization of the expression protocol, the Kazal type serine protease inhibitor (RnKazalPI) was expressed in the soluble fraction of the bacterial lysate. As a perspective, it is intended to use the recombinant protein expression system in insect/baculovirus cells to express the recombinant trypsin-like serine protease (RnTrypsin) in the soluble fraction, and also to determine the characterization of the recombinant Kazal type serine protease inhibitor (RnKazalPI), seeking to contribute to the understanding of the interactions between vector and host and potential pharmacological activities.

Keywords: Triatomine, transcriptome, salivary proteins, recombinant protein, *Escherichia coli*.

1. Introdução

1.1. Doença de Chagas

Presente em 149 países, as doenças tropicais negligenciadas englobam um grupo diversificado de doenças transmissíveis que afetam mais de um bilhão de pessoas localizadas em regiões tropicais e subtropicais do globo (Savioli e Daumerie, 2013). Essas doenças possuem características em comum que são: dispersão geográfica específica que afeta populações com baixo nível socioeconômico, alta morbidade e consequente mortalidade com significativo impacto negativo socioeconômico, dificuldade de diagnóstico, tratamento e controle, além de disponibilidade limitada de recursos e prioridade política (Briceño-León e Méndez Galván, 2007; Daumerie e Savioli, 2010; WHO, 2015). Entre as doenças tropicais negligenciadas encontra-se a doença de Chagas, que está entre as doenças parasitárias mais comuns no mundo (WHO, 2015).

Também conhecida como tripanossomíase americana, a doença é endêmica em 21 países latino-americanos (Salud, 2006; Hotez *et al.*, 2008). Entretanto, devido ao aumento dos movimentos populacionais entre a América Latina e outros continentes, cada vez mais vem sendo detectado a presença da doença em países não endêmicos, como nos Estados Unidos, Canadá, em muitos países europeus e em alguns países do Pacífico Ocidental (Gascon *et al.*, 2010; Bern *et al.*, 2011; Navarro *et al.*, 2012; Jackson *et al.*, 2014). Segundo estimativas da Organização Mundial de Saúde, o número de pessoas infectadas atualmente é de aproximadamente 8 milhões, e acredita-se que mais de 10.000 mortes ocorram todos os anos devido às manifestações clínicas da doença (WHO, 2016).

Carlos Ribeiro Justiniano Chagas foi um médico brasileiro que em 1909 produziu um trabalho único e excepcional na história da medicina. Ele descobriu e descreveu uma nova espécie de protozoário flagelado (agente etiológico) encontrado nos intestinos de insetos hematófagos (vetor), no sangue de animais domésticos e no sangue de uma menina (reservatórios), além de mecanismos de infecção e vários estágios clínicos da doença. Esse protozoário é o *Trypanosoma cruzi*, os vetores são insetos conhecidos como

triatomíneos e a doença foi denominada doença de Chagas em homenagem ao seu descobridor (Chagas, 1909; WHO, 2002; Araujo-Jorge *et al.*, 2017).

Embora a doença de Chagas tenha sido descoberta apenas em 1909, sua história natural começou há muitos anos como uma doença enzoótica de animais silvestres, que ainda persiste no Brasil, principalmente em áreas como a Amazônia. Sua transmissão para os seres humanos se deu acidentalmente como uma antropozoonose quando os homens invadiram ecótopos silvestres através do desmatamento para fins agrícolas, pecuários, entre outros. Devido ao deslocamento de animais silvestres, os insetos hematófagos incapazes de encontrar fonte de alimento ali começaram a buscar em outros locais, colonizando áreas ao redor e dentro de habitações humanas. Alguns se adaptaram a esse novo nicho, alimentando-se de sangue de animais domésticos e também de humanos (Coura, 2007; Coura e Borges-Pereira, 2010).

1.2. Fases da doença de Chagas

A doença de Chagas manifesta-se em 2 fases. Inicialmente, há uma fase aguda, com duração de cerca de 4 a 8 semanas, com alta parasitemia microscopicamente detectável. Durante essa fase, a maioria das pessoas infectadas são assintomáticas ou oligossintomáticas, apresentando sintomas inespecíficos como febre, mal estar, cefaléia, dor muscular, erupção cutânea, entre outros (Rassi Jr *et al.*, 2006; Coura *et al.*, 2007; Rassi Jr *et al.*, 2010). No local da picada, pode ocorrer o desenvolvimento de um processo inflamatório denominado chagoma de inoculação, que é caracterizado como uma lesão cutânea arredondada, indolor ou pouco dolorosa, endurecida e que pode ou não ulcerar (Mazza e Freire, 1940). Se a entrada do parasito se der pela conjuntiva do olho, pode ocorrer o sinal de Romaña, que é um edema não doloroso das pálpebras superiores e inferiores. Depois desse período inicial, a infecção entra na fase crônica, com parasitas escondidos nos tecidos para os quais têm tropismo, especialmente nos músculos cardíaco e sistema digestório. Pessoas com infecção pelo protozoário mas sem sinais ou sintomas da doença, são consideradas como tendo a forma “indeterminada” (Rassi Jr *et al.*, 2006; Coura *et al.*, 2007; Rassi Jr *et al.*,

2010). Aproximadamente um terço dos indivíduos indeterminados progride para a forma sintomática da fase crônica, que consiste em anormalidades no sistema cardiovascular (podendo resultar em insuficiência cardíaca e morte), no gastrointestinal, ou nos dois ao mesmo tempo (Teixeira *et al.*, 2011; Duran-Rehbein *et al.*, 2014). A taxa de mortalidade em 10 anos pode variar entre 10% a 80%, dependendo dos danos sofridos (WHO, 2015). Entre os fatores que determinam a progressão da doença de Chagas devem ser considerados o inóculo inicial do *T. cruzi*, reinfecções, as características biológicas das cepas do parasito, o seu histotropismo e a resposta celular e humoral do hospedeiro (Coura *et al.*, 2007).

1.3. Tratamento da doença de Chagas

A doença de Chagas pode ser clinicamente curável em determinados pacientes durante sua fase aguda e indeterminada (De Andrade *et al.*, 1996; Estani *et al.*, 1998). Porém infelizmente a maioria das pessoas infectadas permanece sem diagnóstico até os estágios posteriores (Chatelain, 2015). Apesar de a doença de Chagas ter sido descoberta há mais de um século e o primeiro tratamento há mais de quarenta anos, atualmente tem-se apenas dois fármacos disponíveis: o nifurtimox e o benzonidazol, sendo este último o único disponível no Brasil. O índice de cura varia muito de acordo com as características individuais de cada pessoa, a fase da doença e também com fatores genéticos do parasito. Na fase aguda as chances de cura são de 70 a 80%. Entretanto, na fase crônica as chances são mais baixas (aproximadamente 10%), tendo o tratamento farmacológico apenas para fins paliativos, para reduzir os sintomas da doença e a proliferação do parasito. Devido ao tempo de tratamento ser prolongado e existir alta incidência de reações adversas, como dor muscular, tonturas, náuseas, até graves distúrbios da hematopoiese, ocorre uma diminuição na adesão e no uso do medicamento, podendo levar até ao abandono da terapêutica (Dias *et al.*, 2016; Alexandre *et al.*, 2018).

1.4. *Trypanosoma cruzi* e seu ciclo de vida

O *Trypanosoma cruzi* é uma espécie do subgênero *Schizotrypanum*, pertencente ao gênero *Trypanosoma*, família Trypanosomatidae, ordem Trypanosomatida, classe Kinetoplastea, subfilo Glycomonada, filo Euglenozoa, sub-reino Eozoa, reino Protozoa e domínio Eukaryota (Hoare, 1972; Cavalier-Smith, 2016). O subgênero *Schizotrypanum* foi adotado para tripanossomos que se multiplicam em estágios intracelulares em vertebrados. Portanto, o nome taxonômico completo é *Trypanosoma (Schizotrypanum) cruzi* (WHO, 2002). Ele é um protozoário flagelado que possui uma única mitocôndria, denominada cinetoplasto, que contém material genético condensado (Stevens et al., 2001; Galvão, 2014) (Figura 1).



Figura 1. Tripomastigota metacíclico de *Trypanosoma cruzi*. FONTE: Imagem cedida por Natália Gil, arquivo pessoal, 2018.

O ciclo de vida do *T. cruzi* é complexo e possui três estágios evolutivos morfológicos e fisiológicos distintos principais em insetos vetores e hospedeiros mamíferos. Esses estágios são denominados amastigotas, epimastigotas e tripomastigotas (Gourbière et al., 2012; Kessler et al., 2017). No inseto, o ciclo de vida se inicia quando ele se alimenta de sangue de um vertebrado contendo tripomastigotas sanguíneos. Quando eles chegam ao intestino do inseto, transformam-se em epimastigotas. No intestino eles se multiplicam e migram para o reto, onde se diferenciam em tripomastigotas metacíclicos, em um processo chamado de metaciclogênese. Quando o inseto infectado for se alimentar novamente, os tripomastigotas metacíclicos

podem ser eliminados através da excreta do inseto e podem invadir a corrente sanguínea do hospedeiro vertebrado através de fissuras na pele, preexistentes ou resultantes da picada, ou também através de tecidos mucosos (Garcia et al., 2007; Gonçalves et al., 2018). Uma vez dentro do hospedeiro vertebrado, os tripomastigotas metacíclicos infectam as células nucleadas como os macrófagos, as células musculares, entre outras. Dentro da célula, os protozoários se transformam em amastigotas, se multiplicam, transformam-se novamente em tripomastigotas e lisam a célula infectada, invadindo células e tecidos adjacentes e espalhando-se pelo sangue e pelo sistema linfático (Brener, 1971; Teixeira *et al.*, 2011; Tomlinson *et al.*, 1995; Deane *et al.*, 1984).

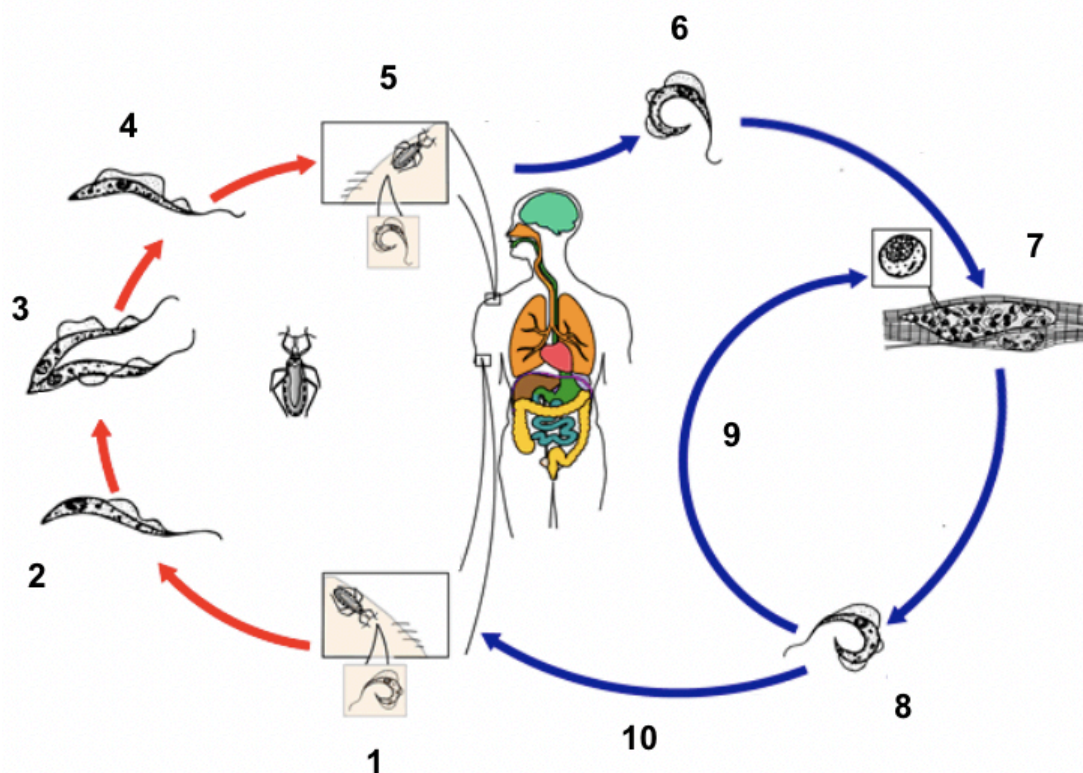


Figura 2. Ciclo de vida do *Trypanosoma cruzi* no vetor triatomíneo e no hospedeiro mamífero. 1) O triatomíneo é infectado quando se alimenta de sangue humano ou animal contendo tripomastigotas sanguíneos. 2) Os parasitos ingeridos se transformam em epimastigotas no intestino médio do vetor. 3) Em seguida, se multiplicam. 4) Quando atingem a porção final do intestino, eles se diferenciam em tripomastigotas metacíclicos infecciosos. 5) Quando o inseto infectado realiza um novo repasto sanguíneo, ele pode liberar os parasitas em suas fezes perto do local da picada. 6) Os tripomastigotas entram no hospedeiro através do local da picada ou de membranas mucosas intactas, como a conjuntiva. 7) Dentro do hospedeiro, os

tripomastigotas metacíclicos invadem células nucleadas próximas ao local da inoculação, onde se diferenciam em amastigotas e multiplicam-se por fissão binária. 8) Em seguida, se diferenciam em tripomastigotas, e depois são liberados na circulação. 9) Os tripomastigotas infectam células de uma variedade de tecidos. 10) O ciclo reinicia-se quando os parasitos circulantes são ingeridos por outro triatomíneo. FONTE: CDC, 2018 (modificado).

O *Trypanosoma cruzi* é transmitido para mais de 150 espécies de mamíferos domésticos e silvestres, além dos seres humanos (Rassi Jr et al., 2010). Alguns animais silvestres podem ter um papel diferenciado na epidemiologia da doença. Um exemplo é o *Didelphis marsupialis*, um dos mais conhecidos reservatórios do *T. cruzi* e que circula do ambiente silvestre para áreas em torno de casas, e até mesmo dentro delas, em busca de alimento. Esse animal não apenas carrega o parasito em sua corrente sanguínea, mas também o abriga em suas glândulas odoríferas, podendo assim transmiti-lo diretamente à outros animais e aos humanos (Deane et al., 1984; Roque e Jansen, 2010). Os animais domésticos como cães e gatos também exercem papel importante na epidemiologia da doença, visto que podem invadir o ambiente selvagem para caçar, infectando-se e trazendo a infecção para lares e áreas ao redor das casas (Coura, 2007). Ter uma multiplicidade de reservatórios pode favorecer a manutenção e a dispersão do patógeno (Begon, 2008). As aves, os répteis, e os anfíbios são refratários à infecção pelo protozoário porque o sangue desses animais, mediado pelo sistema complemento, tem um efeito lítico sob o parasito (WHO, 2002). Além da infecção através dos vetores, existem outros mecanismos ditos principais para a transmissão do parasito: pela via oral através de alimentos contaminados, transfusão sanguínea, e através da placenta ou durante o nascimento. Outros mecanismos, ditos secundários, são: transplante de órgãos, manipulação de animais infectados, sexualmente através de feridas ou fluidos, através de acidentes laboratoriais ou criminalmente (Coura et al., 2007; Coura, 2007).

1.5. Triatomíneos

O vetor da doença de Chagas é um inseto conhecido como triatomíneo. Eles pertencem à ordem Hemiptera, subordem Heteroptera, família Reduviidae e subfamília Triatominae (Lent e Wygodzinsky, 1979). Atualmente, são descritas 153 espécies de triatomíneos (Da Rosa et al., 2017) distribuídos em 18 gêneros, que são organizados em 5 tribos: Aberproseniini, Bolboderini, Cavernicolini, Rhodiniini e Triatomini (Galvão, 2014).

A maioria das espécies de triatomíneos possui hábitos silvestres, vivendo nas copas das palmeiras, em pilhas de rochas, ninhos de pássaros, abrigos de animais como gambás, morcegos e roedores, entre outros. Porém, alguns colonizam moradias rurais e tem sido reportado a presença desses insetos também em apartamentos e casas (Monteiro *et al.*, 2000; WHO, 2002; Rodrigues *et al.*, 2014). As espécies do gênero *Rhodnius* estão associadas principalmente às palmeiras. O gênero *Panstrongylus* é predominantemente associado à tocas e cavidades de árvores, enquanto que o gênero *Triatoma* é fortemente associado à locais rochosos ou tocas de roedores (Gaunt e Miles, 2000).

Os triatomíneos são insetos hematófagos obrigatórios, e necessitam do sangue para sua sobrevivência, crescimento e reprodução (Gubler, 1998; Hecht et al., 2006; Lazzari et al., 2013). Após a eclosão dos ovos, eles passam por cinco estágios ninfais antes de chegarem ao estágio adulto (WHO, 2002), conforme demonstrado na Figura 3.



Figura 3. Estágios de desenvolvimento do triatomíneo *Rhodnius neglectus*, mostrando os cinco estágios ninfais e o adulto. FONTE: Arquivo pessoal, 2018.

As ninfas podem obter uma quantidade de sangue até 12 vezes seu peso corporal não alimentado, enquanto os insetos adultos raramente adquirem uma quantidade de três vezes seu peso corporal sem alimento (Lehane, 2005). A Figura 4 mostra uma ninfa de quinto estágio de *Rhodnius neglectus* após a alimentação em sangue.



Figura 4. Ninfa de quinto estágio de *Rhodnius neglectus* alimentada. FONTE: Arquivo pessoal, 2018.

Cada estágio ninfal requer apenas uma refeição de sangue completa para desencadear a muda para o próximo estágio. O desenvolvimento completo do ovo ao adulto leva de quatro meses a um ano para ser concluído, dependendo da espécie, de fatores ambientais, da disponibilidade de alimento, entre outros (Lehane, 2005).

A detecção da fonte de alimento é obtida através da detecção de correntes de ar com odores distintos, vapor de água e calor (Lazzari et al., 2013). Os insetos não se alimentam aleatoriamente. Além de seu papel como sinal de orientação, o calor também é utilizado para localizar os vasos sanguíneos escondidos sob a pele do hospedeiro (Ferreira et al., 2007), que é feito através de termorreceptores localizados principalmente em suas antenas (Wigglesworth e Gillett, 1934; Flores e Lazzari, 1996).

Todas as espécies de triatomíneos são consideradas potenciais vetores do *T. cruzi* (Otálora-Luna et al., 2015). No entanto, dentre as 153 espécies que já foram descritas até o momento (Ravazi et al., 2017), apenas cerca de vinte delas são consideradas relevantes epidemiologicamente devido à comprovada participação na transmissão do protozoário (Filée et al., 2015; Otálora-Luna et al., 2015). As espécies reconhecidas como vetores

principais na América Latina são *Triatoma infestans*, *Rhodnius prolixus*, *Triatoma dimidiata* e *Triatoma brasiliensis* (Lent e Wygodzinsky, 1979; Barges *et al.*, 2010; Rabinovich *et al.*, 2011; Stevens *et al.*, 2011).

1.5.1. *Rhodnius neglectus*

A espécie *Rhodnius neglectus* também exerce um papel importante na transmissão do *T. cruzi*, apesar de ser considerada como um vetor secundário (Barreto-Santana *et al.*, 2011). No Brasil, essa espécie está amplamente distribuída, sendo encontrada em Goiás, Bahia, Maranhão, Mato Grosso, Mato Grosso do Sul, Minas Gerais, Paraná, Pernambuco, Piauí, Rio de Janeiro, São Paulo, Tocantins e Distrito Federal (Lent e Wygodzinsky, 1979; Gurgel-Gonçalves *et al.*, 2012). Essa espécie é predominantemente silvestre, porém já foi relatada em ambiente peri e intradomiciliares nos estados de Goiás, São Paulo, Tocantins, entre outros (Rodrigues *et al.*, 2014; Gurgel-Gonçalves *et al.*, 2008; Oliveira e Silva, 2007).

As glândulas salivares desses insetos têm grande diversidade quanto ao número, tamanho e forma nas diferentes espécies de triatomíneos. Em geral, localizam-se na cavidade torácica, junto à parte inicial do tubo digestivo. Em todas as espécies de triatomíneos são encontrados três pares de glândulas salivares, com exceção do gênero *Rhodnius*, que só possui dois pares (Carcavallo *et al.*, 1997; Lacombe, 1999; Jurberg *et al.*, 2014). As glândulas pertencentes ao par principal são alongadas e avermelhadas devido à presença de nitroforinas, que são proteínas com um grupo heme férrico, responsável por essa coloração (Champagne *et al.*, 1995; Pacheco, 2014; De Oliveira *et al.*, 2017) (Figura 5).

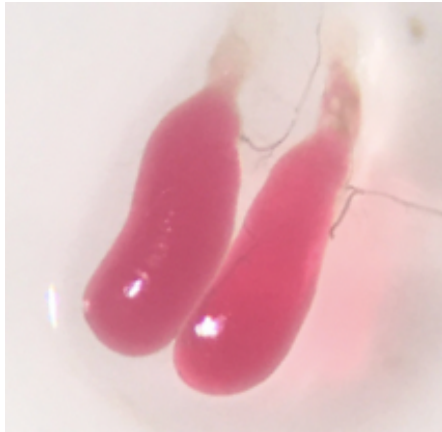


Figura 5. Glândulas salivares de *R. neglectus*. Apresentam coloração avermelhada devido à presença de hemeproteínas. FONTE: Arquivo pessoal, 2018.

Para os triatomíneos, a saliva é de grande importância para sua alimentação (WHO, 2002). Devido ao seu comportamento hematofágico, esses insetos necessitam contrapor o sistema hemostático do hospedeiro para conseguirem se alimentar. Esse sistema inclui várias vias redundantes voltadas para manter o equilíbrio do organismo (Ribeiro, 1987; Ribeiro, 1995; Ribeiro e Arca, 2009).

1.6. Importância das proteínas salivares de artrópodes hematófagos

A hemostasia é um conjunto de processos fisiológicos dinâmicos, complexos e multifatoriais, incluindo a vasoconstrição, a agregação plaquetária e a coagulação sanguínea, que juntos mantêm o equilíbrio do fluxo sanguíneo. Esse sistema é altamente sofisticado, eficiente e inclui várias vias redundantes. Em um indivíduo saudável e com um sistema vascular intacto, o fluxo sanguíneo é contínuo e solúvel (Marder et al., 2012; Hoffman et al., 2013). Em condições fisiológicas normais, o tônus vascular é mantido pelo equilíbrio entre mediadores vasodilatadores e vasoconstritores que regulam a pressão arterial. Em caso de trauma vascular, os componentes celulares do sangue estão prontos para uma resposta rápida e efetiva, desencadeando uma vasoconstrição, que é um estreitamento do vaso sanguíneo devido a uma resposta contrátil da musculatura lisa,

reduzindo o fluxo sanguíneo na região lesionada (Kini et al., 2010; Marder et al., 2012).

Quando a camada de células endoteliais dos vasos sanguíneos é danificada, as plaquetas reconhecem o Fator von Willebrand (vWF) e o colágeno subendotelial exposto por meio de receptores de superfície (como glicoproteínas GPIIb, GPIIb/IIIa e a integrina $\alpha_2\beta_1$). A ligação entre os receptores plaquetários GPIIb e o fator vWF medeia a ligação plaquetária, permitindo a interação do receptor GPIIb/IIIa com o colágeno. Essa interação promove a liberação de mediadores solúveis, como adenosina difosfato (ADP) e tromboxano A₂ (TXA₂), recrutando mais plaquetas para o sítio lesionado (Nieswandt e Watson, 2003; Furie e Furie, 2005; Watson et al., 2005; Varga-Szabo et al., 2008; Gardiner e Andrews, 2014). Ao mesmo tempo, o fator tecidual (TF) desencadeia localmente a formação de trombina que também contribui para a ativação plaquetária (Mackman et al., 2007; Varga-Szabo et al., 2008). Os mediadores solúveis liberados pelas plaquetas ativas geram múltiplas respostas. Além da ativação de mais plaquetas, eles recrutam leucócitos e induzem a secreção de moléculas pró-inflamatórias, contribuindo para o processo de inflamação local (Coughlin, 2005; Jennings, 2009; Jenne e Kubes, 2015).

A vasoconstrição, a agregação plaquetária e a coagulação sanguínea são processos que ocorrem concomitantemente (Versteeg et al., 2013). A coagulação tem início com a exposição da glicoproteína de membrana denominada Fator Tecidual (FT), que funciona como um receptor para o fator VII da coagulação, que quando ativado forma o fator VIIa, que por sua vez ativa os fatores IX e X, os quais são convertidos em fatores IXa e Xa. O fator Xa, na presença de Ca^{2+} e do fator ativado (FVa), converte protrombina em trombina (Ariëns et al., 2002; Bae et al., 2010). Como etapa final, a trombina ativa os fatores V, VIII e XIII da cascata de coagulação e também converte o fibrinogênio em monômeros de fibrina, que finalmente formam um coágulo de fibrina polimerizado elástico e estável (Ariëns et al., 2002; Crawley et al., 2007). A ocorrência de falhas durante o processo de coagulação pode levar a hemorragia ou ao desenvolvimento de trombos que podem dificultar ou

obstruir a circulação sanguínea normal (Rezende, 2010; Spronk *et al.*, 2013; Berger *et al.*, 2014).

Os insetos hematófagos transmissores de microrganismos não são considerados importantes somente devido à transmissão de doenças. Os avanços em pesquisas dos compostos presentes na saliva desses artrópodes têm mostrado que se trata de um potente fluido farmacologicamente ativo que já foi descrito como possuidor de diversas atividades anti-hemostáticas (Ribeiro, 1995; Champagne, 2004; Champagne e Disorders, 2004; Valenzuela *et al.*, 2004; Morita *et al.*, 2006; Ribeiro e Arca, 2009). Além disso, a saliva desses vetores facilita a transmissão de patógenos (Titus e Ribeiro, 1988; Mejia *et al.*, 2006; Oliveira *et al.*, 2006; Gomes *et al.*, 2008; Ribeiro e Arca, 2009).

Para que a alimentação ocorra, o inseto necessita transpor a pele do hospedeiro para encontrar o vaso sanguíneo, e nesse processo, causa lesões e extravasamento de sangue no local, perturbando o equilíbrio hemostático. O organismo do hospedeiro então responde à agressão rapidamente com vasoconstrição, agregação plaquetária, coagulação sanguínea e ativação da resposta imune (Ribeiro, 1987; Ribeiro, 1995). Esses componentes atuam juntos, contendo a perda sanguínea no local da lesão vascular. Sendo assim, o inseto só consegue manter o fluxo de sangue e se alimentar, se conseguir superar esses obstáculos. Para isso, a saliva do inseto é injetada na pele do hospedeiro no local da picada durante todo o período de alimentação (Soares *et al.*, 2006; Sant'anna *et al.*, 2017). A saliva contém uma grande variedade de moléculas farmacológicas potentes, como vasodilatadores, antiagregante plaquetários, anticoagulantes, e imunossupressores (Champagne, 1994; Ribeiro, 1995; Stark e James, 1996).

As funções anti-hemostáticas dessas moléculas são mediadas por famosas famílias de proteínas, tais como serpinas, lipocalinas, apirases, serino proteases, moléculas que contêm o domínio Kazal e kunitz, entre várias outras moléculas (Assumpção *et al.*, 2013).

1.6.1. Vasodilatadores

Algumas moléculas salivares dos triatomíneos hematófagos são vasodilatadores que facilitam a alimentação de duas maneiras: injetados na pele como sondas, eles dilatam os vasos sanguíneos e aceleram a descoberta dos mesmos, e subsequentemente aumentam o fluxo sanguíneo para o local de alimentação, diminuindo o tempo e o número de picadas necessários para adquirir o repasto (Champagne, 2004; 2005; Araujo et al., 2009). Um exemplo dessas moléculas são as nitroforinas I, II, III e IV (NPI - NPV) de *Rhodnius prolixus*. Elas são hemeoproteínas que se ligam e estabilizam reversivelmente ao óxido nítrico (NO) (Andersen et al., 1997). A modificação do pH de aproximadamente 6,0 na saliva para 7,4 no sangue resulta em alterações conformacionais da proteína e liberação de NO, que se difunde nas células musculares lisas vasculares promovendo relaxamento e consequente vasodilatação (Montfort et al., 2000; Bussacos et al., 2011; De Araújo et al., 2012). Outro exemplo é a proteína salivar ligadora de aminas (ABP, amine binding protein) do *R. prolixus* que se liga à serotonina, epinefrina e norepinefrina, inibindo a vasoconstrição (Andersen et al., 2003).

1.6.2. Inibidores da agregação plaquetária

A ativação e agregação plaquetária são direcionadas por respostas mediadas por receptores a sinais associados ao trauma, incluindo ADP, colágeno e trombina. Como as respostas plaquetárias podem ser extremamente rápidas, elas são particularmente relevantes para o processo de repasto sanguíneo, e os artrópodes têm uma variedade de estratégias para inibir essa função (Champagne, 2005). Existem algumas proteínas da família do antígeno 5 salivar que inibem a agregação plaquetária induzida por colágeno. Uma delas é a TIAV (antígeno 5 *T. infestans*) em *Triatoma infestans* (Morita et al., 2006; Assumpção et al., 2008; Assumpção et al., 2013), e outra é a proteína DMAV (antígeno 5 *D. maxima*) de *Dipetalogaster maxima* (Assumpção et al., 2010). Outro exemplo é a palidipina, encontrada no *Triatoma pallidipennis*, que não tem efeito na adesão plaquetária ao colágeno, mas que interfere na agregação das plaquetas inibindo a liberação

de ATP (adenosina trifosfato) pelas mesmas (Noeske-Jungblut et al., 1994). A triabina, também isolada do triatomíneo *Triatoma pallidipennis*, inibe a agregação plaquetária induzida pela trombina (Fuentes-Prior et al., 1997). O triatomíneo *Rhodnius prolixus* produz a (RPAI) (*Rhodnius* platelet aggregation inhibitor), uma molécula que inibe a agregação plaquetária ligando-se ao ADP, impedindo assim que ele exerça sua função (Francischetti et al., 2000).

A estratégia inibidora mais difundida é a presença de uma apirase salivar (ATP / ADP difosfohidrolase) que depleta o ADP (e o ATP) da lesão, interferindo diretamente no processo de agregação plaquetária (Champagne, 2005). Já foram descritas apirases na saliva de diversos triatomíneos, entre eles o *T. infestans* (Faudry et al., 2004; Charneau et al., 2007), *D. maxima* (Assumpção et al., 2010), *P. megistus* (Ribeiro et al., 2015), *R. prolixus* (Sarkis et al., 1986) e *R. neglectus* (Santiago et al., 2016).

1.6.3. Anticoagulantes

Durante a alimentação, os insetos hematófagos devem evitar a coagulação do sangue para que ele não obstrua as partes do seu aparelho sugador e seu o intestino. Sabe-se que a saliva do *R. prolixus* contém uma proteína chamada prolixina-S que atua como anticoagulante por impedir a conversão do fator X em Xa (Hellmann e Hawkins, 1965).

T. pallidipennis também possui um anticoagulante chamado triabina, que inibe especificamente a trombina na cascata de coagulação (Noeske-Jungblut et al., 1995; Montfort et al., 2000).

1.6.4. Imunosupressores

Insetos hematófagos que se alimentam lentamente enfrentam a dificuldade de causar dano tecidual durante o repasto, além de suas proteínas salivares que também resultam em inflamação. Para superar esse problema, esses insetos possuem na saliva proteínas capazes de sequestrar a histamina, a molécula de sinalização secretada pelos mastócitos do

hospedeiro. Em *Rhodnius*, as nitroforinas ligam-se à histamina na porção heme, a mesma localização que o NO se liga, assim evitando que a histamina ative o processo inflamatório (Ribeiro e Walker, 1994; Weichsel et al., 1998; Montfort et al., 2000). A histamina liberada pelo hospedeiro produz dor e coceira, ativando suas respostas defensivas que podem arruinar a alimentação do inseto ou até mesmo sua vida. A neutralização da ação da histamina no local da alimentação é, portanto, uma vantagem para a sobrevivência do inseto (Ribeiro e Walker, 1994).

A proteína salivar ligadora de aminas (ABP, amine binding protein) também se liga à histamina para evitar sua dispersão nos vasos sanguíneos, atuando como um inibidor da resposta inflamatória (Bussacos et al., 2011).

1.7. Estudos das proteínas salivares de animais hematófagos

Estudos sobre as características e funções das proteínas salivares de animais hematófagos tem crescido muito com o avanço da ciência. Alguns resultados já tem demonstrado potencial farmacológico para o desenvolvimento de drogas e vacinas (Ciprandi et al., 2003; Andrade et al., 2005; Titus et al., 2006).

Podemos citar o exemplo da hirudina, um anticoagulante presente na glândula salivar de *Hirudo medicinalis*, uma sanguessuga. Essa proteína de 7 kDa foi primeiro descrita por Haycraft (1884) e depois isolada em 1967 por Markwardt (Markwardt et al., 1967; Ciprandi et al., 2003). A hirudina é um peptídeo inibidor de trombina, composto por uma única cadeia de 65 resíduos de aminoácidos com três pontes dissulfeto intracadeia. Sua porção N-terminal é globular e devido à presença de três pontes dissulfeto, é fortemente compacta. Já a porção C-terminal é linear, com vários resíduos carregados negativamente que reagem com o exosítio I da trombina. Os resíduos da porção N-terminal, por sua vez, interagem com o sítio catalítico. É devido a esse padrão de interação que a hirudina é altamente específica para trombina e não se liga a outras serino proteases (Salzet, 2001). A disponibilidade dessa proteína recombinante tem permitido grandes avanços nos estudos farmacológicos e de uso clínico. O estudo das características de

sua ligação à trombina levou ao desenvolvimento de pequenos compostos sintéticos contendo apenas os domínios necessários para ligação à trombina (Maraganore et al., 1990). Originalmente conhecido como hirulog, ele foi aprovado para uso como anticoagulante em angioplastia coronariana, tendo a capacidade de inibir a atividade da trombina ligada ao coágulo (Gladwell, 2002).

Outro exemplo é o maxadilan, uma proteína vasodilatadora potente descoberta em 1991 (Lerner et al., 1991). O maxadilan é um peptídeo de 6,5 kDa, contendo 61 resíduos de aminoácidos e foi codificado por um gene clonado das glândulas salivares de *Lutzomyia longipalpis*. Além de vasodilatador, ele também possui propriedades imunomoduladoras (Lerner et al., 1991; Lerner e Shoemaker, 1992; Andrade et al., 2005).

A saliva dos artrópodes hematófagos também está relacionada com a produção de anticorpos específicos por humanos e outros vertebrados contra seus componentes (Feingold e Benjamini, 1961; Brummer-Korvenkontio et al., 1994; Wikel, 1996). Estas respostas imunes podem ser usadas como marcadores epidemiológicos de exposição à vetores (Schwartz et al., 1990; Barral et al., 2000), e dão suporte à possibilidade de futuro desenvolvimento de vacinas (Andrade et al., 2005; Titus et al., 2006).

Uma das ferramentas importantes no estudo dos componentes da saliva, os genomas, os transcriptomas (descrição do conjunto de mRNAs), proteomas (descrição das proteínas) e sialomas (do grego sialo - saliva; junção do transcriptoma com proteoma) podem permitir a seleção de transcritos de interesse para a produção de proteínas recombinantes, a fim de que elas possam ter sua função determinada e sua presença na saliva identificada (Ribeiro e Francischetti, 2003; Arcà e Ribeiro, 2018).

Santiago e colaboradores (2016) publicaram o sialoma do triatomíneo *Rhodnius neglectus*, descrevendo as possíveis proteínas secretadas. Entre elas se encontram as lipocalinas, componentes da família do antígeno 5, OBP (odorant-binding protein), proteases e inibidores de serino protease. A partir dos dados do transcriptoma, foram selecionados dois transcritos de

interesse para o presente estudo: o primeiro codifica uma serino protease tipo tripsina e o segundo um inibidor de serino protease tipo Kazal.

1.7.1. Serino proteases

As proteases são enzimas capazes de clivar as ligações peptídicas. Elas podem executar uma enorme variedade de funções fisiológicas complexas, exercendo papel-chave em processos como o catabolismo de proteínas, ativação de zimogênios, liberação de peptídeos farmacologicamente ativos de proteínas precursoras, entre outras funções (Tremacoldi, 2009).

A serino protease é caracterizada pela presença de um resíduo de serina reativo no centro ativo, e o mecanismo catalítico destas proteases envolve a ligação covalente de substratos a este aminoácido (Bond e Butler, 1987). Algumas serino proteases podem funcionar como reguladores da coagulação. A trombina pode participar desta regulação ligando-se à uma proteína de membrana presente nas células endoteliais do hospedeiro chamada trombosmodulina. Este complexo é capaz de ativar a proteína C (serino protease), que atua nos fatores de inativação V e VIII, prejudicando a progressão do trombo (Kalafatis et al., 1994; Kalafatis et al., 1997; Santiago et al., 2016b).

Em artrópodes, as serino proteases desempenham papéis importantes também na fertilização, no desenvolvimento embrionário e nos processos de muda e metamorfose de insetos (Jiang e Kanost, 2000; He et al., 2009; Santiago et al., 2016b). A tripsina é uma serino protease do tipo tripsina de *T. infestans* (Amino et al., 2001). A tripsina é encontrada como um zimogênio no lúmen da glândula salivar de *T. infestans* e é conhecida por ter uma função amidolítica que melhora a captação sanguínea. Essa molécula é ativada pela tripsina quando a saliva é ejetada durante o repasto sanguíneo. Proteínas semelhantes foram observadas na saliva de *T. brasiliensis* (Santos et al., 2007) e de *P. megistus* (Bussacos et al., 2011; De Araújo et al., 2012).

1.7.2. Inibidores de serino proteases do tipo Kazal

Os inibidores de serino proteases compreendem um grupo heterogêneo de moléculas, incluindo compostos de baixo peso molecular, assim como inibidores proteicos de alto peso molecular. Para a família dos inibidores do tipo Kazal com a presença de múltiplos domínios, diversas características estruturais comuns são observadas. Entre elas estão um padrão característico de distribuição de cisteínas, estruturas tridimensionais altamente homólogas, alta potência inibitória e uma pronunciada especificidade contra distintas enzimas-alvo (Kazal et al., 1948; Friedrich et al., 1993; Van De Locht et al., 1995).

Esses inibidores são proteínas que desempenham funções importantes em invertebrados, como por exemplo efeitos vasodilatadores, de inibição da trombina (Schlott et al., 2002), anticoagulantes (Friedrich et al., 1993; Campos et al., 2002; Araujo et al., 2007), antibacterianos (Donpuksa et al., 2009; Kim et al., 2013; Soares et al., 2015) e antitripanossomas (Lovato et al., 2011; Soares et al., 2015).

Os inibidores de protease são proteínas únicas ou com múltiplos domínios que compartilham um motivo conservado (Schlott et al., 2002). Em insetos hematófagos, eles têm diferentes funções. Um inibidor específico de trombina com uma massa molecular de 11 kDa foi purificado do inseto *Rhodnius prolixus* e nomeado de rhodniin (Friedrich et al., 1993). A dipetalogastina do triatomíneo *Dipetalogaster maxima* (Mende et al., 2004) é uma proteína com seis domínios do tipo Kazal, cujas funções principais são as potentes inibições de trombina e de tripsina. O inibidor de trombina de *T. brasiliensis* conhecido com brasiliensina é codificado também por um gene contendo oito domínios de Kazal (Araujo et al., 2007; Bussacos et al., 2011).

2. Justificativa

Na saliva de triatomíneos é possível encontrar moléculas farmacologicamente ativas que auxiliam na obtenção de alimentação sanguínea e aumentam a possibilidade de infecção do hospedeiro pelo *T. cruzi*. Entre essas moléculas estão as proteases e os inibidores de proteases. Este trabalho permitirá a caracterização de uma serino protease tipo tripsina de *R. neglectus*, contribuindo para o entendimento dos mecanismos regulatórios dessa via de processamento de proteínas e de seus papéis na hematofagia do inseto. Permitirá ainda a caracterização de um membro da família de inibidores de serino proteases tipo Kazal, que pode vir a ser um alvo para dificultar o repasto satisfatório e a transmissão vetorial do *T. cruzi*.

O conhecimento sobre essas moléculas salivares de *R. neglectus* poderá ainda confirmar seus potenciais como moléculas de interesse clínico e biotecnológico, por exemplo servindo como marcadores de exposição à picada do inseto ou no desenvolvimento de drogas para o tratamento de doenças.

3. Objetivos

3.1. Objetivo geral

Analisar a expressão de uma serino protease do tipo tripsina (RnTrypsin) e de um inibidor de serino protease tipo Kazal (RnKazalPI) identificados na saliva de *R. neglectus*, vetor da Doença de Chagas, em sistema heterólogo bacteriano.

3.2. Objetivos específicos

3.2.1. Análise *in silico* da serino protease do tipo tripsina (UniProt: A0A0P4VQ38; RnTrypsin) e do inibidor de serino protease tipo Kazal (UniProt: A0A0P4W3X1; RnKazalPI) salivares de *R. neglectus*;

3.2.2. Expressão das proteínas RnTrypsin e RnKazalPI em sistema heterólogo;

3.2.3. Purificação da RnTrypsin recombinante;

3.2.4. Produção de anticorpos policlonais anti-RnTrypsin de *R. neglectus*;

4. Metodologia

As Figuras 6 e 7 mostram as representações esquemáticas da metodologia empregada para realizar e analisar a expressão da RnTrypsin e da RnKazalPI recombinantes.

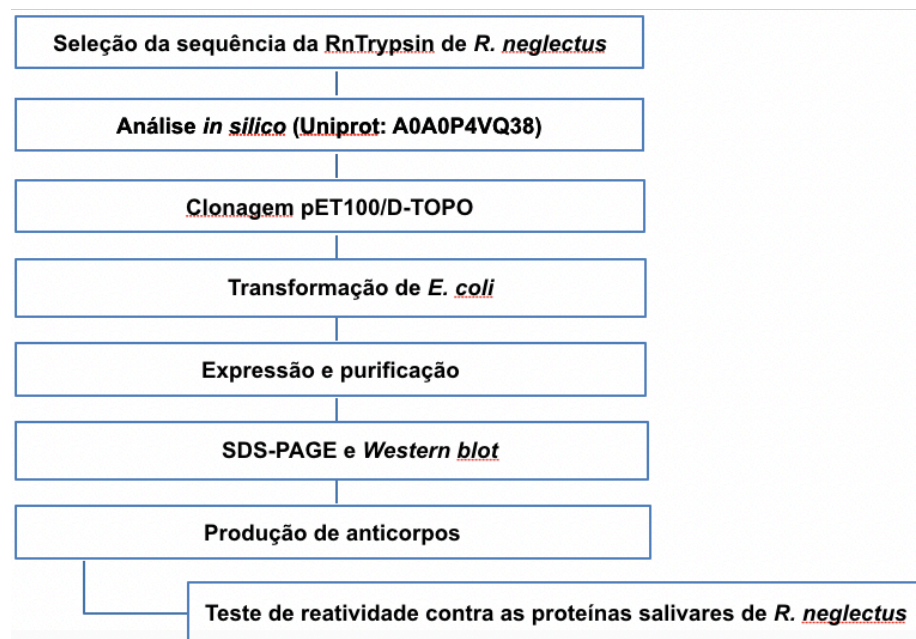


Figura 6. Representação esquemática da metodologia empregada para realizar e analisar a expressão da e RnTrypsin recombinante.

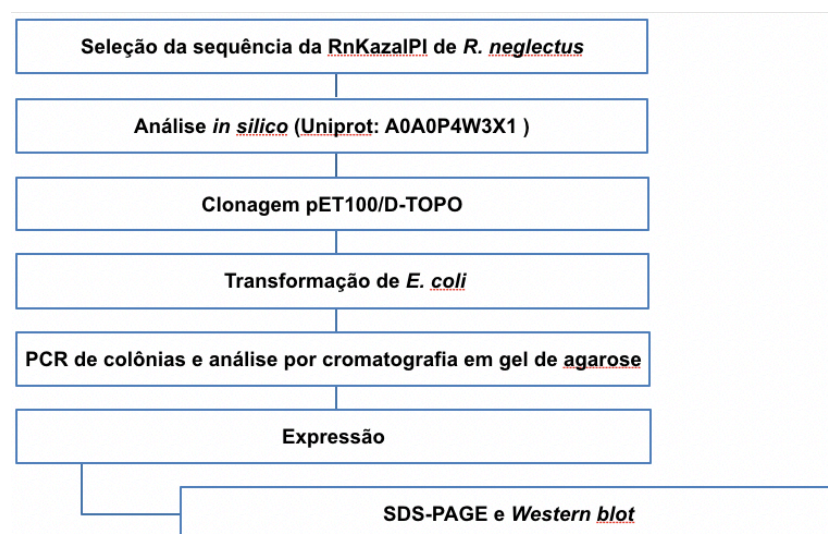


Figura 7. Representação esquemática da metodologia empregada para realizar e analisar a expressão da RnKazalPI recombinante.

4.1. Seleção das sequências de proteínas salivares de *R. neglectus*

A partir dos dados do sialoma de *R. neglectus* produzidos por SANTIAGO *et al.* (2016), foram selecionados dois transcritos: o primeiro, com número de acesso UniProt A0A0P4VQ38, codifica para uma serino protease do tipo tripsina (RnTrypsin) e o segundo, com número de acesso UniProt A0A0P4W3X1, codifica para um inibidor de protease do tipo Kazal (RnKazalPI).

4.2. Análise *in silico*

Uma análise *in silico* das sequências de aminoácidos das proteínas foi realizada visando uma caracterização inicial. Os softwares descritos a seguir foram utilizados: **ProtParam** (<https://web.expasy.org/protparam/>) que calcula vários parâmetros físicos e químicos para uma determinada proteína, como o peso molecular, ponto isoelétrico teórico, composição de aminoácidos, índice de instabilidade, entre outros; **Protein Blast** (Blastp, NCBI, Bethesda, MD, USA) que faz o alinhamento simples de uma sequência de aminoácidos com as sequências de outros organismos com o objetivo de verificar dados de similaridade e identidade entre elas; **MultAlin 5.4.1** (Corpet, 1988) que faz o alinhamento de múltiplas seqüências para sequências de proteínas e ácidos nucléicos; **TargetP 1.1** (Nielsen *et al.*, 1997; Emanuelsson *et al.*, 2000) que prevê a localização de proteínas no interior de células eucarióticas, com base na presença de alguma das sequências N- terminal: peptídeo sinal para via secretora (SP), cloroplasto (cTP) ou mitocôndria (mTP); **SignalP 3.0** (Bendtsen *et al.*, 2004) o qual prediz a presença e a localização dos sítios de clivagem do peptídeo sinal nas sequências de aminoácidos de eucariotos; **NetPhos 3.1** (Blom *et al.*, 1999) que prediz sítios de fosforilação de serina, treonina ou tirosina em proteínas eucarióticas; **NetNGlyc 1.0** e **NetOGlyc 4.0** (Steentoft *et al.*, 2013) os quais predizem sítios de N-glicosilação e O-glicosilação, respectivamente, nas sequências das proteínas.

4.3. Clonagem molecular

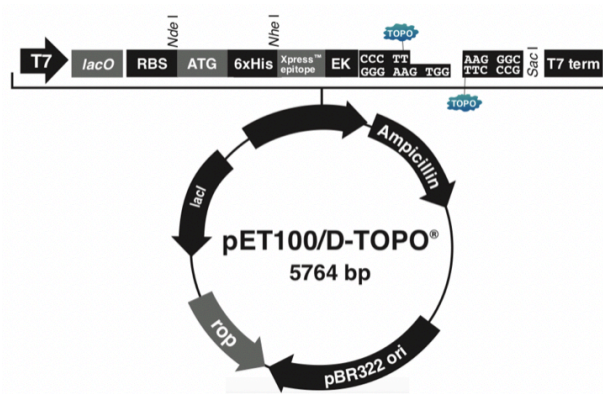
As sequências de DNA da serino protease tipo tripsina (UniProt: A0A0P4VQ38; RnTrypsin) e do inibidor de serino protease tipo Kazal

(UniProt: A0A0P4W3X1; RnKazalPI) foram códon-otimizadas para expressão em sistema bacteriano (**Anexos 1 e 2**), sintetizadas e clonadas no vetor pET100/D-TOPO pela empresa Thermo Fisher (**Anexos 3 e 4**).

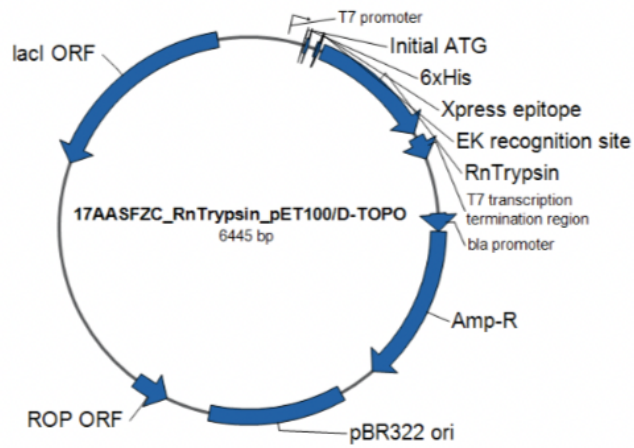
O pET100/D-TOPO possui um promotor T7/lac , uma sequência de DNA de bacteriófago que é reconhecida pela T7 RNA polimerase. Esse promotor é usado para regular a expressão de proteínas (Tabor e Richardson, 1985). Possui ainda uma sequência de operador lac (lacO) imediatamente *downstream* ao promotor T7 e uma origem de replicação do vetor de clonagem pBR322, que reduzem os níveis de expressão basal do gene antes da indução; um sítio RBS (Ribosome Binding Site - sítio de ligação do ribossomo) que é responsável pelo recrutamento ribossomal durante o início da tradução de proteínas; duas sequências que codificam para caudas de fusão na porção N-terminal (uma com seis resíduos de histidina - 6xHis, e a outra com o epítipo XpressTM), para possibilitarem a detecção e purificação das proteínas recombinantes; um sítio para enteroquinase (EK), que possibilita a remoção das caudas após a purificação da proteína recombinante; um terminador T7 que permite a terminação eficiente da transcrição; e um gene de resistência ao antibiótico ampicilina, para facilitar a seleção de transformantes positivos (Figura 8 A).

A sequência nucleotídica da serino protease tipo tripsina sem a região que codifica para o peptídeo sinal possui 681 pares de bases (pb); quando clonada originou um plasmídeo com 6445 pb (Figura 8 B). A sequência nucleotídica do inibidor de serino protease tipo Kazal, sem a sequência do peptídeo sinal, contém 174 pb e originou um plasmídeo com 5938 pb após sua clonagem (Figura 8 C). Ambas foram clonadas direcionalmente no vetor, fusionadas a uma cauda de 6 resíduos de histidina (6xHis) na porção N-terminal e ao epítipo Xpress, o que resultou em uma diferença entre os pesos moleculares observados e esperados das proteínas recombinantes.

A



B



C

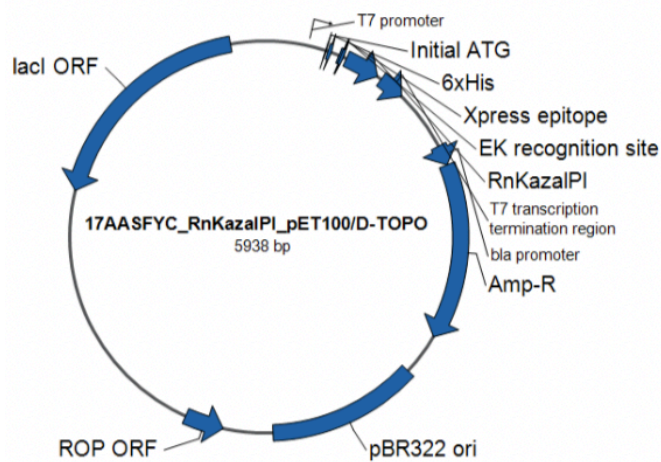


Figura 8. Representações esquemáticas do vetor de expressão pET100/D-TOPO. A - pET100/D-TOPO. B - pET100/D-TOPO contendo a sequência da serino protease tipo tripsina (RnTrypsin). C - pET100/D-TOPO contendo a sequência do inibidor de serino protease tipo Kazal (RnKazalPI). FONTE: ThermoFisher Scientific.

Tabela 1. Características dos insertos utilizados.

Inserto	Sequência nucleotídica (pb*)	Sequência aminoácídica (aa**) com PS**	Sequência aminoácídica (aa**) sem PS	MW esperado (kDa) sem PS	MW esperado (kDa) sem PS com 6xHis
RnTrypsin	681	244	227	27	30
RnKazalPI	174	83	58	9	10

* pares de base; ** resíduos de aminoácidos; PS - peptídeo sinal.

4.4. Transformação de *Escherichia coli* com pET100/D-TOPO_RnTrypsin e pET100/D-TOPO_RnKazalPI

Uma alíquota de 50 μL de *E. coli* competentes das linhagens BL21(DE3), BL21(DE3)pLysS ou Rosetta2(DE3)pLysS foi incubada com 20 ng/ μL do vetor pET100/D-TOPO_RnTrypsin ou pET100/D-TOPO_RnKazalPI por 30 min no gelo. Em seguida, foram submetidas a um choque térmico à 42 °C por 1,5 min e incubadas no gelo por 2 min. Após esse período, 250 μL de meio SOC foram adicionados aos tubos e as bactérias cresceram por 1 h a 37 °C sob agitação de 250 rpm. Ao término dessa incubação, a cultura foi plaqueada em meio Luria Bertani ágar (LB ágar) contendo 100 $\mu\text{g}/\text{mL}$ de ampicilina. As placas foram incubadas à 37 °C por 16 h para o desenvolvimento das colônias.

Às placas onde foram adicionadas as culturas das linhagens BL21(DE3)pLysS ou Rosetta2(DE3)pLysS, adicionou-se também 34 $\mu\text{g}/\text{mL}$ de cloranfenicol além de 100 $\mu\text{g}/\text{mL}$ de ampicilina. Essa adição se deve ao fato de que essas duas linhagens contêm o plasmídeo pLysS, que codifica para a cloranfenicol acetiltransferase, uma enzima de 25,7 kDa que confere resistência ao antibiótico cloranfenicol.

Tabela 2. Características das linhagens de *E. coli* utilizadas nesse estudo.

	BL21(DE3)	BL21(DE3)pLysS	Rosetta(DE3)pLysS
(DE3): possui uma cópia do gene da RNA polimerase T7 sob o controle de promotor induzível por IPTG	X	X	X
pLysS: reduz a expressão basal do gene alvo antes da adição do indutor IPTG		X	X
Otimização da expressão de proteínas que contêm códons raramente usados em <i>E. coli</i>			X

4.5. Identificação de colônias de *E. coli* transformadas com pET100/D-TOPO_RnKazalPI e análise por eletroforese em gel de agarose

Após a transformação de Rosetta(DE3)pLysS com o plasmídeo pET100/D-TOPO_RnKazalPI, seis colônias foram escolhidas aleatoriamente e plaqueadas em meio LB ágar contendo 34 µg/mL de cloranfenicol e 100 µg/mL de ampicilina. Cada ponteira utilizada para recolher as colônias foi agitada vigorosamente em 10 µL de água ultrapura em um microtubo. As amostras foram fervidas por 5 min para lisar as bactérias e desnaturar as DNAses.

Quinze microlitros de uma mistura contendo PCR Master Mix (ThermoFisher Scientific®), primer T7 promotor (concentração final de 0,2 µM), primer T7 terminador (concentração final de 0,2 µM), e água ultrapura esterilizada foram adicionados à cada microtubo contendo o lisado bacteriano de cada colônia. Para o (C+) controle positivo, foi utilizado 9,5 µL de água ultrapura, 15 µL da mistura e 0,5 µL do plasmídeo pET100/D-TOPO_Kazal (concentração final de 10ng). Para o (C-) controle negativo, foi utilizado 10 uL

de água ultrapura esterilizada e 15 µL da mistura. Foram colocados os microtubos em um termociclador ajustado para 35 ciclos, com 1 min à 94°C (desnaturação), 1 min à 55°C (anelamento) e 8 min à 72°C (elongação).

Ao término da PCR, os produtos de DNA gerados foram submetidos à eletroforese em gel de agarose 1,2% contendo 0,5 g/mL de brometo de etídio. A eletroforese foi realizada à temperatura ambiente com voltagem constante de 100 V e o gel visualizado em um fotodocumentador.

4.6. Expressão de RnTrypsin e RnKazalPI recombinantes

Algumas colônias resultantes da transformação foram selecionadas aleatoriamente e inoculadas separadamente em 5 mL de meio Luria Bertani (LB) contendo 100 µg/mL de ampicilina ou 34 µg/mL de cloranfenicol e 100 µg/mL de ampicilina dependendo da linhagem de *E. coli*. As culturas foram incubadas à 37 °C por 16 h sob agitação constante de 250 rpm. Após esse período, 4 mL de cada cultura foram transferidos para 100 mL de meio LB contendo os antibióticos específicos. As culturas foram incubadas à 37 °C sob agitação constante de 250 rpm até atingirem densidade ótica (OD₆₀₀) de 0,4 a 0,6, medida em leitora de microplacas Spectra-Max M5 (Molecular Devices). A expressão das proteínas recombinantes foi induzida pela adição de IPTG. A concentração final de IPTG variou entre 0,2 mM e 1 mM. Após a adição do agente indutor, as culturas foram incubadas variando o tempo (3 a 40 h) e a temperatura (37 a 15 °C), sob agitação constante de 250 rpm.

Decorrido o período de indução, as culturas foram centrifugadas à 2.000 x g por 20 min à 4°C em centrífuga CF16RXII (Hitachi). Os sobrenadantes foram descartados e os sedimentos foram lisados usando solução de lise BugBuster® (Novagen) acrescido de um coquetel de inibidores de protease 1X (cOmplete Protease Inhibitor Cocktail, Roche) e DNase, seguindo as instruções do fabricante. Resumidamente, o lisado foi mantido sob agitação constante por 20 min a temperatura ambiente e em seguida centrifugado à 16.000 x g por 20 min à 4 °C. As frações solúvel e insolúvel foram separadas. As frações insolúveis foram reconstituídas no mesmo volume de água ultrapura. Dez microlitros de cada fração foram

separados por eletroforese em gel de poliacrilamida (SDS-PAGE 12-15%; w/v), conforme descrito abaixo.

Para produção em maior escala, a melhor condição de expressão identificada nos testes de expressão foi utilizada.

4.7. Purificação da RnTrypsin recombinante

Após a confirmação da expressão da proteína RnTrypsin na fração insolúvel de BL21(DE3) (indução com 1 mM de IPTG à 37 °C por 3h sob agitação) pela técnica de *Western blot*, essa fração foi submetida a cromatografia de afinidade em resina His-Bind carregada com níquel (GE Healthcare), projetada para ligar especificamente proteínas contendo uma cauda de histidinas.

A fração insolúvel foi submetida a protocolo para obtenção dos corpos de inclusão utilizando o reagente BugBuster®. Resumidamente, após a obtenção das frações solúvel e insolúvel por centrifugação, à fração insolúvel foi adicionada novamente a solução de lise BugBuster® e o conteúdo homogeneizado através de cuidadosa pipetagem seguida por agitação em vortex por 30 seg. Uma solução de lisozima foi adicionada para uma concentração final de 200 µg/mL, homogeneizou-se novamente em vortex por 1 min, com posterior incubação a temperatura ambiente por 5 min. Em seguida, a solução de lise BugBuster® diluída em água ultrapura (1:10) foi adicionada e homogeneizada em vortex por 1 min. As amostras foram centrifugadas à 16.000 x g por 15 min a 4°C e descartou-se o sobrenadante cuidadosamente. As etapas desde a adição do BugBuster® diluído em água ultrapura (1:10) até o descarte do sobrenadante após centrifugação foram repetidas por mais três vezes. Por fim, o sedimento foi ressuspenso em tampão de ligação 1X (Tris-HCl 20 mM, NaCl 500 mM, pH 7,9), contendo imidazol 5 mM e uréia 6 M.

A amostra foi aplicada em coluna contendo a resina His-Bind carregada com níquel (GE Healthcare, EUA) pré-equilibrada com o tampão de ligação 1X contendo imidazol 5 mM e uréia 6 M. Antes da eluição da RnTrypsin, a coluna foi lavada com 3 mL deste tampão. As proteínas ligadas à resina foram eluídas com o tampão de ligação 1X contendo concentrações

crescentes de imidazol (20mM, 60 mM, 80 mM, 100 mM, 200 mM e 300 mM) e uréia 6 M. As frações obtidas da purificação foram analisadas por SDS-PAGE 12%, e *Western blot*. As frações contendo a proteína eluída foram estocadas a -20 °C até utilização para imunização dos camundongos.

4.8. Eletroforese em SDS-PAGE

A fim de analisar se as proteínas recombinantes foram expressas na fração solúvel ou insolúvel, e posteriormente a pureza das proteínas eluídas da cromatografia de afinidade, as amostras foram submetidas à eletroforese em gel de poliacrilamida 12% (w/v) (no caso da RnTrypsin) e 15% (w/v) (no caso do RnKazalPI) sob condições desnaturantes e redutoras, conforme método descrito por Laemmli (1970). Para a análise da pureza, também foram adicionadas amostras do extrato total da fração insolúvel, a fração não ligada que é a fração obtida durante as lavagens com o tampão de ligação contendo 5 mM de imidazol e 6 M de uréia, além das frações de eluição contendo 20 mM a 300 mM de imidazol .

As amostras foram diluídas em tampão de amostra 5x (Tris HCl 50 mM pH 6,8, glicerol 10%, SDS 2%, azul de bromofenol 0,1%, β -Mercaptoetanol 15 mM e submetidas à uma temperatura de 100 °C por 5 min antes de serem aplicadas no gel. Foram utilizados os marcadores *Bench Mark™ Protein Ladder* (Invitrogen) e *Bench Mark™ Prestained Protein Ladder* (Invitrogen) como padrões de peso molecular. A eletroforese foi realizada à temperatura ambiente em sistema MiniProtean®3 Cell (BioRad), com voltagem constante de 90 V durante passagem pelo gel concentrador e 120 V pelo gel separador, com tampão de corrida Tris- Glicina (Tris HCl 25 mM, pH 8,8, glicina 250 mM e SDS 0,1%).

Os géis foram corados ou transferidos para membrana de nitrocelulose para realização de *Western blot*. Para visualização das proteínas após coloração, os géis foram submersos em solução de coloração azul de Comassie (metanol 45%, ácido acético 10%, *Comassie Brilliant Blue R- 250*, Sigma) por 2 h sob agitação lenta. Posteriormente, essa solução foi substituída por solução descorante (ácido acético 10% e metanol 10%), onde permaneceram até a adequada visualização das bandas proteicas. Para

descontinuar a ação da solução descorante, o gel foi lavado em água ultrapura.

Também foi realizada coloração com nitrato de prata, em quatro etapas. A primeira etapa foi a de fixação, mergulhando o gel em 25 mL de metanol, 6 mL de ácido acético, 19 mL de água ultrapura e 25 µL de formaldeído (37%) durante 1 h. Em seguida, o gel foi lavado em uma solução contendo 25 mL de água ultrapura e 25 mL de etanol durante 20 min. A segunda etapa foi a de sensibilização, deixando o gel mergulhado em solução de 0,01 g de tiosulfato de sódio diluído em 50 mL de água ultrapura durante 1 min. Em seguida, ocorreu uma nova lavagem com água ultrapura durante 1 min. A terceira etapa é a de impregnação, onde o gel ficou mergulhado em uma solução contendo 0,1 g de nitrato de prata, 50 mL de água ultrapura e 37,5 µL de formaldeído (37%) durante 20 min. O gel foi lavado novamente em água ultrapura durante 1 min. Por fim, a quarta etapa foi a de revelação, onde o gel ficou mergulhado em uma solução contendo 50 mL de água ultrapura, 3 g de carbonato de sódio, 0,01 g de tiosulfato e 25 µL de formaldeído (37%), até as bandas aparecerem. Essa última etapa é fotossensível, portanto foi realizada em local escuro.

4.9. Western blot

Para confirmar a expressão das proteínas recombinantes, foi realizada análise por *Western blot*. Essa técnica visa a transferência de proteínas separadas por SDS-PAGE para uma membrana de nitrocelulose (Amersham™ Protran® Supported 0.45 NC, GE Healthcare). Antes do início da transferência, a membrana ficou submersa em tampão (Tris-HCl 25 mM pH 8.3, glicina 192 mM, SDS 0,1% (m/v) e metanol 20% (v/v)) por 10 min à 4°C para equilibrá-la. Para a transferência, utilizou-se o sistema MiniProtean®3 Cell (BioRad) com uma corrente constante de 350 mA por 1 h e 20 min. Quando finalizada, a membrana foi mergulhada em solução Ponceau (Ponceau S 0,2% p/v; ácido acético 1%) por 5 min e depois lavada com água ultrapura para a visualização das bandas.

Para evitar ligações inespecíficas do anticorpo às proteínas transferidas para a membrana, esta foi incubada por 1 h em solução *Tris-*

buffered saline (TBS, Tris 50 mM, pH 7.5 e NaCl 150 mM) contendo 5% (m/v) de leite desnatado. Após esse bloqueio, a membrana foi lavada duas vezes de 10 min cada com aproximadamente 25 mL em cada lavagem de TBS contendo 0,1% de Tween 20, à temperatura ambiente e sob agitação moderada. A membrana foi posteriormente incubada com o anticorpo primário anti-His (Sigma Aldrich), diluído 1:2000 em TBS contendo 0,1% (v/v) de Tween 20 e 1% (m/v) de leite desnatado por 16 h a 4°C. Após essa etapa, a membrana foi lavada quatro vezes de 5 min cada com aproximadamente 25 mL em cada lavagem de TBS contendo 0,1% de Tween 20, à temperatura ambiente e sob agitação moderada. Em seguida, a membrana foi incubada com o anticorpo secundário (*goat anti-mouse IgG horseradish peroxidase conjugated*) diluído 1:30.000 em TBS Tween 20 0,1% e leite desnatado 1% (m/v) por 1,5 h. Finalizada essa etapa, a membrana foi lavada sete vezes de 5 min cada com aproximadamente 25 mL em cada lavagem de TBS Tween 20 0,1%. Por fim, as bandas de proteínas imunorreativas foram visualizadas no equipamento ImageQuant LAS 4000 (GE Healthcare) utilizando o substrato quimioluminescente de peroxidase (*ECLTM Prime Western blot detection reagent*, GE Healthcare).

4.10. Produção de anticorpos policlonais contra a RnTrypsin

A produção de anticorpos policlonais contra a RnTrypsin foi realizada seguindo as diretrizes da Comissão de Ética no Uso Animal da UnB, de acordo com o protocolo UnBDoC no. 27323/2016 (**Anexo 5**). Cinco camundongos fêmeas Balb/c com idade de 10 semanas, mantidos no Biotério do Instituto de Ciências Biológicas, UnB, foram utilizados.

Dez dias antes da primeira imunização, amostras de sangue foram coletadas por punção via retro-orbital com o animal previamente anestesiado com cetamina e xilasina (1:1) por injeção intraperitoneal. O sangue foi centrifugado a 600 x g por 10 min, o soro pré-imune separado, diluído em uma proporção 1:1 em glicerol e armazenado a -20 °C.

A concentração de proteínas nas frações eluídas da purificação contendo a RnTrypsin foi dosada por fluorometria utilizando o equipamento Qubit® 2.0 (Invitrogen). As aplicações foram feitas por via intraperitoneal e

ocorreram em intervalos de 15 em 15 dias. Na primeira imunização foram administrados 10 µg de RnTrypsin emulsionados em adjuvante completo de Freund para um volume final de 100 µL (1:1). Nas duas imunizações seguintes, foram administrados 10 µg de RnTrypsin emulsionados em adjuvante incompleto de Freund para um volume final de 100 µL (1:1). Na quarta e última imunização, 10 µg de RnTrypsin foi diluída em 100 µL PBS estéril. Sete dias após a última imunização, amostras de sangue foram coletadas por punção cardíaca nos camundongos previamente anestesiados. Após a coleta, os animais foram sacrificados por deslocamento cervical. O sangue foi centrifugado a 600 x g por 10 min, o soro imune separado, diluído em uma proporção 1:1 em glicerol e armazenado a -20 °C.

4.11. Colônia de triatomíneos

Os triatomíneos *Rhodnius neglectus* (coletados em Itambaracá, no estado do Paraná, Brasil, em 1982) foram mantidos no biotério do Instituto de Ciências Biológicas da Universidade de Brasília, a 26 +/-1 °C, umidade relativa de 80 % e ciclo claro/escuro de 12/12 h. Os insetos foram alimentados quinzenalmente em sangue de galinhas (*Gallus gallus domesticus*). Os ovos, as ninfas de 1° ao 5° estágios e os adultos foram mantidos separadamente em potes de polietileno para desenvolvimento e acasalamento. Os insetos adultos foram utilizados para a obtenção da saliva.

4.12. Obtenção da saliva

Os insetos adultos foram escolhidos aleatoriamente e colocados em um frasco de polietileno que estava sobre uma camada de gelo, para que com a queda de temperatura eles pudessem entrar em estado letárgico. Após cerca de 5 min, cada triatomíneo foi colocado sobre uma lâmina de vidro, e com o auxílio de pinças cirúrgicas, o inseto foi imobilizado e a cabeça deslocada cuidadosamente expondo assim o par de glândulas. Essas foram recolhidas utilizando a ponta de uma seringa de 0,3 mm, depositadas em um microtubo de 1,5 mL contendo 50 µL de água ultrapura, perfuradas e mantidas em gelo até o término da coleta de todas as glândulas. Em seguida, o extrato salivar foi centrifugado à 16000 x g por 15 min a 4°C para a

obtenção das frações solúvel (sobrenadante) e insolúvel (pellet). Esta última foi descartada. Apenas a fração solúvel foi submetida a SDS-PAGE e transferidas para membrana de nitrocelulose para a realização do teste de reatividade cruzada com o soro anti-RnTrypsin.

4.13. Teste de reatividade dos anticorpos com a saliva de *R.neglectus*

Para averiguar a reatividade dos anticorpos policlonais anti-RnTrypsin, foi realizada a técnica de *Western blot* conforme já descrito, substituindo o anticorpo primário anti-His pelo soro anti-RnTrypsin, nas seguintes diluições: 1/100, 1/200 e 1/400.

5. Resultados

5.1. Serino protease tipo tripsina de *R. neglectus* (RnTrypsin)

5.1.1. Análise *in silico*

O transcrito GDKW01002070.1 (GenBank) descrito no sialoma de *R. neglectus* (SANTIAGO, 2016) codifica uma serino protease tipo tripsina (Uniprot: A0A0P4VQ38) de 244 resíduos de aminoácidos, contendo um peptídeo sinal indicativo de secreção, um peso molecular calculado (MW) de 27 kDa, um ponto isoelétrico teórico (pI) de 9.06 e um índice de instabilidade de 28,19, sendo classificada como estável (**Anexos 6, 7 e 8**). A sequência aminoácídica possui 51 potenciais sítios de fosforilação: seis resíduos de tirosina, 22 de treonina e 23 de serina (**Anexo 9**); dois resíduos de asparagina, nas posições 46 e 138, potenciais de N-glicosilação (**Anexo 10**) e três sítios potenciais de O-glicosilação: 28 (S, serina), 34 (T, treonina) e 52 (T) (**Anexo 11**).

Tabela 3. Resultados da análise *in silico* da serino protease tipo tripsina de *R. neglectus*.

Software	Serino protease tipo tripsina (RnTrypsin)
ProtParam	244 resíduos de aminoácidos, peso molecular 27 kDa, ponto isoelétrico 9.06, índice de instabilidade de 28,19
TargetP	via clássica
SignalP	presença de peptídeo sinal
NetPhos	6 tirosina 22 treonina 23 serina
NetNGlyc	N46 e N138
NetOGlyc	28S 34 e 52T

A sequência aminoacídica da RnTrypsin apresentou maior identidade (71 %) com a sequência AQM58393.1 (GenBank) de uma protease S1 15 do veneno de *Pristhesancus plagipennis* (Walker et al., 2017), um inseto da ordem Heteroptera, como o *R. neglectus*, quando analisada no programa BlastP contra o banco de dados de sequências de proteínas não-redundantes. Contra o mesmo banco, após seleção do organismo (Triatominae – taxid: 70999), RnTrypsin alinhou significativamente com seis sequências: a sequência ABH09442.1 (GenBank) de uma tripsina salivar secretada de *T. brasiliensis*, com a qual apresentou maior identidade (ident; 75 %); seguida da sequência de uma tripsina salivar de *T. infestans* (ABR27829.1; 45 % ident); da Pc97 de *Panstrongylus chinai*, similar a uma serino protease (BBA30652.1; 33 % ident); de uma serino protease de *Panstrongylus megistus* (ADF97756.1; 36 %); de uma serino proteinase de *R. prolixus* (ACF70480.1; 31 %); e de uma proteína não identificada de *T. dimidiata* (BAI53115.1; 38%) (**Anexo 12**). A Figura 9 mostra o alinhamento entre as sequências RnTrypsin, protease S1 15 do veneno de *P. plagipennis* e tripsina salivar secretada de *T. brasiliensis*, na qual é possível observar os resíduos conservados de histidina (H) na posição 93, de ácido aspártico (D) na posição 117 e de serina (S) na posição 195, da tríade catalítica.

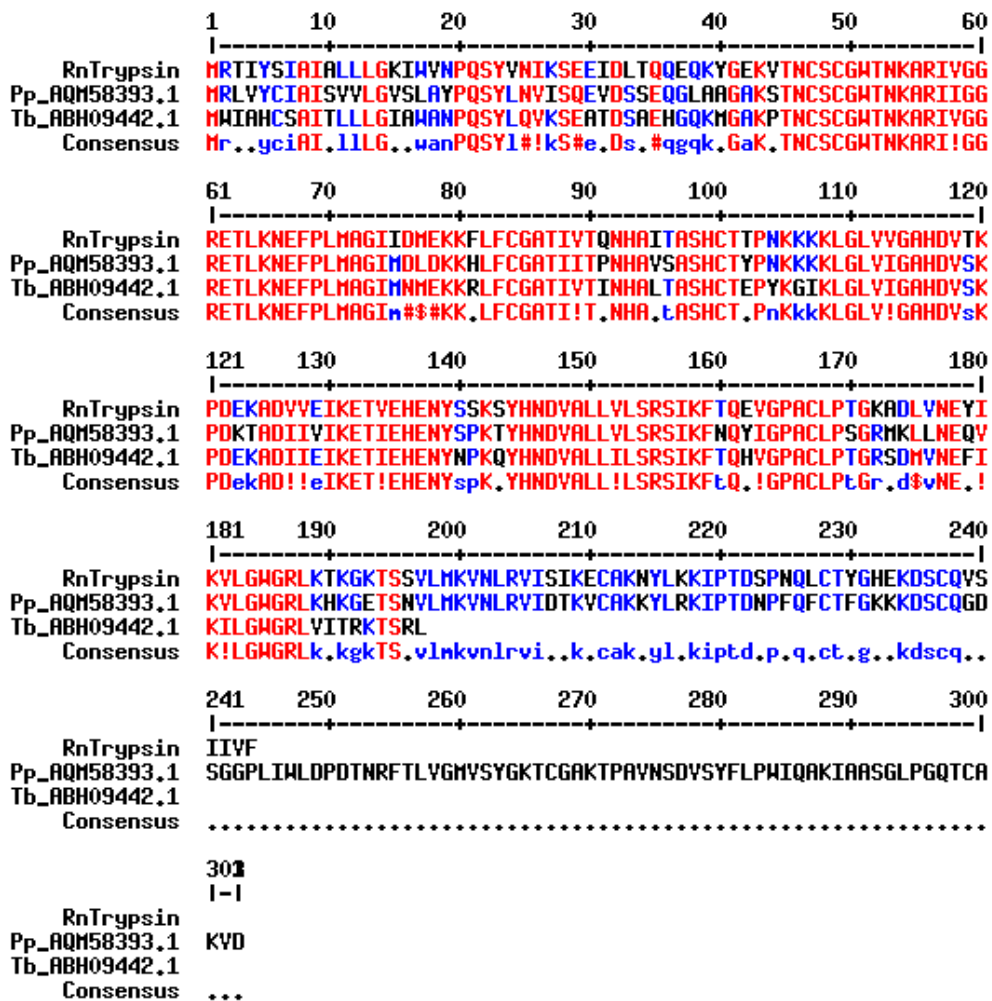


Figura 9. Alinhamento da sequência de aminoácidos da RnTrypsin com serino proteases putativas. O alinhamento foi realizado com as duas sequências do banco de dados de sequências de proteínas não-redundantes do programa BlastP com as quais a RnTrypsin apresentou maior identidade. Pp_AQM58393.1: protease S1 15 do veneno de *P. plagiipennis*; Tb_ABH09442.1: tripsina salivar secretada de *T. brasiliensis*. Resíduos em vermelho e em letra maiúscula na linha *Consensus*: resíduos altamente conservados; resíduos em azul e em letra minúscula na linha *Consensus*: resíduos semiconservados. Alinhamento realizado utilizando o software MultAlin (Corpet, 1988).

5.1.2. Expressão da RnTrypsin

A indução da expressão da RnTrypsin pela *E. coli* BL21(DE3) com 1 mM IPTG por 3 h à 37°C resultou na produção de uma proteína recombinante, fusionada com uma cauda de seis histidinas na porção N-terminal, na fração insolúvel do extrato bacteriano, conforme observado na análise por *Western blot* utilizando anticorpo anti-His (Figura 10 D).

Observou-se uma banda com peso molecular de aproximadamente 30 kDa, que está de acordo com o peso molecular teórico da RnTrypsin contendo a cauda de 6xHis (30.287,36 Da). Tanto na fração solúvel quanto na insolúvel, a RnTrypsin recombinante não foi detectada por *Western blot* antes da adição do agente indutor da expressão (Figura 10 B e D). As Figuras 10 A e C mostram as proteínas totais separadas por eletroforese em gel de poliacrilamida, coradas com azul de Coomassie, das frações solúvel e insolúvel, respectivamente, antes e depois da indução descrita acima. O volume padrão correspondente a 10 µL de amostra foi aplicado em cada poço, a não ser que descrito um volume diferente.

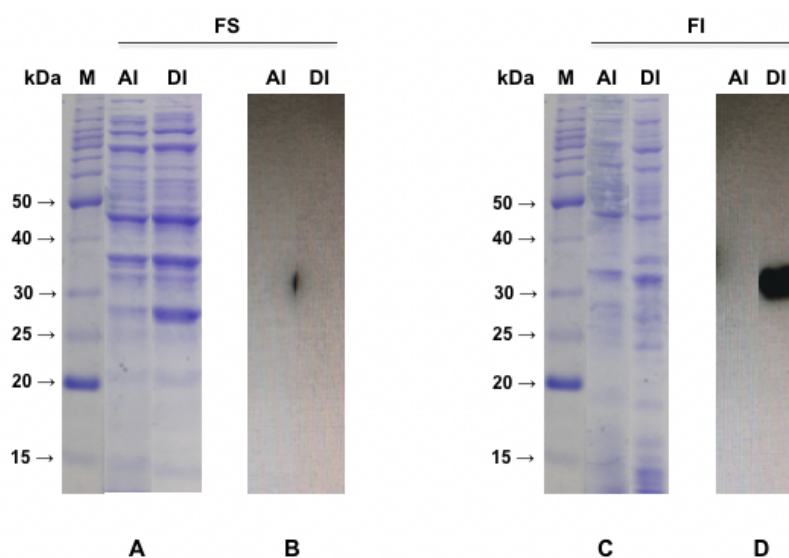


Figura 10. Expressão da proteína recombinante RnTrypsin nas frações solúvel e insolúvel de BL21(DE3) induzidas à 37°C por 3 h com 1 mM IPTG. A - Gel de poliacrilamida 12% corado com azul de Coomassie. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel, (AI) antes da indução, (DI) depois da indução. B - *Western blot*. (FS) fração solúvel, (AI) antes da indução, (DI) depois da indução. C - Gel de poliacrilamida 12% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FI) fração insolúvel, (AI) antes da indução, (DI) depois da indução. D - *Western blot*. (FI) fração insolúvel, (AI) antes da indução, (DI) depois da indução.

Várias tentativas para se obter a RnTrypsin também na fração solúvel foram realizadas. A indução da expressão da RnTrypsin pela mesma

linhagem, BL21(DE3), variando a concentração de IPTG de 1 mM para 0,4 mM, o tempo de indução de 3 h por 22 h e a temperatura de 37°C para 25°C, não resultou na produção de RnTrypsin recombinante detectável nas frações solúvel e insolúvel pela técnica de *Western blot* utilizando anticorpo anti-His (Figura 11 B). A Figura 11 A mostra as proteínas totais, após coloração com azul de Coomassie, separadas por eletroforese em gel de poliacrilamida, das frações solúvel e insolúvel resultantes da indução descrita acima.

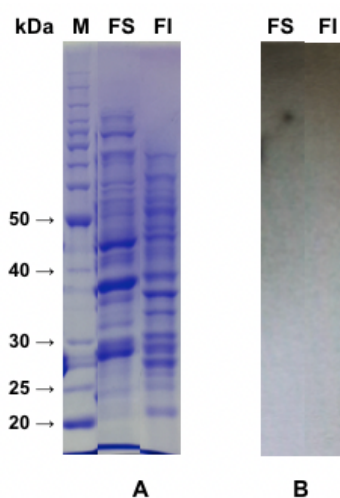


Figura 11. Expressão da proteína recombinante RnTrypsin nas frações solúvel e insolúvel de BL21(DE3) induzidas à 25°C por 22 h com 0,4 mM IPTG. A - Gel de poliacrilamida 12% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel, (FI) fração insolúvel. B - *Western blot*. (FS) fração solúvel, (FI) fração insolúvel.

A indução da expressão da RnTrypsin pela mesma linhagem, BL21(DE3), variando a concentração de IPTG para 0,5 mM, o tempo de indução de 3 h a 5 h à temperatura de 20°C, resultou na produção de RnTrypsin recombinante detectável apenas nas frações insolúveis nos três diferentes períodos de indução testados nesse ensaio (Figura 12 D). Na fração insolúvel após 5 h de indução foi possível observar uma banda abaixo do tamanho esperado para a RnTrypsin recombinante, sugestiva de degradação. Na fração solúvel do controle negativo (C-; BL21(DE3) transformada com pET100/D-TOPO_RnKazalPI) foi possível detectar uma banda de baixa intensidade na faixa de peso molecular entre 15 kDa e 20

kDa pela técnica de *Western blot* utilizando anticorpo anti-His (Figura 12 B), cuja investigação será discutida na seção sobre um inibidor de serino protease tipo Kazal. As Figuras 12 A e C mostram as proteínas totais separadas por eletroforese em gel de poliacrilamida, coradas com azul de Coomassie, das frações solúvel e insolúvel, respectivamente, após indução descrita acima.

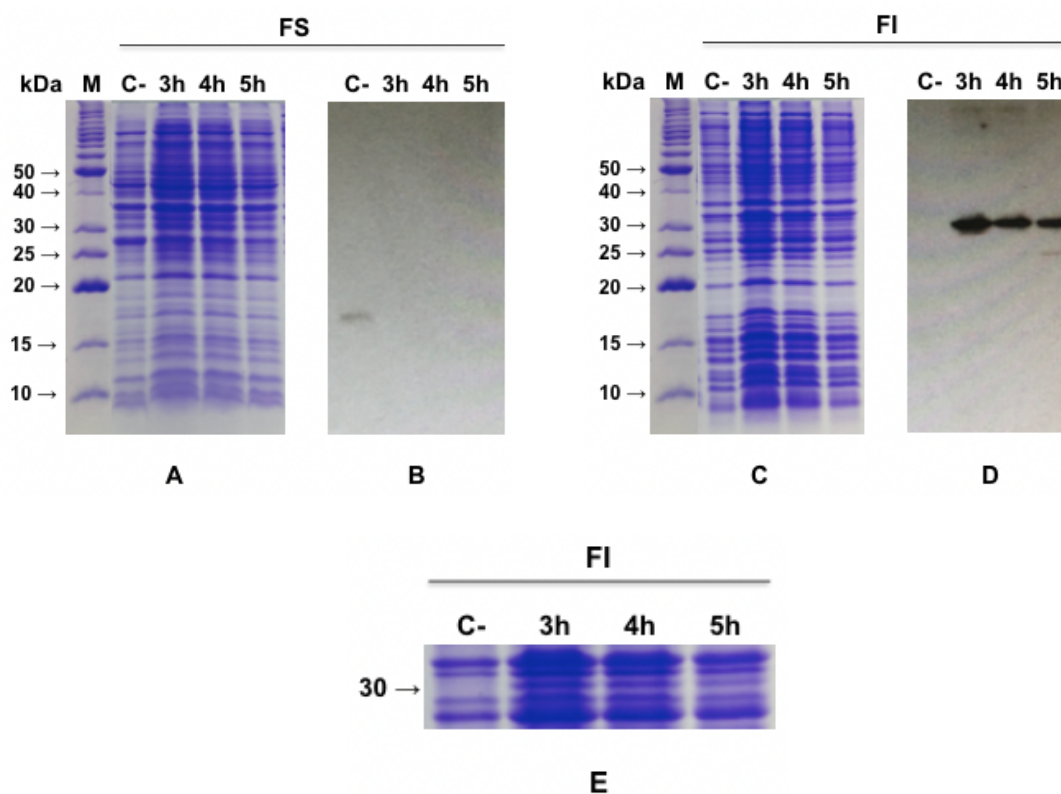


Figura 12. Análise da expressão da proteína recombinante RnTrypsin nas frações solúvel e insolúvel de BL21(DE3) induzidas à 20°C por 3 à 5 h com 0,5 mM IPTG. Volumes variáveis entre 4 e 12 μ L de amostra foram aplicadas em cada poço (3h - 12 μ L, 4h - 8 μ L, 5h - 4 μ L). A e B: correspondem a (FS) fração solúvel. A - Gel de poliacrilamida 15% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (C-) controle negativo: mesmo plasmídeo porém com a sequência da Kazal, (3 à 5h) tempo de indução. B - *Western blot*. (C-) controle negativo, (3 à 5h) tempo de indução. C e D: correspondem a (FI) fração insolúvel. C - Gel de poliacrilamida 15% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (C-) controle negativo: mesmo plasmídeo porém com a sequência da Kazal, (3 à 5h) tempo de indução. D - *Western blot*. (C-) controle negativo, (3 à 5h) tempo de indução. E - Aumento da Figura C para mostrar a banda de aproximadamente 30

kDa que está presente na expressão da (FI) fração insolúvel da RnTrypsin e ausente na expressão da (FI) fração insolúvel do (C-) controle negativo RnKazalPI.

Mantendo-se a temperatura em 20 °C e a concentração do indutor em 0,5 mM, o tempo de indução foi estendido para 16 h, condição em que se observou a RnTrypsin recombinante somente na fração insolúvel (Figura 13 B). A Figura 13 A mostra a separação por eletroforese em gel de poliácridamida das frações solúvel e insolúvel das proteínas totais resultantes da indução descrita acima.

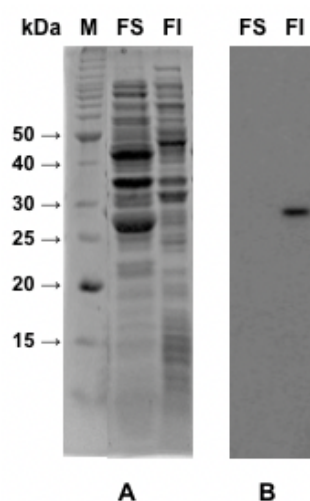


Figura 13. Expressão da proteína recombinante RnTrypsin nas frações solúvel e insolúvel de BL21(DE3) induzidas à 20°C por 16 h com 0,5 mM IPTG. A - Gel de poliácridamida 12% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel, (FI) fração insolúvel. B - *Western blot*. (FS) fração solúvel, (FI) fração insolúvel.

A figura abaixo mostra a indução da expressão da RnTrypsin em BL21(DE3) por 5 h em duas condições diferentes, com 1 mM IPTG à 37°C e com 0,5 mM IPTG à 20°C. A RnTrypsin recombinante só foi observada na fração insolúvel em ambas as situações (Figura 14 B). A Figura 14 A é referente à um gel de poliácridamida contendo as frações solúvel e insolúvel das proteínas totais induzidas como descrito acima, separadas por eletroforese e coradas com azul de Coomassie.

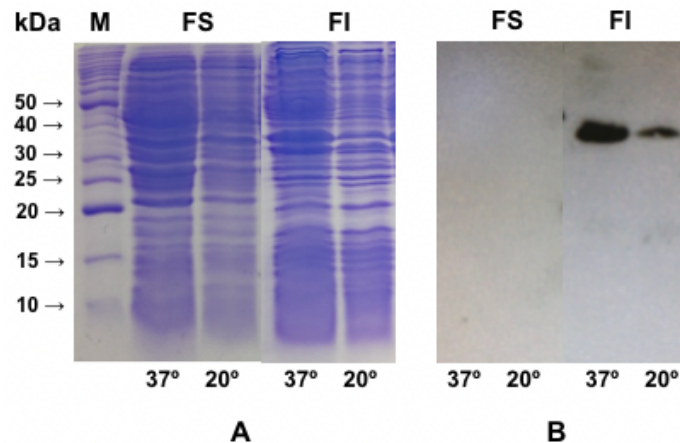


Figura 14. Expressão da proteína recombinante RnTrypsin nas frações solúvel e insolúvel de BL21(DE3) induzidas à 20°C por 5 h com 0,5 mM IPTG e à 37°C por 5 h com 1 mM IPTG. A - Gel de poliacrilamida 15% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel, (FI) fração insolúvel. B - *Western blot*. (FS) fração solúvel, (FI) fração insolúvel.

A indução à uma temperatura ainda menor, por um tempo superior e com uma concentração mais baixa de IPTG foi avaliada na tentativa de se obter a proteína recombinante na fração solúvel. Sendo assim, as condições escolhidas foram as de 18°C por 24 e 40 h com 0,4 mM IPTG (Figura 15) e 15°C por 20 h com 0,2 e 0,4 mM IPTG (Figura 16). Os perfis proteicos nos géis corados com Coomassie não mostraram grandes diferenças entre as duas concentrações utilizadas de indutor. Em ambas as condições de concentração do indutor IPTG testadas à 15°C, a RnTrypsin não foi visualizada em nenhuma das frações (FI e FS) (Figura 16 B). É possível observar na Figura 16 A as frações insolúvel e solúvel das proteínas totais resultantes da indução descrita acima, separadas por eletroforese em gel de poliacrilamida e coradas com azul de Coomassie.

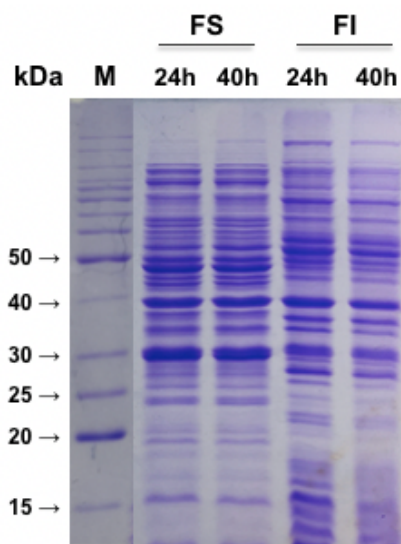


Figura 15. Expressão da proteína recombinante RnTrypsin nas frações solúvel e insolúvel de BL21(DE3) induzidas à 18°C por 24 e 40 h com 0,4 mM IPTG. O volume de 5 μ L de amostra foi aplicada em cada poço. Gel de poliacrilamida 12% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel após 24 e 40h de indução, (FI) fração insolúvel após 24 e 40h de indução.

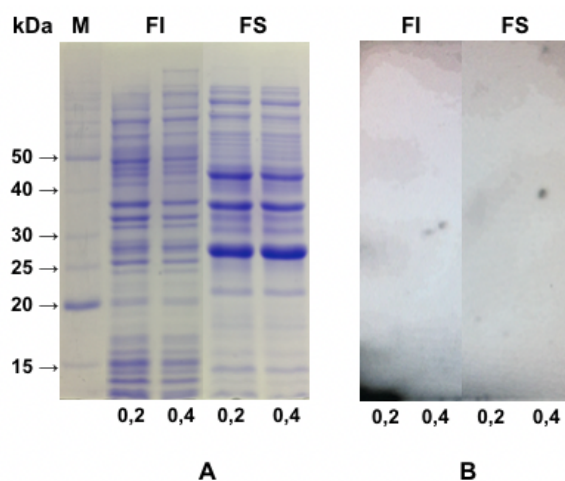


Figura 16. Expressão da proteína recombinante RnTrypsin nas frações solúvel e insolúvel de BL21(DE3) induzidas à 15°C por 20 h com 0,2 e 0,4 mM IPTG. A - Gel de poliacrilamida 12% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FI) fração insolúvel 0,2 e 0,4 mM IPTG, (FS) fração solúvel 0,2 e 0,4 mM IPTG. B - *Western blot*. (FI) fração insolúvel 0,2 e 0,4 mM IPTG, (FS) fração solúvel 0,2 e 0,4 mM IPTG.

Outra variação no protocolo de expressão na tentativa de se obter a RnTrypsin na fração solúvel incluiu o uso de outras linhagens de *E. coli*: a BL21(DE3)pLysS e a Rosetta(DE3)pLysS. A BL21(DE3)pLysS transformada com o pET100/D-TOPO_RnTrypsin não expressou a proteína em nenhuma das frações (FI e FS) nas duas condições testadas: indução com 1 mM IPTG por 3 h à 37°C (Figura 17 B) e com 0,5 mM IPTG por 16 h à 20°C (Figura 17 D). As Figuras 17 A e C correspondem a separação por eletroforese em gel de poliacrilamida das frações solúveis e insolúveis das proteínas produzidas nas condições citadas acima, respectivamente.

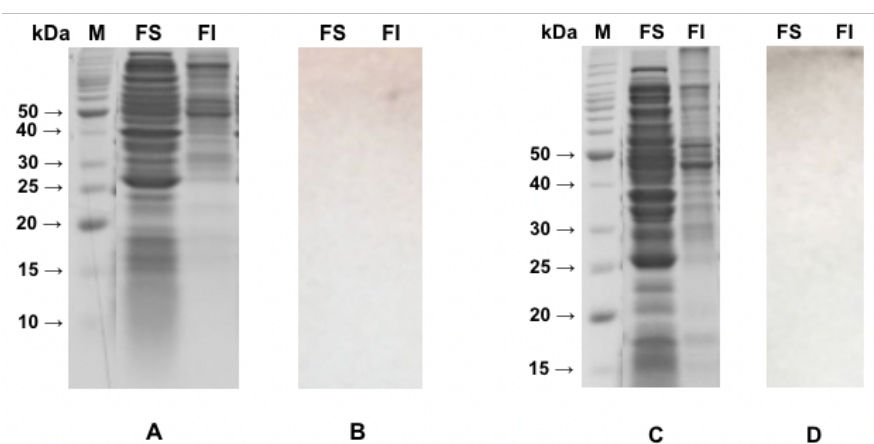


Figura 17. Expressão da proteína recombinante RnTrypsin nas frações solúvel e insolúvel de BL21(DE3)pLysS induzidas à 37°C por 3 h com 1 mM IPTG e à 20°C por 16 h com 0,5 mM IPTG. A e B: 37°C por 3 h com 1 mM IPTG. A - Gel de poliacrilamida 15% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel, (FI) fração insolúvel. B - *Western blot*. (FS) fração solúvel, (FI) fração insolúvel. C e D: 20°C por 16 h com 0,5 mM IPTG. C - Gel de poliacrilamida 12% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel, (FI) fração insolúvel. D - *Western blot*. (FS) fração solúvel, (FI) fração insolúvel.

A linhagem de *E. coli* Rosetta(DE3)pLysS transformada com o pET100/D-TOPO_RnTrypsin expressou a proteína somente na fração insolúvel após indução com 1 mM IPTG por 3 h à 37°C (Figura 18 B). Nas outras condições testadas para essa linhagem, indução com 0,4 mM IPTG

por 22 h à 25°C (Figura 18 D), com 0,5 mM IPTG por 16 h à 20°C (Figura 18 F) e com 0,2 e 0,4 mM IPTG por 22 h a 15°C (Figura 18 H), não foi observada expressão da RnTrypsin recombinante em nenhuma das frações bacterianas. As Figuras 18 A, C, E e G correspondem a separação por eletroforese em gel de poliácridamida das frações solúveis e insolúveis das proteínas produzidas nas condições citadas acima, respectivamente. Todos os géis foram corados com azul de Coomassie Blue.

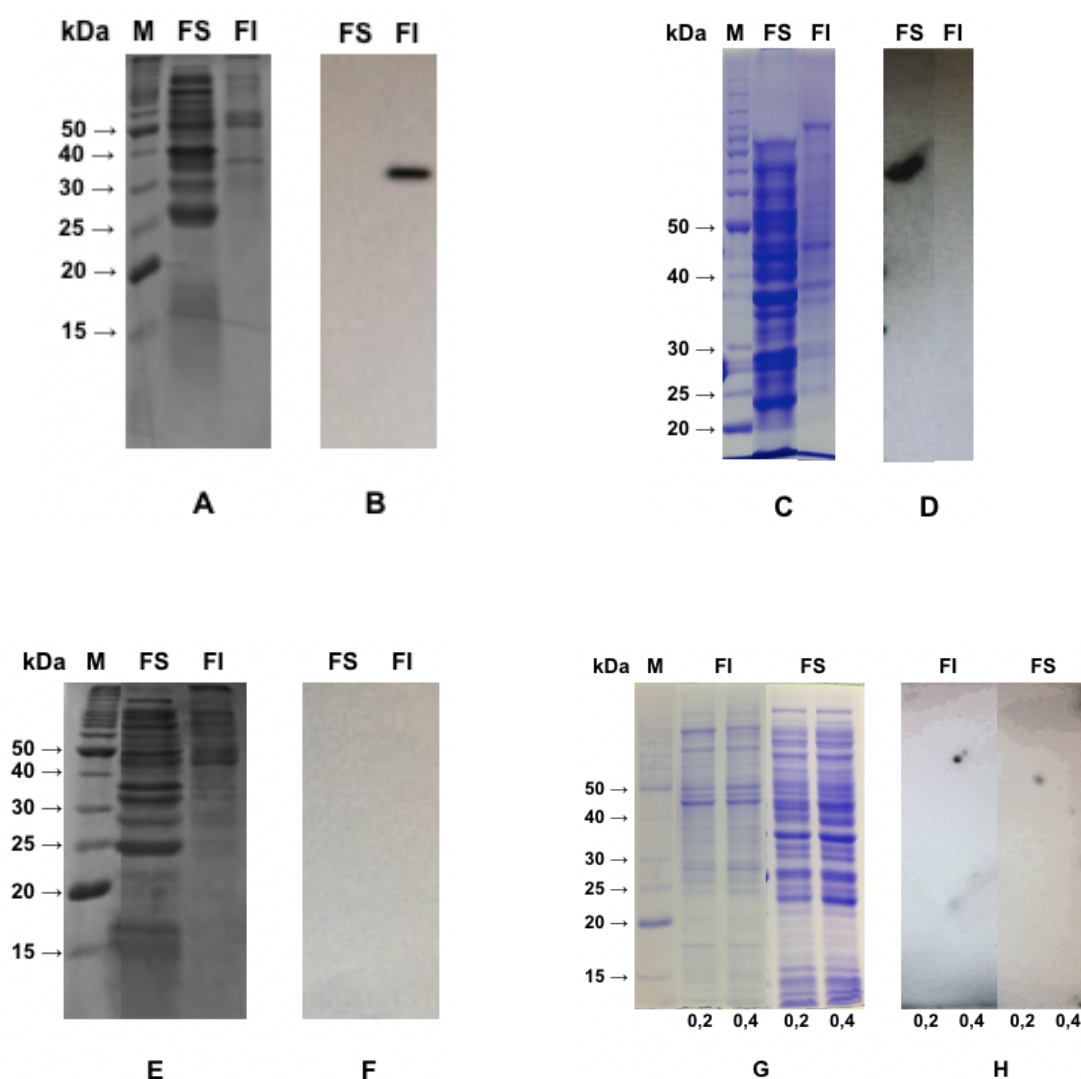


Figura 18. Expressão da proteína recombinante RnTrypsin nas frações solúvel e insolúvel de Rosetta(DE3)pLysS induzidas à 37°C por 3 h com 1 mM IPTG, à 25°C por 22 h com 0,4 mM IPTG, à 20°C por 16 h com 0,5 mM IPTG e à 15°C por 22 h com 0,2 e 0,4 mM IPTG. A - Gel de poliácridamida 15% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel, (FI) fração insolúvel. B - *Western blot*. (FS) fração solúvel, (FI) fração insolúvel. C - Gel de poliácridamida 12% corado com azul de Coomassie Blue. D - *Western blot*. (FS) fração solúvel, (FI) fração insolúvel. E - Gel de poliácridamida 15% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel, (FI) fração insolúvel. F - *Western blot*. (FS) fração solúvel, (FI) fração insolúvel. G - Gel de poliácridamida 12% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FI) fração insolúvel, (FS) fração solúvel. H - *Western blot*. (FI) fração insolúvel, (FS) fração solúvel.

(M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel, (FI) fração insolúvel. D - *Western blot*. (FS) fração solúvel, (FI) fração insolúvel. E - Gel de poliacrilamida 15% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel, (FI) fração insolúvel. F - *Western blot*. (FS) fração solúvel, (FI) fração insolúvel. G - Gel de poliacrilamida 12% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FI) fração insolúvel 0,2 e 0,4 mM IPTG, (FS) fração solúvel 0,2 e 0,4 mM IPTG. H - *Western blot*. (FI) fração insolúvel 0,2 e 0,4 mM IPTG, (FS) fração solúvel 0,2 e 0,4 mM IPTG.

Tabela 4. Condições de expressão da proteína RnTrypsin recombinante em sistema de expressão bacteriano.

Linagem de <i>E. coli</i>	IPTG (mM)	Tempo (horas)	Temperatura (°C)	FS	FI
BL21(DE3)	1	3	37	-	X
BL21(DE3)	0,4	22	25	-	-
BL21(DE3)	0,5	3 à 5	20	-	X
BL21(DE3)	0,5	16	20	-	X
BL21(DE3)	0,2 e 0,4	20	15	-	-
BL21(DE3)pLysS	1	3	37	-	-
BL21(DE3)pLysS	0,5	16	20	-	-
Rosetta(DE3)pLysS	1	3	37	-	X
Rosetta(DE3)pLysS	0,4	22	25	-	-
Rosetta(DE3)pLysS	0,5	16	20	-	-
Rosetta(DE3)pLysS	0,2 e 0,4	22	15	-	-

5.1.3. Purificação da RnTrypsin

Como não foi possível expressar a RnTrypsin recombinante na fração solúvel do lisado bacteriano em nenhuma das três linhagens de *E. coli* sob as diferentes condições testadas, a RnTrypsin recombinante foi purificada dos corpos de inclusão da fração insolúvel da BL21(DE3) induzida com 1 mM de IPTG por 3 h à 37 °C, devido a essas condições serem as condições padrão de indução e também devido ao fato de ter aparecido uma marcação bem evidente no *Western blot* mostrado na Figura 10 D.

A RnTrypsin recombinante se ligou fracamente à matriz na cromatografia de afinidade e começou a ser eluída com uma baixa concentração de imidazol (5mM), como pode ser observado na Figura 19 A. Consequentemente, a amostra tinha uma alta concentração de impurezas. Nas frações da purificação, confirmada pela técnica de *Western blot*, observou-se uma banda com peso molecular de aproximadamente 30 kDa, que está de acordo com o peso molecular teórico da RnTrypsin contendo a cauda de 6xHis (30.287,36 Da) (Figura 19 B). Essas frações foram utilizadas para a produção de anticorpos policlonais.

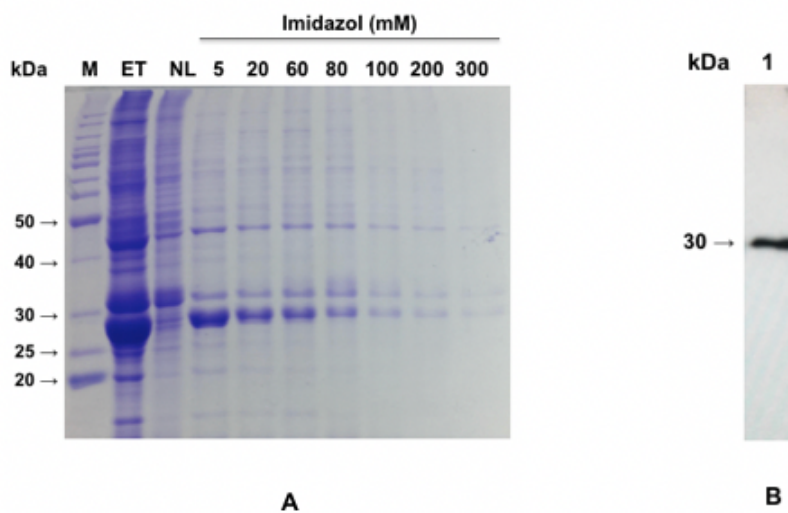


Figura 19. Purificação dos corpos de inclusão da fração insolúvel de RnTrypsin por cromatografia de afinidade em níquel, utilizando His-tag; e análise por *Western blot*, utilizando anticorpo primário anti-His, da proteína recombinante RnTrypsin eluída com 5 mM de imidazol. A - Gel de poli-acrilamida 12% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (ET) extrato total, (NL) fração não ligada, (5 a 300) diferentes concentrações de imidazol e suas eluições. B - *Western blot* da eluição na fração contendo 5 mM de Imidazol.

5.1.4. Produção de anticorpos

A serino protease do tipo tripsina salivar nativa de *R. neglectus*, sem o peptídeo sinal, apresenta peso molecular de aproximadamente 25.148,01 Da. Apenas o soro pré-imune do camundongo 4 reconheceu uma banda na altura do peso molecular esperado (Figura 20 A). Portanto, a produção de anticorpos policlonais contra a RnTrypsin recombinante capazes de reconhecer a serino protease do tipo tripsina nativa na saliva de *R. neglectus* não foi observada no soro de nenhum dos animais imunizados nas diluições avaliadas (Figura 20).

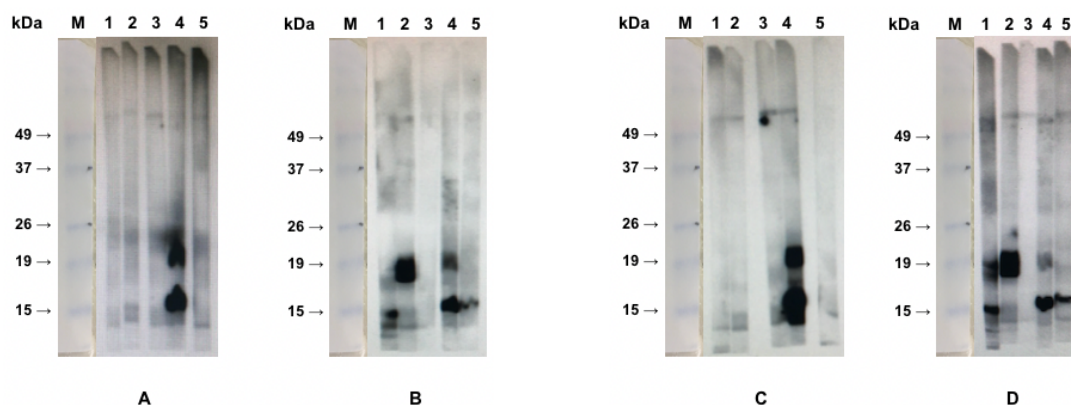


Figura 20. *Western blot* dos soros pré-imune e imune anti-RnTrypsin produzidos por 5 camundongos, diluídos 1:400 e 1:100. Existe aproximadamente 30 µg de saliva total de *R. neglectus* por fita. A e B: diluição 1:400. A - Soros pré-imune. (M) marcador (kDa) BenchMark™ Pre-Stained Protein Ladder – Invitrogen, (1 à 5) numeração referente aos camundongos. B - Soros imune. (M) marcador (kDa) BenchMark™ Pre-Stained Protein Ladder – Invitrogen, (1 à 5) numeração referente aos camundongos. C e D: diluição 1:100. C - Soros pré-imune. (M) marcador (kDa) BenchMark™ Pre-Stained Protein Ladder – Invitrogen, (1 à 5) numeração referente aos camundongos. D - Soros imune. (M) marcador (kDa) BenchMark™ Pre-Stained Protein Ladder – Invitrogen, (1 à 5) numeração referente aos camundongos.

Entretanto, houve produção de anticorpos contra as proteínas inoculadas durante as imunizações, entre elas a RnTrypsin, como pode ser observado no soro do camundongo 2 (Figura 21).

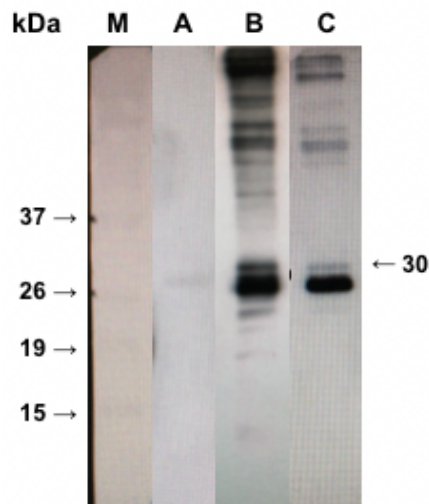


Figura 21. *Western blot* dos soros pré-imune e imune anti-RnTrypsin do camundongo nº 2, diluídos 1:200, e do anticorpo Anti-His, diluído 1:2000, utilizados para avaliar reatividade com a eluição na fração contendo 5 mM de Imidazol resultante da purificação. (M) marcador (kDa) BenchMark™ Pre-Stained Protein Ladder – Invitrogen. A - Soro pré-imune. B - Soro imune. C - Anticorpo Anti-His.

5.1.5. Perfil da saliva de *R. neglectus*

A Figura 22 mostra o perfil eletroforético da saliva de *R. neglectus* em diferentes diluições. Há uma concentração de proteínas entre 15 kDa e 40 kDa. Entretanto, na altura de 25 kDa, peso molecular esperado para a RnTrypsin secretada, mesmo no gel corado por nitrato de prata, é possível observar uma escassez de bandas, sugerindo que a serino protease salivar de *R. neglectus* é pouco expressa em comparação com outras proteínas salivares.

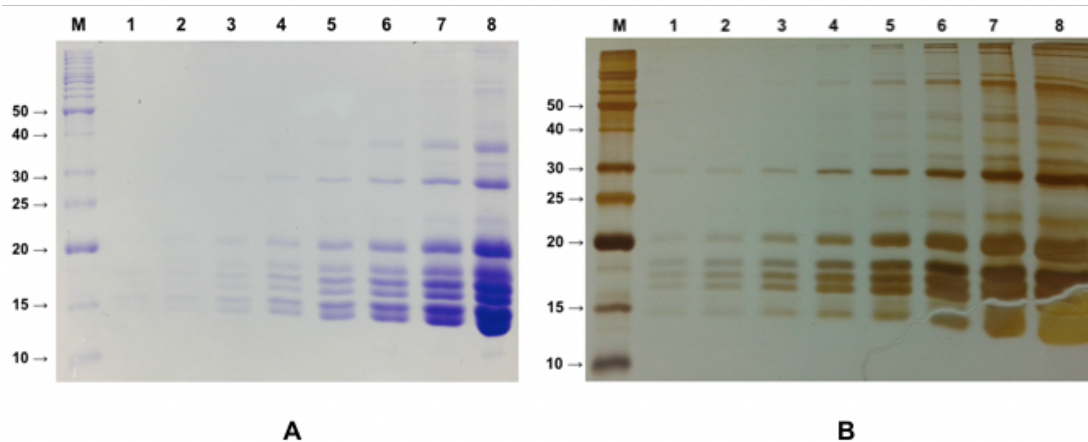


Figura 22. Expressão das proteínas totais do extrato salivar de *R. neglectus* em gel de poliacrilamida 12%. A - (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (1 a 8) diluições da amostra; corado com azul de Coomassie Blue. B -

(M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (1 a 8) diluições da amostra; corado com nitrato de prata.

5.2. Inibidor de serino protease tipo Kazal de *R. neglectus*

5.2.1. Análise *in silico*

O transcrito GDKW01000509.1 (GenBank) descrito no sialoma de *R. neglectus* (SANTIAGO, 2016) codifica para um inibidor de serino protease tipo Kazal (Uniprot: A0A0P4W3X1) de 83 resíduos de aminoácidos, contendo um peptídeo sinal indicativo de secreção, um peso molecular calculado (MW) de 9.381,68 Da, um ponto isoelétrico teórico de 4.39 e um o índice de instabilidade de 39,43, sendo classificada como estável (**Anexos 13, 14 e 15**). A sequência aminoácídica possui 29 potenciais sítios de fosforilação: dois resíduos de tirosina, 13 de treonina e 14 de serina (**Anexo 16**); nenhum sítio predito de N-glicosilação (**Anexo 17**) ou O-glicosilação (**Anexo 18**)

Tabela 5. Resultados da análise *in silico* do inibidor de serino protease tipo Kazal de *R. neglectus*.

Software	Inibidor de serino protease tipo Kazal (RnKazalPI)
ProtParam	83 resíduos de aminoácidos, peso molecular 9 kDa, ponto isoelétrico 4.39, índice de instabilidade de 39,43
TargetP	via clássica
SignalP	presença de peptídeo sinal
NetPhos	2 tirosina 13 treonina 14 serina
NetNGlyc	-
NetOGlyc	-

Quando analisada no programa BlastP contra o banco de dados de sequências de proteínas não-redundantes, com ou sem seleção do organismo (Triatominae – taxid: 70999), a sequência aminoacídica da RnKazalPI alinhou significativamente com quatro sequências: BBA30668.1 (GenBank) da Pc266, similar a um inibidor de serino protease tipo Kazal de *Panstrongylus chinai*, com a qual apresentou maior identidade (ident; 48 %); seguida das sequências de inibidores de serino protease tipo Kazal de *Triatoma infestans* (ABR27937.1 - 46 % ident; e ABR27896.1 – 42 %); e da Pc77 de *P. chinai*, similar a um inibidor de serino protease tipo Kazal (BBA30649.1; 47 % ident) (**Anexo 19**). A Figura 23 mostra o alinhamento entre as cinco sequências.

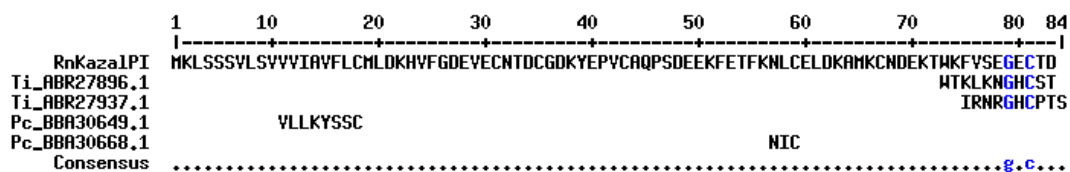


Figura 23. Alinhamento da sequência de aminoácidos da RnKaza1PI com inibidores de serino proteases tipo Kazal putativos. O alinhamento foi realizado com as sequências do banco de dados de sequências de proteínas não-redundantes do programa BlastP com as quais a RnKaza1PI apresentou maior identidade. Pc_BBA30668.1: Pc266, similar a um inibidor de serino protease tipo Kazal de *P. chinai*; Ti_ABR27937.1 e Ti_ABR27896.1: inibidores de serino protease tipo Kazal de *T. infestans*; e Pc_BBA30649.1: Pc77 de *P. chinai*, similar a um inibidor de serino protease tipo Kazal. Resíduos em azul e em letra minúscula na linha *Consensus*: resíduos semiconservados. Alinhamento realizado utilizando o programa MultAlin (Corpet, 1988).

5.2.2. Expressão da RnKaza1PI

A indução da expressão da RnKaza1PI pela *E. coli* BL21(DE3) com 1 mM IPTG por 3 h à 37°C não resultou na produção da proteína recombinante detectável nas frações solúvel e insolúvel pela técnica de *Western blot* utilizando anticorpo anti-His (Figura 24 B). A Figura 24 A mostra as frações solúvel e insolúvel das proteínas totais resultantes da indução descrita acima, separadas por eletroforese em gel de poliacrilamida.

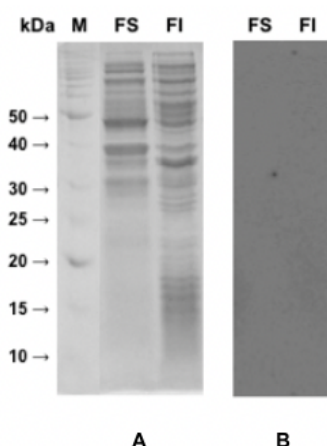


Figura 24. Expressão da proteína recombinante RnKaza1PI nas frações solúvel e insolúvel de BL21(DE3) induzidas à 37°C por 3 h com 1 mM IPTG. A - Gel de poliacrilamida 12% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel, (FI) fração insolúvel. B - *Western blot*. (FS) fração solúvel, (FI) fração insolúvel.

Uma primeira variação nas condições de indução foi testada, buscando produzir a RnKazalPI à 20°C por 5 h com 0,5 mM IPTG. Uma marcação foi visualizada na (FS) fração solúvel do *Western blot* correspondente ao peso molecular aproximado de 18 kDa, podendo indicar a presença da proteína expressa em forma de dímero (Figura 25 B). Para ter a conclusão, é necessário a realização de mais testes. A Figura 25 A mostra as frações solúvel e insolúvel das proteínas totais resultantes da indução descrita acima, separadas por eletroforese em gel de poliacrilamida e coradas com azul de Coomassie.

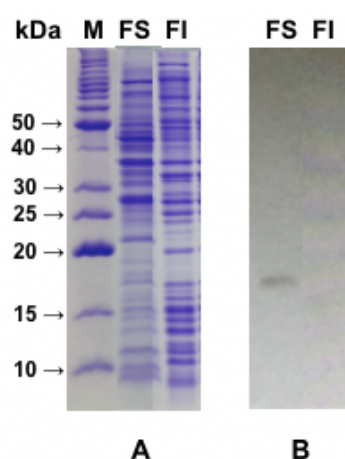


Figura 25. Expressão da proteína recombinante RnKazalPI nas frações solúvel e insolúvel de BL21(DE3) induzidas à 20°C por 5 h com 0,5 mM IPTG. O volume de 4 μ L de amostra foi aplicada em cada poço. A - Gel de poliacrilamida 15% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel, (FI) fração insolúvel. B - *Western blot*. (FS) fração solúvel, (FI) fração insolúvel.

Quando as condições de temperatura e concentração do indutor foram mantidas mas o tempo de indução foi aumentado de 5 h para 16h, nenhuma marcação foi observada (Figura 26 B). É possível observar na Figura 26 A as frações solúvel e insolúvel das proteínas totais resultantes da indução descrita acima, separadas por eletroforese em gel de poliacrilamida e coradas com azul de Coomassie.

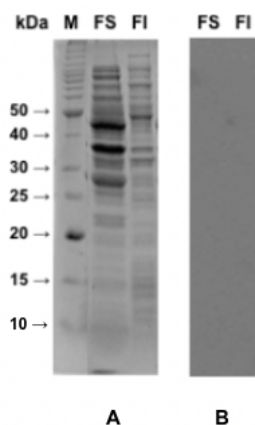


Figura 26. Expressão da proteína recombinante RnKazalPI nas frações solúvel e insolúvel de BL21(DE3) induzidas à 20°C por 16 h com 0,5 mM IPTG. A - Gel de poliacrilamida 12% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel, (FI) fração insolúvel. B - *Western blot*. (FS) fração solúvel, (FI) fração insolúvel.

Outras variações no protocolo de expressão na tentativa de se obter a RnKazalPI na fração solúvel incluíram o uso de outras linhagens de *E. coli*: a BL21(DE3)pLysS e a Rosetta(DE3)pLysS. A BL21(DE3)pLysS transformada com o pET100/D-TOPO_RnKazalPI não expressou a proteína em nenhuma das frações (FI e FS) nas duas condições testadas: indução com 1 mM IPTG por 3 h à 37°C (Figura 27 B) e com 0,5 mM IPTG por 16 h à 20°C (Figura 27 D). As Figuras 27 A e C correspondem a separação por eletroforese em gel de poliacrilamida das frações solúveis e insolúveis das proteínas produzidas nas condições citadas acima, respectivamente.

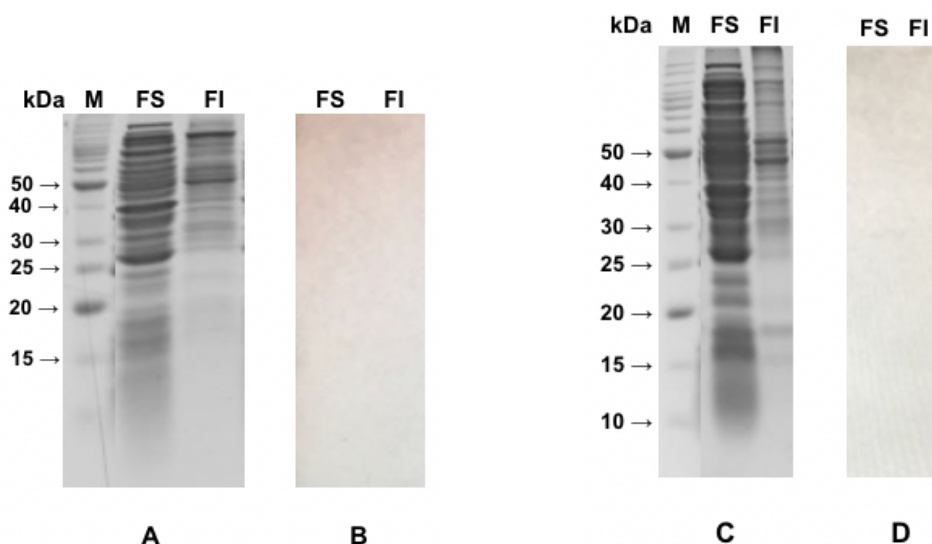


Figura 27. Expressão da proteína recombinante RnKazalPI nas frações solúvel e insolúvel de BL21(DE3)pLysS induzidas à 37°C por 3 h com 1 mM IPTG e à 20°C por 16 h com 0,5 mM IPTG. A - Gel de poliacrilamida 15% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel, (FI) fração insolúvel. B - *Western blot*. (FS) fração solúvel, (FI) fração insolúvel. C - Gel de poliacrilamida 12% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel, (FI) fração insolúvel. D - *Western blot*. (FS) fração solúvel, (FI) fração insolúvel.

Em uma última tentativa, foi utilizado a Rosetta(DE3)pLysS para tentar expressar a proteína recombinante RnKazalPI. As primeiras condições de indução foram com 1 mM IPTG por 3 h à 37°C. É possível ver uma pequena marcação na (FI) fração insolúvel perto da altura esperada do peso molecular da RnKazalPI. Porém, como a marcação tem baixa intensidade e está um pouco abaixo da altura correspondente ao peso molecular de 10 kDa, é necessário um novo teste para ter a confirmação (Figura 28 B). Outras condições de indução foram testadas: com 0,5 mM IPTG por 16 h à 20°C. Entretanto, não foi possível visualizar a proteína RnKazalPI em nenhuma das frações (FI e FS) (Figura 28 D). É possível observar a separação por eletroforese em gel de poliacrilamida das frações solúveis e insolúveis das proteínas produzidas nas condições citadas acima, respectivamente (Figuras 28 A e C).

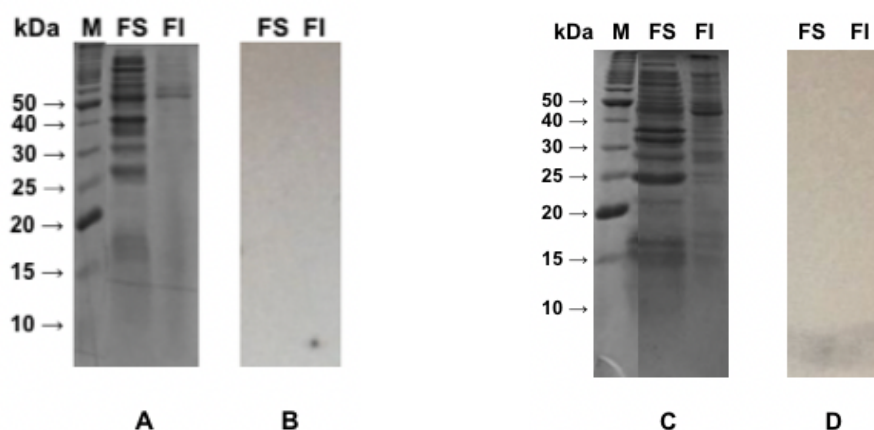


Figura 28. Expressão da proteína recombinante RnKazalPI nas frações solúvel e insolúvel de Rosetta(DE3)pLysS induzidas à 37°C por 3 h com 1 mM IPTG e à 20°C por 16 h com 0,5 mM IPTG. A - Gel de poliacrilamida 15% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel, (FI) fração insolúvel. B - *Western blot*. (FS) fração solúvel, (FI) fração insolúvel. C - Gel de poliacrilamida 15% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel, (FI) fração insolúvel. D - *Western blot*. (FS) fração solúvel, (FI) fração insolúvel.

A expressão do RnKazalPI foi avaliada também após transformação da linhagem Rosetta(DE3)pLysS de *E. coli* com o plasmídeo pET100/D-TOPO_RnKazalPI. A PCR de colônia de seis colônias resultantes da transformação mostra a amplificação de uma banda de DNA entre 400 pb e 500 pb, coerente com o tamanho esperado do inserto (453 pb) (Figura 29).

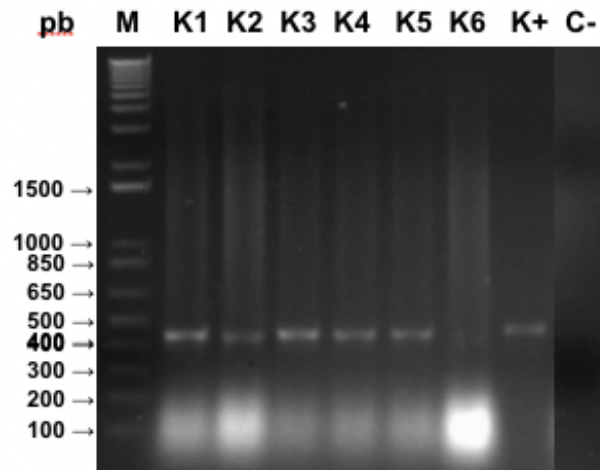


Figura 29. Identificação de colônias transformadas de *E. coli* Rosetta(DE3)pLysS transformada com o plasmídeo pET100/D-TOPO_RnKazalPI. Gel de agarose 1,2% corado com brometo de etídeo. (M) marcador (pb) 1 Kb Plus DNA Ladder – Invitrogen, (K1 à K6) colônias escolhidas aleatoriamente, (K+) controle positivo contendo 10 ng do plasmídeo pET100/D-TOPO_Kazal, (C-) controle negativo.

O clone K3 foi induzido com 1 mM IPTG por 3 h à 18°C. A RnKazalPI recombinante foi então expressa na fração solúvel dessa linhagem bacteriana (Figura 30).

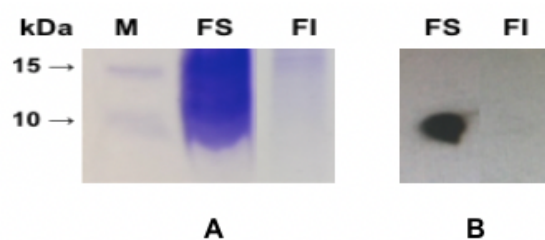


Figura 30. Expressão da proteína recombinante RnKazalPI nas frações solúvel e insolúvel de Rosetta(DE3)pLysS (mesmo clone da K3 - vide figura 25) induzidas à 18°C por 3 h com 1 mM IPTG. A - Gel de poliacrilamida 15% corado com azul de Coomassie Blue. (M) marcador (kDa) BenchMark™ Protein Ladder – Invitrogen, (FS) fração solúvel, (FI) fração insolúvel. B - *Western blot*. (FS) fração solúvel, (FI) fração insolúvel.

Tabela 6. Condições de expressão da proteína RnKazalPI recombinante em sistema de expressão bacteriano.

Linhagem de <i>E. coli</i>	IPTG (mM)	Tempo (horas)	Temperatura (°C)	FS	FI
BL21(DE3)	1	3	37	-	-
BL21(DE3)	0,5	5	20	-	-
BL21(DE3)	0,5	16	20	-	-
BL21(DE3)pLysS	1	3	37	-	-
BL21(DE3)pLysS	0,5	16	20	-	-
Rosetta(DE3)pLysS	1	3	37	-	-
Rosetta(DE3)pLysS	0,5	16	20	-	-
Rosetta(DE3)pLysS	1	3	18	X	-

6. Discussão

As proteínas salivares de triatomíneos, insetos hematófagos, têm grande potencial biotecnológico devido à suas possíveis funções antihemostáticas. Nesse estudo, uma serino protease tipo tripsina (RnTrypsin) e um inibidor de serino protease tipo Kazal (RnKazalPI) salivares de *R. neglectus* foram parcialmente caracterizados *in silico* e expressos na forma recombinante em sistema de expressão bacteriano. A primeira proteína foi expressa apenas de forma insolúvel. Já a segunda foi observada, através da técnica de Western blot, na fração solúvel do lisado bacteriano.

Inicialmente, foram utilizados softwares para fazer previsões sobre as sequências de aminoácidos que permitiram obter informações, mesmo que hipotéticas, sobre as proteínas ainda não caracterizadas. Os resultados gerados ajudaram a guiar o trabalho experimental e a entender melhor alguns dos resultados obtidos nas tentativas de expressão das proteínas recombinantes. Por exemplo, a glicosilação é uma modificação pós-traducional importante, capaz de influenciar o enovelamento, endereçamento, solubilidade, antigenicidade, atividade, vida média e as interações das proteínas (Xu e Ng, 2015). As proteínas secretadas, como as estudadas nessa dissertação, são comumente glicosiladas no retículo endoplasmático. A sequência de aminoácidos do RnKazalPI não apresenta sítios putativos de glicosilação, sugerindo que o peptídeo não sofre esse processo pós-traducional, o que pode ter facilitado sua expressão na fração solúvel pela *E. coli*. No entanto, a sequência da RnTrypsin apresenta dois resíduos potenciais de *N*-glicosilação (asparagina (N) 46 e 138) e três sítios potenciais de *O*-glicosilação (S28, T34 e T52), fato que pode ter impedido sua expressão de forma solúvel e ativa pela bactéria, pois embora as manoses (*O*-glicosilação) sejam estruturalmente mais simples, os glicanos, que se ligam à asparagina (*N*-glicosilação), são estruturas oligoméricas importantes para o enovelamento das proteínas (Xu e Ng, 2015).

O sistema de expressão em *Escherichia coli* é amplamente utilizado para a produção de proteínas recombinantes devido à facilidade de

manipulação genética, o baixo custo e o crescimento bacteriano rápido (Costa *et al.*, 2014; Jia *et al.*, 2016). No entanto, é necessário encontrar a combinação adequada de componentes e condições (linhagens de expressão, vetores de expressão, condições de crescimento, purificação, etc) para se obter uma proteína recombinante altamente expressa, solúvel, que possa ser purificada em grandes quantidades para realização de estudos funcionais (Stevens, 2000; Malhotra, 2009; Jia *et al.*, 2016). As linhagens de *E. coli* testadas nesse trabalho foram as BL21(DE3), BL21(DE3)pLysS e Rosetta2(DE3)pLysS, que são capazes de expressar proteínas recombinantes utilizando o promotor T7 presente no vetor pET100/D-TOPO. Elas possuem uma cópia do gene da T7 RNA polimerase sob o controle do promotor *lacUV5* em seu cromossomo, promotor induzível por IPTG (Isopropyl-1-thio- β -D-galactopyranoside) (Manual; Clos e Brandau, 1994). O plasmídeo pLysS contém o gene que codifica a lisozima T7. Baixos níveis de expressão dessa lisozima, que é um inibidor da T7 RNA polimerase, resulta na redução da expressão basal do gene alvo antes da adição de IPTG ao meio de cultura (Rosano e Ceccarelli, 2014). A linhagem Rosetta2(DE3)pLysS é derivada da linhagem BL21(DE3) com o intuito de melhorar a expressão de proteínas que contêm códons que são raramente usados em *E. coli*, por meio da expressão de tRNAs para os códons AUA, AGG, AGA, CUA, CCC e GGA, que codificam para os aminoácidos isoleucina, arginina, arginina, leucina, prolina e glicina, respectivamente (Alves e Souza, 2013; Jia, B. e Jeon, C. O., 2016).

Um fator que pode influenciar consideravelmente a produção de proteínas recombinantes bem como as quantidades acumuladas e a estabilidade das linhagens de expressão, são as diferentes combinações de vetor e célula hospedeira aplicadas. Nessas combinações o estresse relativo da proteína de interesse para a célula hospedeira, os níveis basais e induzidos de RNA polimerase T7 na linhagem de expressão, as características do vetor de expressão que afetam os níveis basais e induzidos do RNAm alvo, a estabilidade tanto do RNAm quanto da proteína de interesse dentro da célula, entre outros, são fatores que devem ser considerados (Studier, 2018).

Por esses motivos, várias condições foram testadas nesse trabalho a fim de expressar a RnTrypsin e o RnKazalPI, variando-se as linhagens de expressão para superar o viés de códons (mesmo tendo otimizado as sequências de códons nos cDNAs da RnTrypsin e do RnKazalPI para melhor refletirem o uso de códons da bactéria); a concentração do indutor; a temperatura de crescimento das linhagens, pois algumas proteínas que são insolúveis quando produzidas a 37 °C são consideravelmente mais solúveis quando produzidas em temperaturas mais baixas, como à temperatura ambiente ou abaixo de 18°C, devido ao fato das culturas demorarem mais tempo para atingirem a autoindução, a saturação (Studier, 2018), e o tempo de cultivo.

Mesmo assim, a RnTrypsin foi expressa em corpos de inclusão, o que pode ser devido à formação incorreta de pontes dissulfeto ou ao enovelamento incorreto da proteína. Outras possibilidades para solucionar esse problema, não testadas nessa dissertação, seriam: 1) realizar a co-expressão da RnTrypsin com chaperonas ou suplementar o meio com chaperonas químicas e co-fatores; 2) mudar de sistema de expressão, em decorrência da necessidade de modificação pós-traducional (o que pode ter impedido a obtenção da proteína de interesse), já que se trata da expressão de proteína de organismo eucariótico em sistema procariótico. Um exemplo de sistema de expressão de proteínas recombinantes é com células de inseto e baculovírus.

O objetivo da produção de anticorpos anti-RnTrypsin era de avaliar a reatividade cruzada com serino proteases de outras espécies de triatomíneos. Como a RnTrypsin recombinante se ligou fracamente à matriz na cromatografia de afinidade e começou a ser eluída com uma baixa concentração de imidazol (5mM), a amostra tinha uma alta concentração de impurezas. Mesmo após a verificação, pela técnica de Western blot, de uma banda com peso molecular de aproximadamente 30 kDa na fração eluída utilizada para fazer as imunizações nos camundongos, a RnTrypsin recombinante não foi suficientemente imunogênica para induzir a produção de anticorpos pelos animais imunizados, o que impossibilitou a análise por

Western blot de seu padrão de expressão na saliva de *R. neglectus* e de outras espécies de triatomíneos.

7. Conclusão

Nesse estudo, realizou-se uma caracterização *in silico* da serino protease do tipo tripsina (RnTrypsin; UniProt: A0A0P4VQ38) e do inibidor de serino protease tipo Kazal (RnKazalPI; UniProt: A0A0P4W3X1) salivares de *R. neglectus*. Ambas as proteínas foram expressas em sistema heterólogo utilizando *E. coli* como célula hospedeira, porém a RnTrypsin recombinante só foi observada na fração insolúvel do lisado bacteriano. Mesmo após sua purificação, não foi suficientemente imunogênica para induzir a produção de anticorpos pelos animais imunizados, o que impossibilitou a análise por *Western blot* de seu padrão de expressão na saliva de *R. neglectus*.

8. Perspectivas

1. Utilizar o sistema de expressão de proteínas recombinantes em células de inseto / baculovírus para expressar a RnTrypsin;
2. Purificar e dar continuidade à caracterização do RnKazalPI.

9. Referências bibliográficas

ALEXANDRE, J. P. D. O.; TESTON, A. P. M.; JÚNIOR, G. Z. J. R. U. R. Tratamento Etiológico da Doença de Chagas: um antigo problema de saúde pública. v. 20, n. 2, 2018. ISSN 2178-2571.

ALVES, E. A.; SOUZA, D. S. Biología molecular. In: (Ed.): EPSJV, 2013. ISBN 8598768413.

AMINO, R.; TANAKA, A. S.; SCHENKMAN, S. J. I. B. A. M. B. Triapsin, an unusual activatable serine protease from the saliva of the hematophagous vector of Chagas' disease *Triatoma infestans* (Hemiptera: Reduviidae). v. 31, n. 4-5, p. 465-472, 2001. ISSN 0965-1748.

ANDERSEN, J. F. et al. Nitric oxide binding and crystallization of recombinant nitrophorin I, a nitric oxide transport protein from the blood-sucking bug *Rhodnius prolixus*. v. 36, n. 15, p. 4423-4428, 1997. ISSN 0006-2960.

_____. Inhibition of hemostasis by a high affinity biogenic amine-binding protein from the saliva of a blood-feeding insect. v. 278, n. 7, p. 4611-4617, 2003. ISSN 0021-9258.

ANDRADE, B. B. et al. Haematophagous arthropod saliva and host defense system: a tale of tear and blood. v. 77, n. 4, p. 665-693, 2005. ISSN 0001-3765.

ARAUJO, R. N. et al. Brasiliensin: A novel intestinal thrombin inhibitor from *Triatoma brasiliensis* (Hemiptera: Reduviidae) with an important role in blood intake. v. 37, n. 12, p. 1351-1358, 2007. ISSN 0020-7519.

_____. The role of salivary nitrophorins in the ingestion of blood by the triatomine bug *Rhodnius prolixus* (Reduviidae: Triatominae). v. 39, n. 2, p. 83-89, 2009. ISSN 0965-1748.

ARAUJO-JORGE, T.; TELLERIA, J.; DALENZ, J. History of the discovery of the American Trypanosomiasis (Chagas disease). In: (Ed.). **American Trypanosomiasis Chagas Disease (Second Edition)**: Elsevier, 2017. p.1-22.

ARCÀ, B.; RIBEIRO, J. M. J. C. O. I. I. S. Saliva of hematophagous insects: a multifaceted toolkit. 2018. ISSN 2214-5745.

ARIËNS, R. A. et al. Role of factor XIII in fibrin clot formation and effects of genetic polymorphisms. **Blood**, v. 100, n. 3, p. 743-54, Aug 2002. ISSN 0006-4971. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/12130481> >.

ASSUMPÇÃO, T. C. et al. Salivary antigen-5/CAP family members are Cu²⁺-dependent antioxidant enzymes that scavenge O₂^{•-} and inhibit collagen-

induced platelet aggregation and neutrophil oxidative burst. **J Biol Chem**, v. 288, n. 20, p. 14341-61, May 2013. ISSN 1083-351X. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/23564450> >.

ASSUMPÇÃO, T. C. F. et al. Insight into the Salivary the Transcriptome and Proteome of *Dipetalogaster maxima*. **Journal of Proteome research**, p. 669-679, 2010.

_____. An insight into the sialome of the blood-sucking bug *Triatoma infestans* a vector of Chagas' disease. **Insect Biochemistry and Molecular Biology**, v. 38, p. 213-232, 2008.

BAE, J. S.; KIM, I. S.; REZAIE, A. R. J. J. O. C. P. Thrombin down-regulates the TGF- β -mediated synthesis of collagen and fibronectin by human proximal tubule epithelial cells through the EPCR-dependent activation of PAR-1. v. 225, n. 1, p. 233-239, 2010. ISSN 0021-9541.

BARGUES, M. D.; SCHOFIELD, C.; DUJARDIN, J.-P. Classification and phylogeny of the Triatominae. In: (Ed.). **American Trypanosomiasis**: Elsevier, 2010. p.117-147.

BARRAL, A. et al. Human immune response to sand fly salivary gland antigens: a useful epidemiological marker? , v. 62, n. 6, p. 740-745, 2000. ISSN 0002-9637.

BARRETO-SANTANA, D. et al. Biología comparativa e comportamento alimentar de *Rhodnius neglectus* e *Rhodnius robustus* (Triatominae) sob condicoes de laboratorio. v. 44, n. 4, p. 490-495, 2011. ISSN 0037-8682.

BEER, L. A.; SPEICHER, D. W. Protein detection in gels using fixation. **Current protocols in protein science**, v. 91, n. 1, p. 10.5. 1-10.5. 20, 2018. ISSN 1934-3655.

BEGON, M. Effects of host diversity on disease dynamics. **Infectious disease ecology: effects of ecosystems on disease and of disease on ecosystems**, p. 12-29, 2008.

BENDTSEN, J. D. et al. Improved prediction of signal peptides: SignalP 3.0. **Journal of molecular biology**, v. 340, n. 4, p. 783-795, 2004. ISSN 0022-2836.

BERGER, M. et al. Hemostasia: uma breve revisão. v. 11, n. 1, 2014. ISSN 1983-0882.

BERGKESSEL, M.; GUTHRIE, C. Colony PCR. In: (Ed.). **Methods in enzymology**: Elsevier, v.529, 2013. p. 299-309. ISBN 0076-6879.

BERN, C. et al. *Trypanosoma cruzi* and Chagas' disease in the United States. **Clinical microbiology reviews**, v. 24, n. 4, p. 655-681, 2011. ISSN 0893-8512.

- BLOM, N.; GAMMELTOFT, S.; BRUNAK, S. J. J. O. M. B. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites1. v. 294, n. 5, p. 1351-1362, 1999. ISSN 0022-2836.
- BOND, J. S.; BUTLER, P. E. J. A. R. O. B. Intracellular proteases. v. 56, n. 1, p. 333-364, 1987. ISSN 0066-4154.
- BRENER, Z. Comparative studies of different strains of *Trypanosoma cruzi*. **Annals of Tropical Medicine & Parasitology**, v. 59, n. 1, p. 19-26, 1965.
- BRENER, Z. J. R. D. I. D. M. T. D. S. P. Life cycle of *Trypanosoma cruzi*. v. 13, n. 3, p. 171-8, 1971. ISSN 0036-4665.
- BRICEÑO-LEÓN, R.; MÉNDEZ GALVÁN, J. The social determinants of Chagas disease and the transformations of Latin America. **Memórias do Instituto Oswaldo Cruz**, v. 102, p. 109-112, 2007. ISSN 0074-0276.
- BRUMMER-KORVENKONTIO, H. et al. Detection of mosquito saliva-specific IgE and IgG4 antibodies by immunoblotting. v. 93, n. 3, p. 551-555, 1994. ISSN 0091-6749.
- BUSSACOS, A. C. et al. Redundancy of proteins in the salivary glands of *Panstrongylus megistus* secures prolonged procurement for blood meals. v. 74, n. 9, p. 1693-1700, 2011. ISSN 1874-3919.
- _____. Diversity of anti-haemostatic proteins in the salivary glands of *Rhodnius* species transmitters of Chagas disease in the greater Amazon. v. 74, n. 9, p. 1664-1672, 2011. ISSN 1874-3919.
- CAMPOS, I. et al. Infestin, a thrombin inhibitor presents in *Triatoma infestans* midgut, a Chagas' disease vector: gene cloning, expression and characterization of the inhibitor. v. 32, n. 9, p. 991-997, 2002. ISSN 0965-1748.
- CARCAVALLO, R. U. et al. Atlas dos vetores da doença de Chagas nas Américas. In: (Ed.). **Atlas dos vetores da doença de Chagas nas Américas**, 1997. p.1217-1217.
- CAVALIER-SMITH, T. Higher classification and phylogeny of Euglenozoa. **Eur J Protistol**, v. 56, p. 250-276, Oct 2016. ISSN 1618-0429. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/27889663> >.
- CDC (Centers for Disease Control & Prevention). Division of Parasitic Disease, Laboratory Identification of Parasites of Public Health Concern, National Center for Infectious Diseases. Parasites - American Trypanosomiasis (also known as Chagas Disease). Disponível em <<https://www.cdc.gov/parasites/chagas/biology.html>>. Acesso em 14 agosto 2018.

CHAGAS, C. Nova tripanozomíaze humana: estudos sobre a morfologia e o ciclo evolutivo do *Schizotrypanum cruzi* n. gen., n. sp., agente etiológico de nova entidade morbida do homem. **Memórias do Instituto Oswaldo Cruz**, v. 1, n. 2, p. 159-218, 1909. ISSN 0074-0276.

CHAMPAGNE, D. The role of salivary vasodilators in bloodfeeding and parasite transmission. **Parasitology Today**, v. 10, n. 11, p. 430-433, 1994. ISSN 0169-4758.

CHAMPAGNE, D. E. Antihemostatic strategies of blood-feeding arthropods. **Curr Drug Targets Cardiovasc Haematol Disord**, v. 4, n. 4, p. 375-96, Dec 2004. ISSN 1568-0061. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15578959> >.

_____. Antihemostatic molecules from saliva of blood-feeding arthropods. **Pathophysiol Haemost Thromb**, v. 34, n. 4-5, p. 221-7, 2005. ISSN 1424-8832. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/16707932> >.

CHAMPAGNE, D. E.; NUSSENZVEIG, R. H.; RIBEIRO, J. M. J. J. O. B. C. Purification, partial characterization, and cloning of nitric oxide-carrying heme proteins (nitrophorins) from salivary glands of the blood-sucking insect *Rhodnius prolixus*. v. 270, n. 15, p. 8691-8695, 1995. ISSN 0021-9258.

CHAMPAGNE, D. E. J. C. D. T.-C.; DISORDERS, H. Antihemostatic strategies of blood-feeding arthropods. v. 4, n. 4, p. 375-396, 2004. ISSN 1568-0061.

CHARNEAU, S.; JUNQUEIRA, M.; COSTA, C. M. **The saliva proteome of the blood-feeding insect *Triatoma infestans* is rich in platelet-aggregation inhibitors**. International Journal of Mass Spectrometry. v. 268, p. 265–276 p. 2007.

CHATELAIN, E. Chagas disease drug discovery: toward a new era. **Journal of biomolecular screening**, v. 20, n. 1, p. 22-35, 2015. ISSN 1087-0571.

CIPRANDI, A.; HORN, F.; TERMIGNONI, C. J. R. B. D. H. E. H. S. P., SP. VOL. 25, N. 4 , P. 250-262. Saliva de animais hematófagos: fonte de novos anticoagulantes. 2003. ISSN 1516-8484.

CLOS, J.; BRANDAU, S. J. P. E. A. P. pJC20 and pJC40-two high-copy-number vectors for T7 RNA polymerase-dependent expression of recombinant genes in *Escherichia coli*. v. 5, n. 2, p. 133-137, 1994. ISSN 1046-5928.

CORPET, F. J. N. A. R. Multiple sequence alignment with hierarchical clustering. v. 16, n. 22, p. 10881-10890, 1988. ISSN 1362-4962.

COSTA, S. et al. Fusion tags for protein solubility, purification and immunogenicity in *Escherichia coli*: the novel Fh8 system. **Frontiers in microbiology**, v. 5, p. 63, 2014. ISSN 1664-302X.

COUGHLIN, S. R. J. J. O. T. A. H. Protease-activated receptors in hemostasis, thrombosis and vascular biology. v. 3, n. 8, p. 1800-1814, 2005. ISSN 1538-7933.

COURA, J. et al. Uma visão sistêmica da endemia chagásica. **La enfermedad de Chagas a la puerta de los**, v. 100, p. 25-35, 2007.

COURA, J. R. Chagas disease: what is known and what is needed-A background article. **Memórias do Instituto Oswaldo Cruz**, v. 102, p. 113-122, 2007. ISSN 0074-0276.

COURA, J. R.; BORGES-PEREIRA, J. Chagas disease: 100 years after its discovery. A systemic review. **Acta tropica**, v. 115, n. 1-2, p. 5-13, 2010. ISSN 0001-706X.

CRAWLEY, J. et al. The central role of thrombin in hemostasis. v. 5, p. 95-101, 2007. ISSN 1538-7933.

DA ROSA, J. A. et al. A new species of *Rhodnius* from Brazil (Hemiptera, Reduviidae, Triatominae). **Zookeys**, n. 675, p. 1-25, 2017. ISSN 1313-2989. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/28769676> >.

DAUMERIE, D.; SAVIOLI, L. **Working to overcome the global impact of neglected tropical diseases: first WHO report on neglected tropical diseases**. World Health Organization, 2010. ISBN 9241564091.

DE ANDRADE, A. L. S. S. et al. Randomised trial of efficacy of benznidazole in treatment of early *Trypanosoma cruzi* infection. **The Lancet**, v. 348, n. 9039, p. 1407-1413, 1996. ISSN 0140-6736.

DE ARAÚJO, C. N. et al. Interactome: Smart hematophagous triatomine salivary gland molecules counteract human hemostasis during meal acquisition. **J Proteomics**, v. 75, n. 13, p. 3829-41, Jul 2012. ISSN 1876-7737. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/22579750> >.

DE OLIVEIRA, J. et al. Study of the Salivary Glands in Triatominae (Hemiptera, Reduviidae, Triatominae): Their Color and Application to the Chagas Disease Vector Evolution. v. 97, n. 3, p. 771-773, 2017. ISSN 0002-9637.

DEANE, M. P.; LENZI, H. L.; JANSEN, A. *Trypanosoma cruzi*: vertebrate and invertebrate cycles in the same mammal host, the opossum *Didelphis marsupialis*. **Memórias do Instituto Oswaldo Cruz**, v. 79, n. 4, p. 513-515, 1984. ISSN 0074-0276.

DIAS, J. C. P. et al. II Consenso Brasileiro em doença de Chagas, 2015. v. 25, p. 7-86, 2016. ISSN 2237-9622.

DISEASE, W. E. C. O. T. C. O. C.; ORGANIZATION, W. H. **Control of Chagas disease: Second report of the WHO Expert Committee.** World Health Organization, 2002. ISBN 9241209054.

DONPUDSA, S. et al. Domain inhibitory and bacteriostatic activities of the five-domain Kazal-type serine proteinase inhibitor from black tiger shrimp *Penaeus monodon*. v. 33, n. 4, p. 481-488, 2009. ISSN 0145-305X.

DURAN-REHBEIN, G. A. et al. Mammalian cellular culture models of *Trypanosoma cruzi* infection: a review of the published literature. v. 21, 2014.

EMANUELSSON, O. et al. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. v. 300, n. 4, p. 1005-1016, 2000. ISSN 0022-2836.

ESTANI, S. S. et al. Efficacy of chemotherapy with benznidazole in children in the indeterminate phase of Chagas' disease. **The American journal of tropical medicine and hygiene**, v. 59, n. 4, p. 526-529, 1998. ISSN 0002-9637.

FAUDRY, E. et al. *Triatoma infestans* apyrases belong to the 5'-nucleotidase family. **Journal of Biological Chemistry**, v. 279, n. 19, p. 19607-19613, 2004. ISSN 0021-9258.

FEINGOLD, B. F.; BENJAMINI, E. J. A. O. A. Allergy to flea bites. Clinical and experimental observations. v. 19, p. 1275, 1961. ISSN 0003-4738.

FERREIRA, R. A. et al. Do haematophagous bugs assess skin surface temperature to detect blood vessels? **PLoS One**, v. 2, n. 9, p. e932, 2007. ISSN 1932-6203.

FILÉE, J. et al. Mariner transposons are sailing in the genome of the blood-sucking bug *Rhodnius prolixus*. v. 16, n. 1, p. 1061, 2015. ISSN 1471-2164.

FLORES, G. B.; LAZZARI, C. R. The role of the antennae in *Triatoma infestans*: orientation towards thermal sources. **Journal of insect physiology**, v. 42, n. 5, p. 433-440, 1996. ISSN 0022-1910.

FRANCISCHETTI, I. M. et al. Purification, cloning, expression, and mechanism of action of a novel platelet aggregation inhibitor from the salivary gland of the blood-sucking bug, *Rhodnius prolixus*. v. 275, n. 17, p. 12639-12650, 2000. ISSN 0021-9258.

FRIEDRICH, T. et al. A Kazal-type inhibitor with thrombin specificity from *Rhodnius prolixus*. v. 268, n. 22, p. 16216-16222, 1993. ISSN 0021-9258.

FUENTES-PRIOR, P. et al. Structure of the thrombin complex with triabin, a lipocalin-like exosite-binding inhibitor derived from a triatomine bug. v. 94, n. 22, p. 11845-11850, 1997. ISSN 0027-8424.

FURIE, B.; FURIE, B. C. J. T. J. O. C. I. Thrombus formation in vivo. v. 115, n. 12, p. 3355-3362, 2005. ISSN 0021-9738.

GALVÃO, C. **Vetores da doença de chagas no Brasil**. SciELO-Sociedade Brasileira de Zoologia, 2014. ISBN 8598203092.

GARCIA, E. S. et al. Exploring the role of insect host factors in the dynamics of Trypanosoma cruzi-Rhodnius prolixus interactions. **J Insect Physiol**, v. 53, n. 1, p. 11-21, Jan 2007. ISSN 0022-1910. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/17141801> >.

GARDINER, E. E.; ANDREWS, R. K. Structure and function of platelet receptors initiating blood clotting. In: (Ed.). **A Systems Biology Approach to Blood**: Springer, 2014. p.263-275.

GASCON, J.; BERN, C.; PINAZO, M.-J. Chagas disease in Spain, the United States and other non-endemic countries. **Acta tropica**, v. 115, n. 1-2, p. 22-27, 2010. ISSN 0001-706X.

GAUNT, M.; MILES, M. J. M. D. I. O. C. The ecotopes and evolution of triatomine bugs (Triatominae) and their associated trypanosomes. v. 95, n. 4, p. 557-565, 2000. ISSN 0074-0276.

GLADWELL, T. D. J. C. T. Bivalirudin: a direct thrombin inhibitor. v. 24, n. 1, p. 38-58, 2002. ISSN 0149-2918.

GOMES, R. et al. Immunity to a salivary protein of a sand fly vector protects against the fatal outcome of visceral leishmaniasis in a hamster model. v. 105, n. 22, p. 7845-7850, 2008. ISSN 0027-8424.

GONÇALVES, C. S. et al. Revisiting the Trypanosoma cruzi metacyclogenesis: morphological and ultrastructural analyses during cell differentiation. v. 11, n. 1, p. 83, 2018. ISSN 1756-3305.

GOURBIÈRE, S. et al. Genetics and evolution of triatomines: from phylogeny to vector control. **Heredity (Edinb)**, v. 108, n. 3, p. 190-202, Mar 2012. ISSN 1365-2540. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/21897436> >.

GUBLER, D. J. Resurgent vector-borne diseases as a global health problem. **Emerging infectious diseases**, v. 4, n. 3, p. 442, 1998.

GURGEL-GONÇALVES, Rodrigo et al. Is Rhodnius prolixus (Triatominae) invading houses in central Brazil?. **Acta Tropica**, v. 107, n. 2, p. 90-98, 2008.

GURGEL-GONÇALVES, Rodrigo et al. Infestation of Mauritia flexuosa palms by triatomines (Hemiptera: Reduviidae), vectors of Trypanosoma cruzi and Trypanosoma rangeli in the Brazilian savanna. **Acta tropica**, v. 121, n. 2, p. 105-111, 2012.

GÜRTLER, R. E.; CARDINAL, M. J. A. T. Reservoir host competence and the role of domestic and commensal hosts in the transmission of *Trypanosoma cruzi*. v. 151, p. 32-50, 2015. ISSN 0001-706X.

HAYCRAFT, J. B. J. A. F. E. P. U. P. Ueber die Einwirkung eines Secretes des officinellen Blutegels auf die Gerinnbarkeit des Blutes. v. 18, n. 3-4, p. 209-217, 1884. ISSN 0365-2009.

HE, W. Y. et al. Cloning, expression and localization of a trypsin-like serine protease in the spruce budworm, *Choristoneura fumiferana*. v. 16, n. 6, p. 455-464, 2009. ISSN 1672-9609.

HECHT, M. M. et al. *Triatoma infestans* chooses to feed upon immune prey. **The American journal of tropical medicine and hygiene**, v. 75, n. 5, p. 893-900, 2006. ISSN 0002-9637.

HELLMANN, K.; HAWKINS, R. I. J. N. Prolixin-S and Prolixin-G; two anticoagulants from *Rhodnius prolixus* Stål. v. 207, n. 4994, p. 265, 1965. ISSN 1476-4687.

HOARE, C. A. The trypanosomes of mammals. A Zoological Monograph. **The trypanosomes of mammals. A zoological monograph.**, 1972.

HOFFMAN, R. et al. **Hematology: basic principles and practice**. Elsevier Health Sciences, 2013. ISBN 1437729282.

HOTEZ, P. J. et al. The neglected tropical diseases of Latin America and the Caribbean: a review of disease burden and distribution and a roadmap for control and elimination. **PLoS Negl Trop Dis**, v. 2, n. 9, p. e300, 2008. ISSN 1935-2735. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/18820747> >.

JACKSON, Y.; PINTO, A.; PETT, S. Chagas disease in Australia and New Zealand: risks and needs for public health interventions. **Tropical Medicine & International Health**, v. 19, n. 2, p. 212-218, 2014. ISSN 1360-2276.

JENNE, C. N.; KUBES, P. J. P. Platelets in inflammation and infection. v. 26, n. 4, p. 286-292, 2015. ISSN 0953-7104.

JENNINGS, L. K. J. T. A. J. O. C. Role of platelets in atherothrombosis. v. 103, n. 3, p. 4A-10A, 2009. ISSN 0002-9149.

JIA, B.; JEON, C. O. High-throughput recombinant protein expression in *Escherichia coli*: current status and future perspectives. **Open biology**, v. 6, n. 8, p. 160196, 2016. ISSN 2046-2441.

JIA, B.; JEON, C. O. J. O. B. High-throughput recombinant protein expression in *Escherichia coli*: current status and future perspectives. v. 6, n. 8, p. 160196, 2016. ISSN 2046-2441.

- JIANG, H.; KANOST, M. R. J. I. B. A. M. B. The clip-domain family of serine proteinases in arthropods. v. 30, n. 2, p. 95-105, 2000. ISSN 0965-1748.
- JURBERG, J. et al. **Atlas Iconográfico dos triatomíneos do Brasil (Vetores da doença de Chagas)**: Laboratorio Nacional e Internacional de Referencia em Taxonomia de Triatomíneos, Instituto Oswaldo Cruz 2014.
- KALAFATIS, M. et al. The regulation of clotting factors. v. 7, n. 3, 1997. ISSN 1045-4403.
- KALAFATIS, M.; RAND, M. D.; MANN, K. G. J. J. O. B. C. The mechanism of inactivation of human factor V and human factor Va by activated protein C. v. 269, n. 50, p. 31869-31880, 1994. ISSN 0021-9258.
- Kazal, L. A.; SPICER, D. S.; BRAHINSKY, R. A. J. J. O. T. A. C. S. Isolation of a Crystalline Trypsin Inhibitor-Anticoagulant Protein from Pancreas1a. v. 70, n. 9, p. 3034-3040, 1948. ISSN 0002-7863.
- KESSLER, R. L. et al. Recently differentiated epimastigotes from *Trypanosoma cruzi* are infective to the mammalian host. **Mol Microbiol**, v. 104, n. 5, p. 712-736, Jun 2017. ISSN 1365-2958. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/28240790> >.
- KIM, B. Y. et al. Antimicrobial activity of a honeybee (*Apis cerana*) venom Kazal-type serine protease inhibitor. v. 76, p. 110-117, 2013. ISSN 0041-0101.
- KINI, R. M. et al. Toxins and Hemostasis. v. 20011, p. 600, 2010.
- LACOMBE, D. J. M. D. I. O. C. Anatomy and histology of salivary glands of triatomine bugs. v. 94, n. 4, p. 557-564, 1999. ISSN 0074-0276.
- LAEMMLI, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **nature**, v. 227, n. 5259, p. 680, 1970. ISSN 1476-4687.
- LAZZARI, C. R.; PEREIRA, M. H.; LORENZO, M. G. Behavioural biology of Chagas disease vectors. **Memorias do Instituto Oswaldo Cruz**, v. 108, p. 34-47, 2013. ISSN 0074-0276.
- LEHANE, M. J. **The biology of blood-sucking in insects**. Cambridge University Press, 2005. ISBN 0521836085.
- LENT, H.; WYGODZINSKY, P. Revision of the Triatominae (Hemiptera, Reduviidae), and their significance as vectors of Chagas' disease. **Bulletin of the American museum of Natural History**, v. 163, n. 3, p. 123-520, 1979.
- LERNER, E. et al. Isolation of maxadilan, a potent vasodilatory peptide from the salivary glands of the sand fly *Lutzomyia longipalpis*. v. 266, n. 17, p. 11234-11236, 1991. ISSN 0021-9258.

LERNER, E.; SHOEMAKER, C. J. J. O. B. C. Maxadilan. Cloning and functional expression of the gene encoding this potent vasodilator peptide. v. 267, n. 2, p. 1062-1066, 1992. ISSN 0021-9258.

LOVATO, D. V. et al. Infestin 1R, an intestinal subtilisin inhibitor from *Triatoma infestans* able to impair mammalian cell invasion by *Trypanosoma cruzi*. v. 129, n. 4, p. 362-367, 2011. ISSN 0014-4894.

MACKMAN, N.; TILLEY, R. E.; KEY, N. S. J. A., THROMBOSIS, AND VASCULAR BIOLOGY. Role of the extrinsic pathway of blood coagulation in hemostasis and thrombosis. v. 27, n. 8, p. 1687-1693, 2007. ISSN 1079-5642.

MALHOTRA, A. Tagging for protein expression. In: (Ed.). **Methods in enzymology**: Elsevier, v.463, 2009. p. 239-258. ISBN 0076-6879.

MANUAL, I. BL21(DE3) Competent Cells, BL21(DE3)pLysS Competent Cells, and BL21 Competent Cells.

MARAGANORE, J. et al. Design and characterization of hirulogs: a novel class of bivalent peptide inhibitors of thrombin. v. 29, n. 30, p. 7095-7101, 1990. ISSN 0006-2960.

MARDER, V. J. et al. **Hemostasis and thrombosis: basic principles and clinical practice**. Lippincott Williams & Wilkins, 2012. ISBN 1451177690.

MARKWARDT, F. et al. Isolation of hiruin from the medicinal leech. v. 22, n. 6, p. 239, 1967. ISSN 0031-7144.

MAZZA, S.; FREIRE, R. Manifestaciones cutáneas en Enfermedad de Chagas. **Chagoma de Inoculación, Chagomas Metastásicos y Chagomas Hematógenos**. *Rev. MEPPRA*, n. 46, p. 3-38, 1940.

MEJIA, J. S.; BISHOP, J. V.; TITUS, R. G. J. T. I. P. Is it possible to develop pan-arthropod vaccines? , v. 22, n. 8, p. 367-370, 2006. ISSN 1471-4922.

MENDE, K.; LANGE, U.; NOWAK, G. J. I. B. A. M. B. Three recombinant serine proteinase inhibitors expressed from the coding region of the thrombin inhibitor dipetalogastin. v. 34, n. 9, p. 971-979, 2004. ISSN 0965-1748.

MINUZZI-SOUZA, T. T. et al. Vector-borne transmission of *Trypanosoma cruzi* among captive Neotropical primates in a Brazilian zoo. **Parasit Vectors**, v. 9, p. 39, Jan 2016. ISSN 1756-3305. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/26813657> >.

MONTEIRO, F. A. et al. Phylogeny and molecular taxonomy of the *Rhodniini* derived from mitochondrial and nuclear DNA sequences. **Am J Trop Med Hyg**, v. 62, n. 4, p. 460-5, Apr 2000. ISSN 0002-9637. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/11220761> >.

MONTFORT, W. R.; WEICHSEL, A.; ANDERSEN, J. F. J. B. E. B. A.-P. S. A. M. E. Nitrophorins and related antihemostatic lipocalins from *Rhodnius prolixus* and other blood-sucking arthropods. v. 1482, n. 1-2, p. 110-118, 2000. ISSN 0167-4838.

MORITA, A. et al. Identification and characterization of a collagen-induced platelet aggregation inhibitor, triplatin, from salivary glands of the assassin bug, *Triatoma infestans*. v. 273, n. 13, p. 2955-2962, 2006. ISSN 1742-464X.

NAVARRO, M. et al. Chagas disease in Spain: need for further public health measures. **PLoS neglected tropical diseases**, v. 6, n. 12, p. e 1962, 2012. ISSN 1935-2735.

NIELSEN, H. et al. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. v. 10, n. 1, p. 1-6, 1997. ISSN 1741-0134.

NIESWANDT, B.; WATSON, S. P. J. B. Platelet-collagen interaction: is GPVI the central receptor? , v. 102, n. 2, p. 449-461, 2003. ISSN 0006-4971.

NOESKE-JUNGBLUT, C. et al. Triabin, a highly potent exosite inhibitor of thrombin. v. 270, n. 48, p. 28629-28634, 1995. ISSN 0021-9258.

_____. An inhibitor of collagen-induced platelet aggregation from the saliva of *Triatoma pallidipennis*. v. 269, n. 7, p. 5050-5053, 1994. ISSN 0021-9258.

OLIVEIRA, F. et al. From transcriptome to immunome: identification of DTH inducing proteins from a *Phlebotomus ariasi* salivary gland cDNA library. v. 24, n. 3, p. 374-390, 2006. ISSN 0264-410X.

OLIVEIRA, A. W. S. de; SILVA, I. G. da. Distribuição geográfica e indicadores entomológicos de triatomíneos sinantrópicos capturados no Estado de Goiás. **Rev Soc Bras Med Trop**, v. 40, n. 2, p. 204-8, 2007.

OTÁLORA-LUNA, F. et al. Evolution of hematophagous habit in Triatominae (Heteroptera: Reduviidae). **Revista chilena de historia natural**, v. 88, n. 1, p. 4, 2015. ISSN 0717-6317.

PACHECO, D. E. Nitroforinas salivares de *Rhodnius prolixus* (Hemiptera: Reduviidae): avaliação do RNAi parental e do papel da albumina do hospedeiro nasua atividade biológica. 2014.

RABINOVICH, J. E. et al. Ecological patterns of blood-feeding by kissing-bugs (Hemiptera: Reduviidae: Triatominae). v. 106, n. 4, p. 479-494, 2011. ISSN 0074-0276.

RASSI JR, A. et al. Development and validation of a risk score for predicting death in Chagas' heart disease. **New England Journal of Medicine**, v. 355, n. 8, p. 799-808, 2006. ISSN 0028-4793.

RASSI JR, A.; RASSI, A.; MARIN-NETO, J. A. Chagas disease. **The Lancet**, v. 375, n. 9723, p. 1388-1402, 2010. ISSN 0140-6736.

RAVAZI, A. et al. Cytogenetic analysis in different populations of *Rhodnius prolixus* and *R. nasutus* from different countries of South America. **Braz J Biol**, p. 0, May 2017. ISSN 1678-4375. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/28562785> >.

REZENDE, S. M. Distúrbios da hemostasia: doenças hemorrágicas. 2010.

RIBEIRO, J. Blood-feeding arthropods: live syringes or invertebrate pharmacologists? **Infectious agents and disease**, v. 4, n. 3, p. 143-152, 1995. ISSN 1056-2044.

RIBEIRO, J.; WALKER, F. A. J. J. O. E. M. High affinity histamine-binding and antihistaminic activity of the salivary nitric oxide-carrying heme protein (nitrophorin) of *Rhodnius prolixus*. v. 180, n. 6, p. 2251-2257, 1994. ISSN 0022-1007.

RIBEIRO, J. M. **Role of saliva in blood-feeding by arthropods**: *Annu. Rev. Entomol.* 32: 463-78 p. 1987.

RIBEIRO, J. M.; ARCA, B. J. A. I. I. P. From sialomes to the sialoverse: an insight into salivary potion of blood-feeding insects. v. 37, p. 59-118, 2009. ISSN 0065-2806.

RIBEIRO, J. M.; FRANCISCHETTI, I. M. Role of arthropod saliva in blood feeding: sialome and post-sialome perspectives. **Annu Rev Entomol**, v. 48, p. 73-88, 2003. ISSN 0066-4170. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/12194906> >.

RIBEIRO, J. M.; SCHWARZ, A.; FRANCISCHETTI, I. M. A Deep Insight Into the Sialotranscriptome of the Chagas Disease Vector, *Panstrongylus megistus* (Hemiptera: Heteroptera). **J Med Entomol**, v. 52, n. 3, p. 351-8, May 2015. ISSN 0022-2585. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/26334808> >.

RODRIGUES, V. L. et al. Colonization of palm trees by *Rhodnius neglectus* and household and invasion in an urban area, Araçatuba, São Paulo State, Brazil. **Rev Inst Med Trop Sao Paulo**, v. 56, n. 3, p. 213-8, 2014 May-Jun 2014. ISSN 1678-9946. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/24878999> >.

ROQUE, A.; JANSEN, A. Reservatórios do *Trypanosoma cruzi* e sua relação com os vetores. **Vetores da doença de Chagas no Brasil**. Editora Fiocruz, Rio de Janeiro, 2010.

ROSANO, G. L.; CECCARELLI, E. A. Recombinant protein expression in *Escherichia coli*: advances and challenges. **Frontiers in microbiology**, v. 5, p. 172, 2014. ISSN 1664-302X.

SALUD, O. P. D. L. **Estimación cuantitativa de la enfermedad de Chagas en las Américas**: Organización Panamericana de la Salud Montevideo, Uruguay, 2006.

SALZET, M. J. F. L. Anticoagulants and inhibitors of platelet aggregation derived from leeches. v. 492, n. 3, p. 187-192, 2001. ISSN 0014-5793.

SAMUELS, A. M. et al. Epidemiology of and impact of insecticide spraying on Chagas disease in communities in the Bolivian Chaco. **PLoS neglected tropical diseases**, v. 7, n. 8, p. e2358, 2013. ISSN 1935-2735.

SANTIAGO, P. B. et al. A deep insight into the sialome of *Rhodnius neglectus*, a vector of Chagas disease. **PLoS neglected tropical diseases**, v. 10, n. 4, p. e0004581, 2016. ISSN 1935-2735.

SANTOS, A. et al. The sialotranscriptome of blood-sucking bug *Triatoma brasiliensis* (Hemiptera, Triatominae). **Insect Biochem. Mol. Biol.**, v. 37, p. 702-712, 2007.

SANT'ANNA, M. R. V. et al. Triatomines (Hemiptera, Reduviidae) blood intake: Physical constraints and biological adaptations. **Journal of insect physiology**, v. 97, p. 20-26, 2017. ISSN 0022-1910.

SARKIS, J.; GUIMARAES, J.; RIBEIRO, J. Salivary apyrase of *Rhodnius prolixus*. Kinetics and purification. **Biochemical Journal**, v. 233, n. 3, p. 885-891, 1986. ISSN 0264-6021.

SAVIOLI, L.; DAUMERIE, D. **Sustaining the drive to overcome the global impact of neglected tropical diseases: second WHO report on neglected tropical diseases**. World Health Organization, 2013. ISBN 9241564547.

SCHLOTT, B. et al. Interaction of Kazal-type inhibitor domains with serine proteinases: biochemical and structural studies. v. 318, n. 2, p. 533-546, 2002. ISSN 0022-2836.

SCHWARTZ, B. S.; RIBEIRO, J. M.; GOLDSTEIN, M. D. J. A. J. O. E. Anti-tick antibodies: an epidemiologic tool in Lyme disease research. v. 132, n. 1, p. 58-66, 1990. ISSN 1476-6256.

SOARES, A. C. et al. Salivation pattern of *Rhodnius prolixus* (Reduviidae; Triatominae) in mouse skin. v. 52, n. 5, p. 468-472, 2006. ISSN 0022-1910.

SOARES, T. S. et al. A Kazal-type inhibitor is modulated by *Trypanosoma cruzi* to control microbiota inside the anterior midgut of *Rhodnius prolixus*. **Biochimie**, v. 112, p. 41-48, 2015. ISSN 0300-9084.

SPRONK, H. M.; BORISSOFF, J. I.; TEN CATE, H. J. C. A. R. New insights into modulation of thrombin formation. v. 15, n. 11, p. 363, 2013. ISSN 1523-3804.

STARK, K. R.; JAMES, A. A. Salivary gland anticoagulants in culicine and anopheline mosquitoes (Diptera: Culicidae). **Journal of medical entomology**, v. 33, n. 4, p. 645-650, 1996. ISSN 1938-2928.

STEENTOFT, C. et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. v. 32, n. 10, p. 1478-1488, 2013. ISSN 0261-4189.

STEVENS, J. et al. The molecular evolution of Trypanosomatidae. **Advances in parasitology**, v. 48, p. 2-56, 2001. ISSN 0065-308X.

STEVENS, L. et al. Kissing bugs. The vectors of Chagas. In: (Ed.). **Advances in parasitology**: Elsevier, v.75, 2011. p.169-192. ISBN 0065-308X.

STEVENS, R. C. Design of high-throughput methods of protein production for structural biology. **Structure**, v. 8, n. 9, p. R177-R185, 2000. ISSN 0969-2126.

STUDIER, F. W. T7 Expression Systems for Inducible Production of Proteins from Cloned Genes in E. coli. **Current protocols in molecular biology**, p. e63, 2018. ISSN 1934-3639.

SWITZER III, R. C.; MERRIL, C. R.; SHIFRIN, S. A highly sensitive silver stain for detecting proteins and peptides in polyacrylamide gels. **Analytical biochemistry**, v. 98, n. 1, p. 231-237, 1979. ISSN 0003-2697.

TABOR, S.; RICHARDSON, C. C. J. P. O. T. N. A. O. S. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. v. 82, n. 4, p. 1074-1078, 1985. ISSN 0027-8424.

TEIXEIRA, A. R. et al. Pathogenesis of chagas' disease: parasite persistence and autoimmunity. **Clinical microbiology reviews**, v. 24, n. 3, p. 592-630, 2011. ISSN 0893-8512.

TITUS, R.; BISHOP, J.; MEJIA, J. J. P. I. The immunomodulatory factors of arthropod saliva and the potential for these factors to serve as vaccine targets to prevent pathogen transmission. v. 28, n. 4, p. 131-141, 2006. ISSN 0141-9838.

TITUS, R. G.; RIBEIRO, J. Salivary gland lysates from the sand fly *Lutzomyia longipalpis* enhance *Leishmania* infectivity. **Science**, v. 239, n. 4845, p. 1306-1308, 1988. ISSN 0036-8075.

TOMLINSON, S. et al. The induction of *Trypanosoma cruzi* trypomastigote to amastigote transformation by low pH. **Parasitology**, v. 110, n. 5, p. 547-554, 1995. ISSN 1469-8161.

TREMACOLDI, C. R. Proteases e Inibidores de Proteases na Defesa de

Plantas Contra Pragas. Embrapa Amazônia Oriental Documentos 353. 1 ed. Belém: Embrapa, 2009, 45 p.

VALENZUELA, J. G. et al. Identification of the most abundant secreted proteins from the salivary glands of the sand fly *Lutzomyia longipalpis*, vector of *Leishmania chagasi*. v. 207, n. 21, p. 3717-3729, 2004. ISSN 0022-0949.

VAN DE LOCHT, A. et al. Two heads are better than one: crystal structure of the insect derived double domain Kazal inhibitor rhodniin in complex with thrombin. v. 14, n. 21, p. 5149-5157, 1995. ISSN 0261-4189.

VARGA-SZABO, D.; PLEINES, I.; NIESWANDT, B. J. A., THROMBOSIS, AND VASCULAR BIOLOGY. Cell adhesion mechanisms in platelets. v. 28, n. 3, p. 403-412, 2008. ISSN 1079-5642.

VERSTEEG, H. H. et al. New fundamentals in hemostasis. **Physiol Rev**, v. 93, n. 1, p. 327-58, Jan 2013. ISSN 1522-1210. Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/23303912> >.

WALKER, A. A. et al. Melt with this kiss: Paralyzing and liquefying venom of the assassin bug *Pristhesancus plagipennis* (Hemiptera: Reduviidae). p. mcp. M116. 063321, 2017. ISSN 1535-9476.

WATSON, S. et al. GPVI and integrin $\alpha\text{IIb}\beta\text{3}$ signaling in platelets. v. 3, n. 8, p. 1752-1762, 2005. ISSN 1538-7933.

WEICHSEL, A. et al. Crystal structures of a nitric oxide transport protein from a blood-sucking insect. v. 5, n. 4, p. 304, 1998. ISSN 1545-9985.

WHO. **Control of Chagas disease: second report of the WHO expert committee**. World Health Organization, 2002. ISBN 9241209054.

WHO. Chagas disease in Latin America: an epidemiological update based on 2010 estimates. **Weekly Epidemiological Record= Relevé épidémiologique hebdomadaire**, v. 90, n. 06, p. 33-44, 2015.

WHO. **Chagas disease (American trypanosomiasis). Fact sheet N 340 (2015)**, 2016.

WIGGLESWORTH, V.; GILLET, J. The function of the antennae in *Rhodnius prolixus* (Hemiptera) and the mechanism of orientation to the host. **Journal of Experimental Biology**, v. 11, n. 2, p. 120-139, 1934. ISSN 0022-0949.

WIKEL, S. K. J. A. R. O. E. Host immunity to ticks. v. 41, n. 1, p. 1-22, 1996. ISSN 0066-4170.

XU, C.; NG, D. T. J. N. R. M. C. B. Glycosylation-directed quality control of protein folding. v. 16, n. 12, p. 742, 2015. ISSN 1471-0080.

ZHAO, X.; LI, G.; LIANG, S. Several affinity tags commonly used in chromatographic purification. **Journal of analytical methods in chemistry**, v. 2013, 2013. ISSN 2090-8865.

10. Anexos

Anexo 1. Códon-otimização da RnTrypsin para expressão em sistema bacteriano.

Anexo 2. Códon-otimização da RnKazalPI para expressão em sistema bacteriano.

Anexo 3. Clonagem da RnTrypsin no vetor pET100/D-TOPO.

Anexo 4. Clonagem da RnKazalPI no vetor pET100/D-TOPO.

Anexo 5. Declaração do Comitê de Ética no Uso Animal da Universidade de Brasília - UnB.

Anexo 6. Resultados do software SignalP para a RnTrypsin.

Anexo 7. Resultados do software TargetP para a RnTrypsin.

Anexo 8. Resultados do software ProtParam para a RnTrypsin.

Anexo 9. Resultados do software NetPhos para a RnTrypsin.

Anexo 10. Resultados do software NetNGlyc para a RnTrypsin.

Anexo 11. Resultados do software NetOGlyc para a RnTrypsin.

Anexo 12. Resultados do software Protein Blast para a RnTrypsin.

Anexo 13. Resultados do software SignalP para a RnKazalPI.

Anexo 14. Resultados do software TargetP para a RnKazalPI.

Anexo 15. Resultados do software ProtParam para a RnKazalPI.

Anexo 16. Resultados do software NetPhos para a RnKazalPI.

Anexo 17. Resultados do software NetNGlyc para a RnKazalPI.

Anexo 18. Resultados do software NetOGlyc para a RnKazalPI.

Anexo 19. Resultados do software Protein Blast para a RnKazalPI

Anexo 1

Códon-otimização da RnTrypsin para expressão em sistema bacteriano.

Your GeneOptimizer Assisted Sequence Analysis

invitrogen
by Thermo Fisher Scientific

Customer: **UNB, Carla Nunes de Araújo**
Gene name: **RnTrypsin**
Scientist: **Markus Baumgartner**
Optimized for expression in: ***Escherichia coli***
Analysis created: **2/16/2017** (4:20 PM)

Construct No: **17AASFZC**
Gene length: **681 bp**
Ref. No: **2081427**



Gene design according to DIN EN ISO 9001:2008

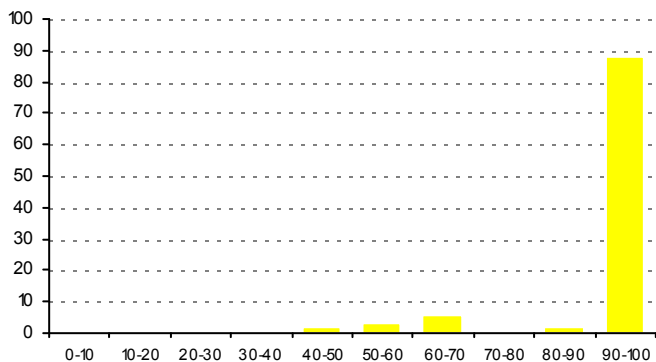
The codon usage was adapted to the codon bias of *Escherichia coli* genes.
In addition, regions of very high (> 80 %) or very low (< 30 %) GC content have been avoided where possible.

During the optimization process the following cis-acting sequence motifs were avoided where applicable:

- internal TATA-boxes, chi-sites and ribosomal entry sites
- AT-rich or GC-rich sequence stretches
- RNA instability motifs
- repeat sequences and RNA secondary structures
- (cryptic) splice donor and acceptor sites in higher eucaryotes

Optimized with GeneOptimizer

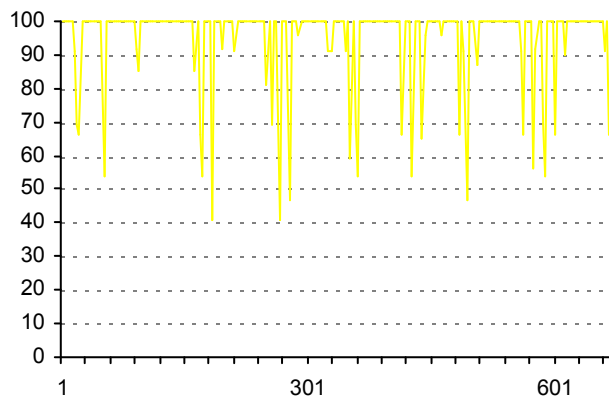
Codon Quality Distribution



Codon Adaptation Index : **0.94**

Optimized with GeneOptimizer

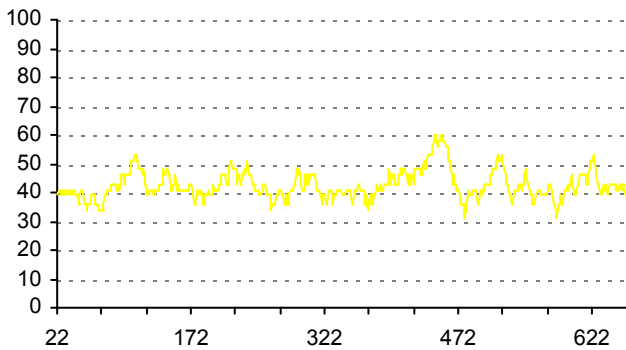
Codon Quality Plot



The histograms show the percentage of sequence codons which fall into a certain quality class. The quality value of the most frequently used codon for a given amino acid in the desired expression system is set to 100, the remaining codons are scaled accordingly (see also Sharp, P.M., Li, W.H., Nucleic Acids Res. 15 (3), 1987). The plots show the quality of the used codon at the indicated codon position.

Optimized with GeneOptimizer

GC Content



Average GC content: **42 %**

The plot shows the GC content in a 40 bp window centered at the indicated nucleotide position.

Summary

The optimization was successful: Negative cis-acting sites (such as splice sites, TATA-boxes, etc.) which may negatively influence expression were eliminated wherever possible. GC content was adjusted to prolong mRNA half life. Codon usage was adapted to the bias of *Escherichia coli* resulting in a CAI* value of **0.94**. The optimized gene should therefore allow high and stable expression rates in *Escherichia coli*.

Please note: for mammalian, insect and yeast expression, we recommend to add a KOZAK sequence upstream of the construct.

* CAI: The parameter CAI (codon adaptation index) describes how well the codons match the codon usage preference of the target organism. Thus, a CAI of 1.0 would be perfect. However, a CAI of > 0.9 is considered as very good (i.e. allowing high expression).

** In some rare cases negatively cis acting motifs cannot be removed due to the particular amino acid sequence or restriction site and codon quality constraints. In all of these cases the optimization scientist chooses an optimal compromise for each individual site.

GeneArt™'s gene optimization process is certified according to the standards of **ISO 9001** (registration no. 12 100 24212). Our internal quality management system covers all aspects from gene optimization to gene synthesis, DNA sequencing as well as sales and service. I certify to the best of my knowledge that: (1) no special safety precautions are needed for the production of the below mentioned DNA and (2) the below mentioned DNA sequence will be used in accordance with applicable laws and regulations.

I agree that Thermo Fisher Scientific shall synthesize the following sequence 17AASFZC according to the conditions described in the quotation, including the presence or absence of:

- stop codons, as per my request or lack thereof
- 5', 3' and internal restriction sites for downstream cloning and processing, as per my request or lack thereof

date

signature

Optimized Sequence

Customer: **UNB,Carla Nunes de Araújo**

Gene name: **RnTrypsin**

Construct No: **17AASFZC**

Gene length: **681 bp**

Optimized for expression in: ***Escherichia coli***

Analysis created: **2/16/2017 (4:20 PM)**

```
AATCCGCAGAGCTATGTGAACATCAAAGCGAAGAAATTGATCTGACCCAGCAAGAACAG
1  -----+-----+-----+-----+-----+-----+-----+
TTAGGCGTCTCGATACACTTGTAGTTTTTCGCTTCTTTAACTAGACTGGGTCGTTCTTGTC
N P Q S Y V N I K S E E I D L T Q Q E Q
1 3 5 7 9 11 13 15 17 19

AAATATGGTGAAAAAGTTACCAATTGTAGCTGCGGCTGGACCAATAAAGCACGTATTGTT
61 -----+-----+-----+-----+-----+-----+-----+
TTTATACCACCTTTTTCAATGGTTAACATCGACGCCGACCTGGTTATTTTCGTGCATAACAA
K Y G E K V T N C S C G W T N K A R I V
21 23 25 27 29 31 33 35 37 39

GGTGGTTCGTGAAACCCTGAAAAATGAATTTCCGCTGATGGCAGGCATTATCGACATGGAA
121 -----+-----+-----+-----+-----+-----+-----+
CCACCAGCACTTTGGGACTTTTTACTTAAAGGCGACTACCGTCCGTAATAGCTGTACCTT
G G R E T L K N E F P L M A G I I D M E
41 43 45 47 49 51 53 55 57 59

AAAAAGTTTCTGTTTTGCGGTGCAACCATTGTGACCCAGAATCATGCAATTACCGCAAGC
181 -----+-----+-----+-----+-----+-----+-----+
TTTTTCAAAGACAAAACGCCACGTTGGTAACACTGGGTCCTTAGTACGTTAATGGCGTTTCG
K K F L F C G A T I V T Q N H A I T A S
61 63 65 67 69 71 73 75 77 79

CATGTACCACACCGAACAAAAAAGAAACTGGGTTTAGTTGTTGGTGCCCATGATGTT
241 -----+-----+-----+-----+-----+-----+-----+
GTAACATGGTGTGGCTTGTTTTTTTTTCTTTGACCCAAATCAACAACCACGGGTACTACAA
H C T T P N K K K K L G L V V G A H D V
81 83 85 87 89 91 93 95 97 99

ACCAAACCGGATGAAAAAGCAGATGTGGTGGAAATTAAGAAACCGTGGAACACGAAAAC
301 -----+-----+-----+-----+-----+-----+-----+
TGGTTTGGCCTACTTTTTTCGTCTACACCACCTTTAATTTCTTTGGCACCTTGTGCTTTTG
T K P D E K A D V V E I K E T V E H E N
101 103 105 107 109 111 113 115 117 119
```

361 TACAGCAGCAAAAGCTATCATAATGATGTTGCACTGCTGGTTCTGAGCCGTAGCATCAAA
 -----+-----+-----+-----+-----+-----+
 ATGTCGTCGTTTTTCGATAGTATTACTACAACGTGACGACCAAGACTCGGCATCGTAGTTT
Y S S K S Y H N D V A L L V L S R S I K
 121 123 125 127 129 131 133 135 137 139

AgeI

421 TTTACCCAAGAAGTTGGTCCCTGCCTGTCTGCCGACCGGTAAAGCCGATCTGGTTAATGAA
 -----+-----+-----+-----+-----+-----+
 AAATGGGTTCTTCAACCAGGACGGACAGACGGCTGGCCATTTTCGGCTAGACCAATTACTT
F T Q E V G P A C L P T G K A D L V N E
 141 143 145 147 149 151 153 155 157 159

481 TATATCAAAGTGTTAGGTTGGGGTCGCCTGAAAACCAAAGGTAAAACCAGCAGCGTTCTG
 -----+-----+-----+-----+-----+-----+
 ATATAGTTTTCACAATCCAACCCAGCGGACTTTTGGTTTTCCATTTTGGTCGTCGCAAGAC
Y I K V L G W G R L K T K G K T S S V L
 161 163 165 167 169 171 173 175 177 179

BclI

541 ATGAAAGTTAATCTGCGTGTGATCAGCATTAAGAGTGC GCAAAAACTACCTGAAAAAA
 -----+-----+-----+-----+-----+-----+
 TACTTTCAATTAGACGCACACTAGTCGTAATTTCTCACGCGTTTTTTGATGGACTTTTTT
M K V N L R V I S I K E C A K N Y L K K
 181 183 185 187 189 191 193 195 197 199

PvuII

601 ATCCCAGCCGATAGTCCGAATCAGCTGTGTACCTATGGTCATGAAAAAGATAGCTGTCAG
 -----+-----+-----+-----+-----+-----+
 TAGGGCTGGCTATCAGGCTTAGTCGACACATGGATAACCAGTACTTTTTTCTATCGACAGTC
I P T D S P N Q L C T Y G H E K D S C Q
 201 203 205 207 209 211 213 215 217 219

661 GTGAGCATCATCGTGTTTTAA
 -----+-----+
 CACTCGTAGTAGCACAAAATT
V S I I V F *
 221 223 225 227

Biosecurity Compliance

We routinely screen all incoming orders for biosecurity according to international guidelines. We therefore reserve the right to refuse any order or to involve government authorities, if a DNA sequence may be of dual-use nature.

Further reading: Buegl *et al.* (2007), DNA synthesis and biological security, Nature Biotechnology
 Free pdf available at <http://dx.doi.org/doi:10.1038/nbt0607-627>

Anexo 2

Códon-otimização da RnKazalPI para expressão em sistema bacteriano.

Your GeneOptimizer Assisted Sequence Analysis

invitrogen
by Thermo Fisher Scientific

Customer: **UNB, Carla Nunes de Araújo**
Gene name: **RnKazalPI**
Scientist: **Markus Baumgartner**
Optimized for expression in: ***Escherichia coli***
Analysis created: **2/16/2017** (4:20 PM)

Construct No: **17AASFYC**
Gene length: **174 bp**
Ref. No: **2081426**



Gene design according to DIN EN ISO 9001:2008

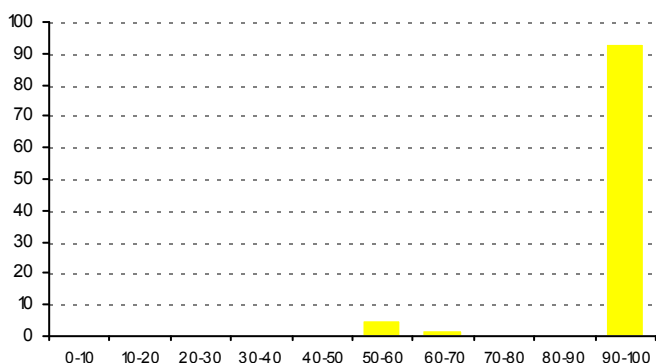
The codon usage was adapted to the codon bias of *Escherichia coli* genes.
In addition, regions of very high (> 80 %) or very low (< 30 %) GC content have been avoided where possible.

During the optimization process the following cis-acting sequence motifs were avoided where applicable:

- internal TATA-boxes, chi-sites and ribosomal entry sites
- AT-rich or GC-rich sequence stretches
- RNA instability motifs
- repeat sequences and RNA secondary structures
- (cryptic) splice donor and acceptor sites in higher eucaryotes

Optimized with GeneOptimizer

Codon Quality Distribution



Codon Adaptation Index : **0.95**

Optimized with GeneOptimizer

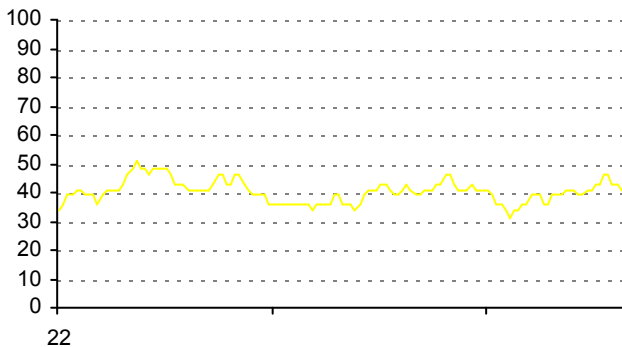
Codon Quality Plot



The histograms show the percentage of sequence codons which fall into a certain quality class. The quality value of the most frequently used codon for a given amino acid in the desired expression system is set to 100, the remaining codons are scaled accordingly (see also Sharp, P.M., Li, W.H., Nucleic Acids Res. 15 (3), 1987). The plots show the quality of the used codon at the indicated codon position.

Optimized with GeneOptimizer

GC Content



Average GC content: **40 %**

The plot shows the GC content in a 40 bp window centered at the indicated nucleotide position.

Summary

The optimization was successful: Negative cis-acting sites (such as splice sites, TATA-boxes, etc.) which may negatively influence expression were eliminated wherever possible. GC content was adjusted to prolong mRNA half life. Codon usage was adapted to the bias of *Escherichia coli* resulting in a CAI* value of **0.95**. The optimized gene should therefore allow high and stable expression rates in *Escherichia coli*.

Please note: for mammalian, insect and yeast expression, we recommend to add a KOZAK sequence upstream of the construct.

* CAI: The parameter CAI (codon adaptation index) describes how well the codons match the codon usage preference of the target organism. Thus, a CAI of 1.0 would be perfect. However, a CAI of > 0.9 is considered as very good (i.e. allowing high expression).

** In some rare cases negatively cis acting motifs cannot be removed due to the particular amino acid sequence or restriction site and codon quality constraints. In all of these cases the optimization scientist chooses an optimal compromise for each individual site.

GeneArt™'s gene optimization process is certified according to the standards of **ISO 9001** (registration no. 12 100 24212). Our internal quality management system covers all aspects from gene optimization to gene synthesis, DNA sequencing as well as sales and service. I certify to the best of my knowledge that: (1) no special safety precautions are needed for the production of the below mentioned DNA and (2) the below mentioned DNA sequence will be used in accordance with applicable laws and regulations.

I agree that Thermo Fisher Scientific shall synthesize the following sequence 17AASFYC according to the conditions described in the quotation, including the presence or absence of:

- stop codons, as per my request or lack thereof
- 5', 3' and internal restriction sites for downstream cloning and processing, as per my request or lack thereof

date

signature

Optimized Sequence

Customer: **UNB,Carla Nunes de Araújo**

Gene name: **RnKazalPI**

Construct No: **17AASFYC**

Gene length: **174 bp**

Optimized for expression in: ***Escherichia coli***

Analysis created: **2/16/2017 (4:20 PM)**

AgeI

GATGAAGTGGAATGTAATACCGATTGCGGTGATAAATATGAACCGGTTTGTGCACAGCCG
1 -----+-----+-----+-----+-----+-----+-----+
CTACTTCACCTTACATTATGGCTAACGCCACTATTTATACTTGGCCAAACACGTGTCCGGC
D E V E C N T D C G D K Y E P V C A Q P
1 3 5 7 9 11 13 15 17 19

BstBI

AGTGATGAAGAAAAATTCGAAACCTTTAAAAACCTGTGCGAGCTGGATAAAGCCATGAAA
61 -----+-----+-----+-----+-----+-----+-----+
TCACTACTTCTTTTTAAGCTTTGGAAATTTTTGGACACGCTCGACCTATTTTCGGTACTTT
S D E E K F E T F K N L C E L D K A M K
21 23 25 27 29 31 33 35 37 39

TGCAATGATGAGAAAACCTGGAAATTTGTGAGCGAAGGTGAATGCACCGATTAA
121 -----+-----+-----+-----+-----+-----+-----+
ACGTTACTACTCTTTTGGACCTTTAAACACTCGCTTCCACTTACGTGGCTAATT
C N D E K T W K F V S E G E C T D *
41 43 45 47 49 51 53 55 57

Biosecurity Compliance

We routinely screen all incoming orders for biosecurity according to international guidelines. We therefore reserve the right to refuse any order or to involve government authorities, if a DNA sequence may be of dual-use nature.

Further reading: Buegl *et al.* (2007), DNA synthesis and biological security, Nature Biotechnology
Free pdf available at <http://dx.doi.org/doi:10.1038/nbt0607-627>

Anexo 3

Clonagem da RnTrypsin no vetor
pET100/D-TOPO.

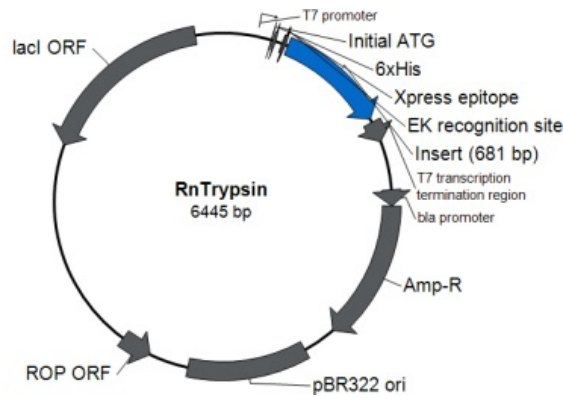
Sequence name: RnTrypsin

Express cloning vector: pET100/D-TOPO

Sequence type: DNA

TSE free: No

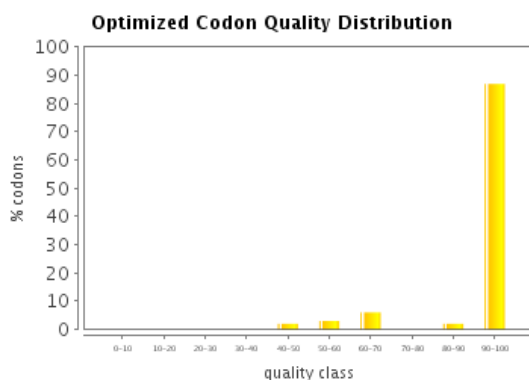
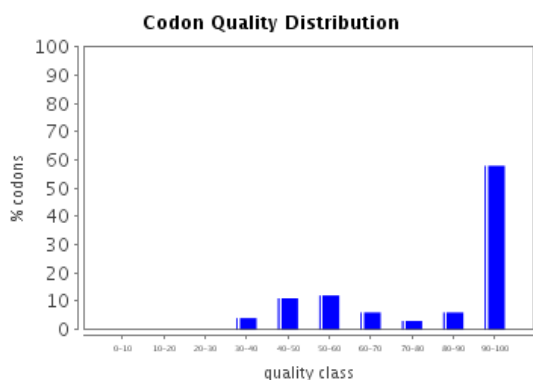
Biosafety level: Level 1



Sequence name / optimized for
RnTrypsin/ Escherichia coli

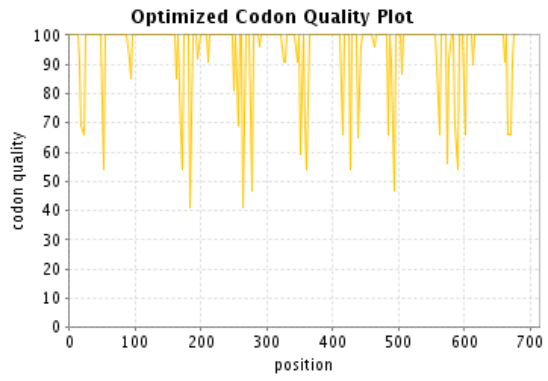
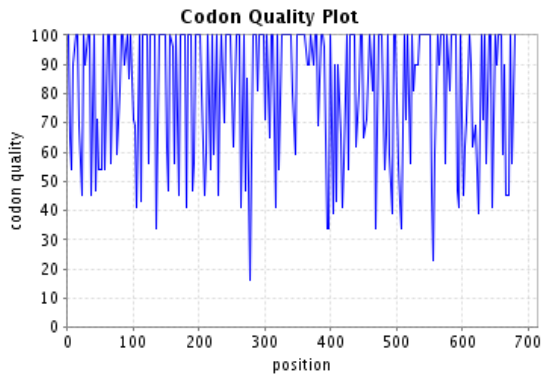
ORF	Protected sites	Protected areas	Motifs to avoid
1-681	[AAT...TAA]		

1.	N P Q S Y V N I K S E E I D L T Q Q E Q K Y G
	AATCCGCAGAGCTATGTGAACATCAAAAAGCGAAGAAATTGATCTGACCCAGCAAGAACAGAAATATGGT
70.	E K V T N C S C G W T N K A R I V G G R E T L
	GAAAAAGTTACCAATTGTAGCTGCGGCTGGACCAATAAAGCACGTATTGTTGGTGGTCGTGAAACCCCTG
139.	K N E F P L M A G I I D M E K K F L F C G A T
	AAAAATGAATTTCCGCTGATGGCAGGCATTATCGACATGGAAAAAAGTTTCTGTTTTGCGGTGCAACC
208.	I V T Q N H A I T A S H C T T P N K K K K L G
	ATTGTGACCCAGAATCATGCAATTACCGCAAGCCATTGTACCACACCGAACAACAAAAAAGAAATCGGGT
277.	L V V G A H D V T K P D E K A D V V E I K E T
	TTAGTTGTTGGTGCCCATGATGTTACCAAACCGGATGAAAAAGCAGATGTGGTGGAAATTAAGAAACC
346.	V E H E N Y S S K S Y H N D V A L L V L S R S
	GTGGAACACGAAAACACTACAGCAGCAAAAAGCTATCATAATGATGTTGCACTGCTGGTCTGAGCCGTAGC
415.	I K F T Q E V G P A C L P T G K A D L V N E Y
	ATCAAATTTACCCAAGAAGTTGGTCTGCCTGTCTGCCGACCGGTAAGCCGATCTGGTTAATGAATAT
484.	I K V L G W G R L K T K G K T S S V L M K V N
	ATCAAAGTGTTAGGTTGGGGTGCCTGAAAACCAAAGGTAACCAGCAGCGTTCTGATGAAAAGTTAAT
553.	L R V I S I K E C A K N Y L K K I P T D S P N
	CTGCGTGTGATCAGCATTAAAGAGTGCGCAAAAACACTACCTGAAAAAAATCCCGACCGATAGTCCGAAT
622.	Q L C T Y G H E K D S C Q V S I I V F *
	CAGCTGTGTACCTATGGTCATGAAAAGATAGCTGTCAGGTGAGCATCATCGTGTTTTAA

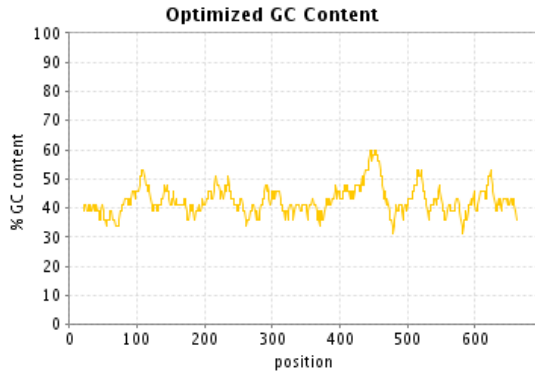
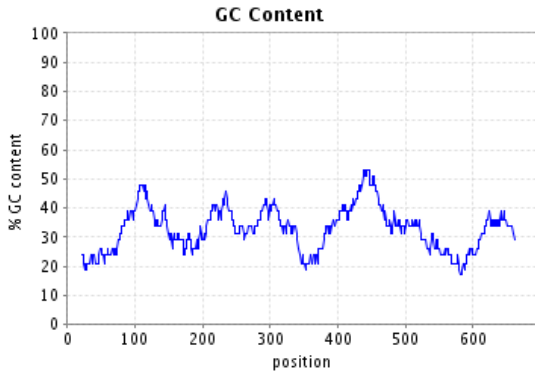


The histograms show the percentage of sequence codons which fall into a certain quality class. The quality value of the most frequently used codon for a

given amino acid in the desired expression system is set to 100, the remaining codons are scaled accordingly (see also Sharp, P.M., Li, W.H., Nucleic Acids Res. 15 (3),1987).



The plots show the quality of the used codon at the indicated codon position.



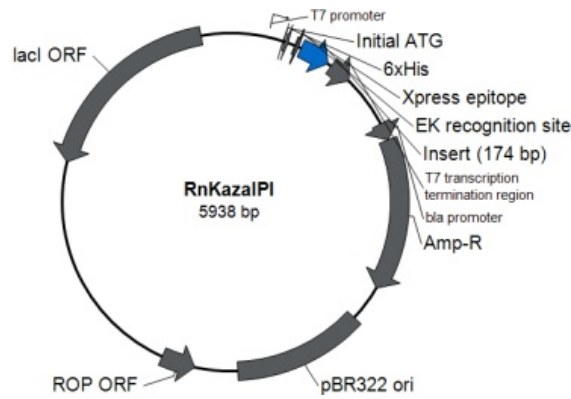
The plots show the GC content in a 40 bp window centered at the indicated nucleotide position.

Anexo 4

Clonagem da RnKazalPI no vetor
pET100/D-TOPO.

Sequence name: RnKazalPI
 Sequence type: DNA
 Biosafety level: Level 1

Express cloning vector: pET100/D-TOPO
 TSE free: No

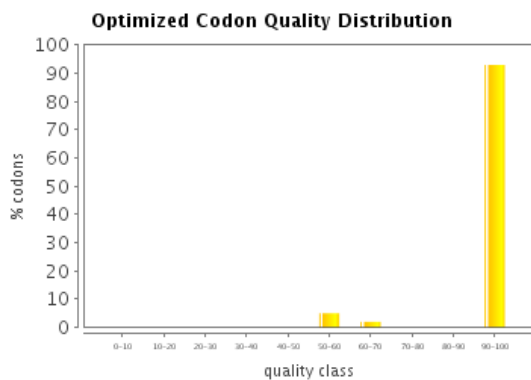
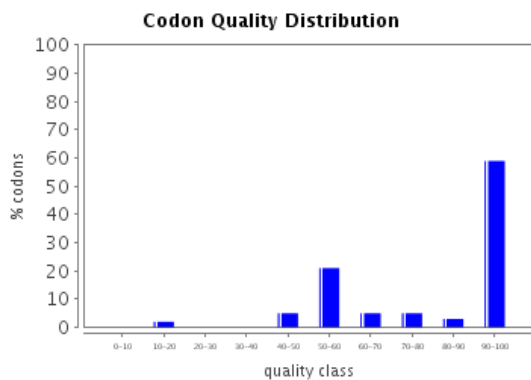


Sequence name / optimized for
RnKazalPI/ Escherichia coli

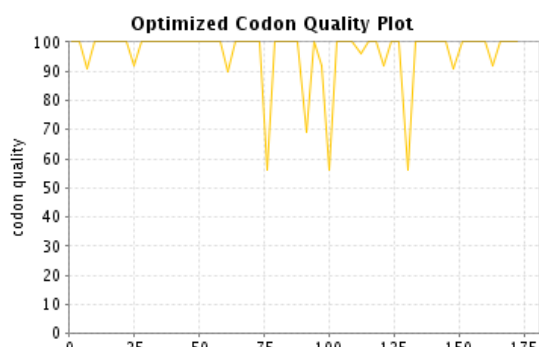
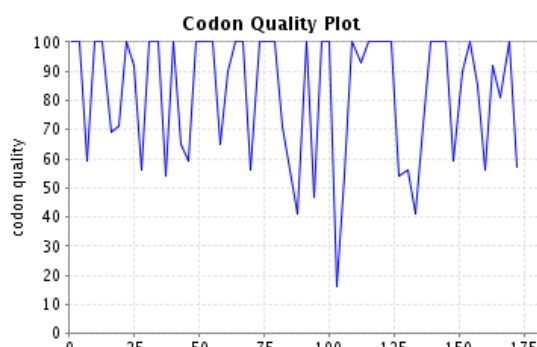
ORF	Protected sites	Protected areas	Motifs to avoid
	1-174 [GAT...TAA]		

```

1.  D E V E C N T D C G D K Y E P V C A Q P S D E
    GATGAAGTGGAATGTAATACCGATTGCGGTGATAAATATGAACCGTTTGTGCACAGCCGAGTGATGAA
70. E K F E T F K N L C E L D K A M K C N D E K T
    GAAAAATTGAAACCTTTAAAAACCTGTGCGAGCTGGATAAAGCCATGAAATGCAATGATGAGAAAACC
139. W K F V S E G E C T D *
    TGGAAATTTGTGAGCGAAGGTGAATGCACCGATTAA
  
```



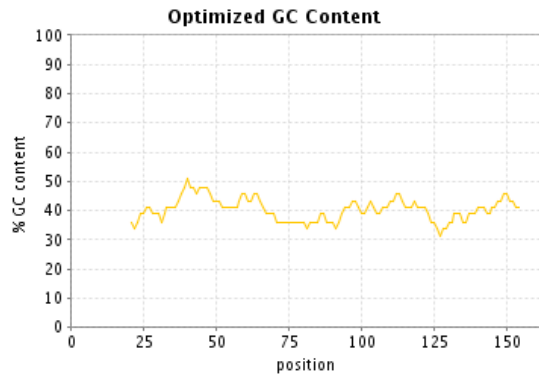
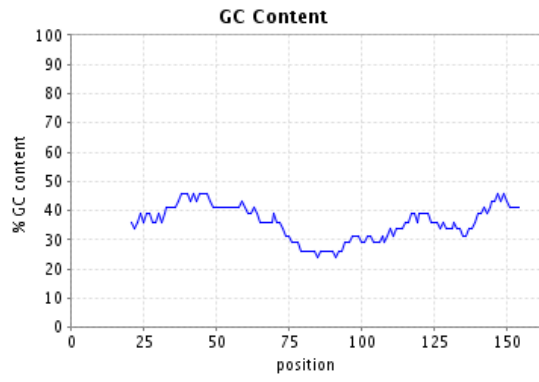
The histograms show the percentage of sequence codons which fall into a certain quality class. The quality value of the most frequently used codon for a given amino acid in the desired expression system is set to 100, the remaining codons are scaled accordingly (see also Sharp, P.M., Li, W.H., Nucleic Acids Res. 15 (3),1987).



0 25 50 75 100 125 150
position

0 25 50 75 100 125 150
position

The plots show the quality of the used codon at the indicated codon position.



The plots show the GC content in a 40 bp window centered at the indicated nucleotide position.

Anexo 5

Declaração do Comitê de Ética no Uso
Animal da Universidade de Brasília - UnB.



Universidade de Brasília
Instituto de Ciências Biológicas
Comitê de Ética no Uso Animal

Brasília, 10 de março de 2016.



DECLARAÇÃO

Declaramos que o projeto intitulado "ADAMTS E ANTÍGENO 5 SALIVARES DE RHODNIUS NEGLECTUS, VETOR DE TRYPANOSOMA CRUZI: IDENTIFICAÇÃO DO PAPEL DESSAS MOLÉCULAS NA INTERAÇÃO VETOR-HOSPEDEIRO", UnBDoC n.º 27323/2016, sob responsabilidade da Professora Carla Nunes de Araújo foi avaliado e aprovado pela Comissão de Ética no Uso Animal (CEUA) da Universidade de Brasília. Este projeto foi aprovado para utilização de *Mus musculus*: 10. A presente aprovação é válida pelo período de 2/04/2016 a 2/04/2018.

Profa. Dra. Paula Diniz Galera
Coordenadora da CEUA - UnB



*Este documento se restringe à avaliação ética do projeto supracitado e não substitui outras licenças e permissões que porventura se façam necessárias.

Anexo 6

Resultados do software SignalP para a
RnTrypsin.

Sequência FASTA - Uniprot

```
>tr|A0A0P4VQ38|A0A0P4VQ38_9HEMI Putative secreted salivary trypsin
OS=Rhodnius neglectus OX=72488 PE=2 SV=1
MRTIYSIAIALLLGKIWVNPQSYVNIKSEEIDLTQQEQKYGEKVTNCSCGWTNKARIVGG
RETLKNEFFPLMAGIIDMEKKFLFCGATIVTQNHAI TASHCTTPNKKKKLGLVVG AHDVTK
PDEKADVVEIKETVEHENYSSKSYHNDVALLVLSRSIKFTQEVGPACLP TGKADLVNEYI
KVLGWGRLKTKGKTSSVLMKVNLRVISIKECAKNYLKKIPTDSPNQLCTYGHEKDSCQVS
IIVF
```

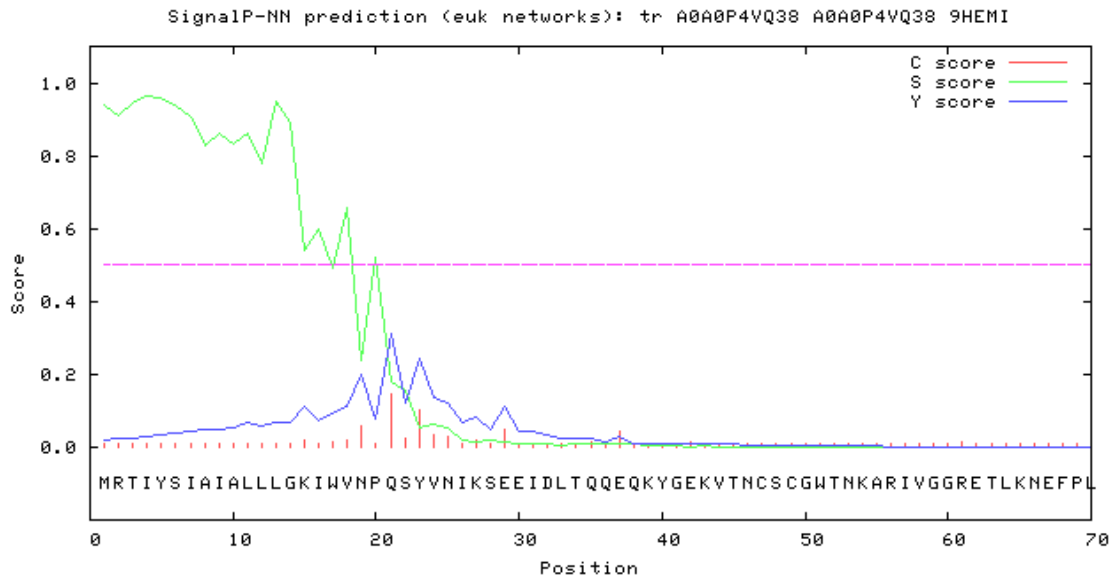
SignalP 3.0 Server - prediction results

Technical University of Denmark

Using neural networks (NN) and hidden Markov models (HMM) trained on eukaryotes

```
>tr_A0A0P4VQ38_A0A0P4VQ38_9HEMI Putative secreted salivary trypsin
OS_Rhodnius neglectus OX_72488 PE_2 SV_1
```

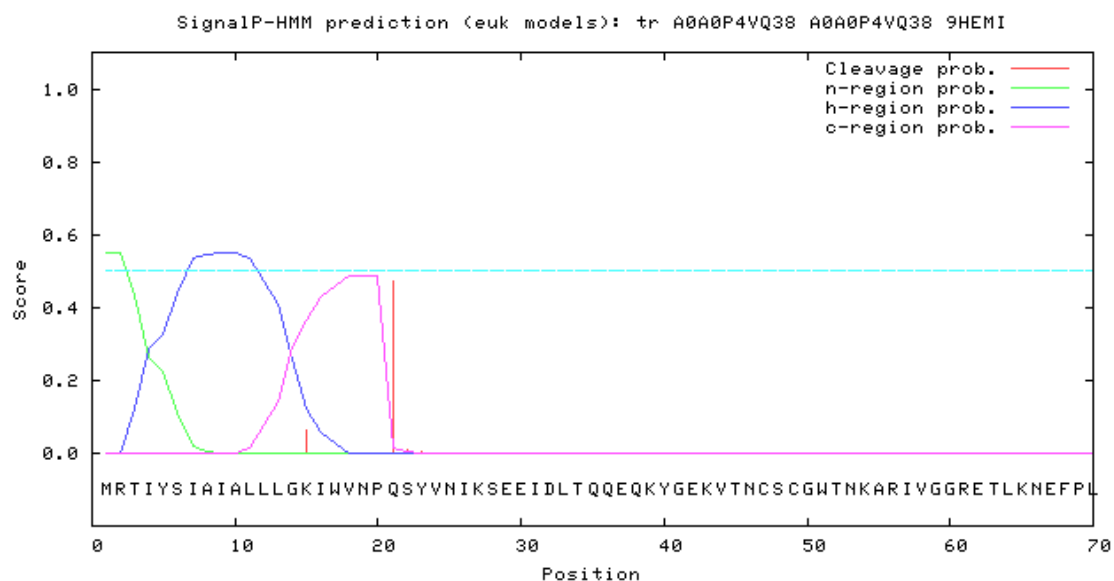
SignalP-NN result:



data

```
>tr_A0A0P4VQ38_A0A0P4 length = 70
# Measure Position Value Cutoff signal peptide?
max. C 21 0.146 0.32 NO
max. Y 21 0.312 0.33 NO
max. S 4 0.963 0.87 YES
mean S 1-20 0.781 0.48 YES
D 1-20 0.546 0.43 YES
# Most likely cleavage site between pos. 20 and 21: VNP-QS
```

SignalP-HMM result:



[data](#)

>tr_A0A0P4VQ38_A0A0P4VQ38_9HEMI

Prediction: Signal peptide

Signal peptide probability: 0.550

Signal anchor probability: 0.000

Max cleavage site probability: 0.474 between pos. 20 and 21

Anexo 7

Resultados do software TargetP para a
RnTrypsin.

TargetP 1.1 Server - prediction results

Technical University of Denmark

```
### targetp v1.1 prediction results #####  
Number of query sequences: 1  
Cleavage site predictions not included.  
Using NON-PLANT networks.
```

Name	Len	mTP	SP	other	Loc	RC
tr_A0A0P4VQ38_A0A0P4	244	0.037	0.879	0.119	S	2
cutoff		0.000	0.000	0.000		

[Explain](#) the output. Go [back](#).

Anexo 8

Resultados do software ProtParam para a
RnTrypsin.

ProtParam

User-provided sequence:

```
      10      20      30      40      50      60
MRTIYSIAIA LLLGKIWVNP QSYVNIKSEE IDLTQQEQKY GEKVTNCSCG WTNKARIVGG

      70      80      90     100     110     120
RETLKNEFPL MAGIIDMEKK FLFCGATIVT QNHAITASHC TTPNKKKKLG LVVGAHDVTK

     130     140     150     160     170     180
PDEKADVVEI KETVEHENYS SKSYHNDVAL LVLSRSIKFT QEVGPACLPT GKADLVNEYI

     190     200     210     220     230     240
KVLGWGRLKT KGKTSSVLMK VNLRVISIKE CAKNYLKKIP TDSPNQLCTY GHEKDSCQVS
```

IIVF

Number of amino acids: 244

Molecular weight: 27191.59

Theoretical pI: 9.06

Amino acid composition: [CSV format](#)

Ala (A)	13	5.3%
Arg (R)	6	2.5%
Asn (N)	13	5.3%
Asp (D)	9	3.7%
Cys (C)	8	3.3%
Gln (Q)	8	3.3%
Glu (E)	16	6.6%
Gly (G)	15	6.1%
His (H)	6	2.5%
Ile (I)	19	7.8%
Leu (L)	20	8.2%
Lys (K)	28	11.5%
Met (M)	4	1.6%
Phe (F)	5	2.0%
Pro (P)	8	3.3%
Ser (S)	16	6.6%
Thr (T)	18	7.4%
Trp (W)	3	1.2%
Tyr (Y)	8	3.3%
Val (V)	21	8.6%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 25

Total number of positively charged residues (Arg + Lys): 34

Atomic composition:

Carbon	C	1210
Hydrogen	H	1964
Nitrogen	N	326
Oxygen	O	358
Sulfur	S	12

Formula: C₁₂₁₀H₁₉₆₄N₃₂₆O₃₅₈S₁₂

Total number of atoms: 3870

Extinction coefficients:

Extinction coefficients are in units of $M^{-1} \text{ cm}^{-1}$, at 280 nm measured in water.

Ext. coefficient 28920
Abs 0.1% (=1 g/l) 1.064, assuming all pairs of Cys residues form cystines

Ext. coefficient 28420
Abs 0.1% (=1 g/l) 1.045, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
 >20 hours (yeast, in vivo).
 >10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 28.19
This classifies the protein as stable.

Aliphatic index: 92.62

Grand average of hydropathicity (GRAVY): -0.242

Anexo 9

Resultados do software NetPhos para a
RnTrypsin.

NetPhos 3.1 Server - prediction results

Technical University of Denmark

```

>Sequence      316 amino acids
#
# netphos-3.1b prediction results
#
# Sequence      # x   Context      Score   Kinase      Answer
# -----
# Sequence      1 T   ----TRAAP    0.477   PKC         .
# Sequence      1 T   ----TRAAP    0.451   GSK3        .
# Sequence      1 T   ----TRAAP    0.439   CaM-II      .
# Sequence      1 T   ----TRAAP    0.411   cdc2        .
# Sequence      1 T   ----TRAAP    0.364   CKI         .
# Sequence      1 T   ----TRAAP    0.349   DNAPK       .
# Sequence      1 T   ----TRAAP    0.313   p38MAPK     .
# Sequence      1 T   ----TRAAP    0.307   PKG         .
# Sequence      1 T   ----TRAAP    0.269   CKII        .
# Sequence      1 T   ----TRAAP    0.256   ATM         .
# Sequence      1 T   ----TRAAP    0.189   cdk5        .
# Sequence      1 T   ----TRAAP    0.184   PKA         .
# Sequence      1 T   ----TRAAP    0.180   RSK         .
# Sequence      1 T   ----TRAAP    0.103   PKB         .
# Sequence      1 T   ----TRAAP    0.039   unsp        .
#
# Sequence      19 T   MIPXTATIV    0.463   GSK3        .
# Sequence      19 T   MIPXTATIV    0.424   CaM-II      .
# Sequence      19 T   MIPXTATIV    0.382   cdc2        .
# Sequence      19 T   MIPXTATIV    0.374   CKI         .
# Sequence      19 T   MIPXTATIV    0.372   p38MAPK     .
# Sequence      19 T   MIPXTATIV    0.344   DNAPK       .
# Sequence      19 T   MIPXTATIV    0.341   CKII        .
# Sequence      19 T   MIPXTATIV    0.274   PKG         .
# Sequence      19 T   MIPXTATIV    0.250   ATM         .
# Sequence      19 T   MIPXTATIV    0.220   RSK         .
# Sequence      19 T   MIPXTATIV    0.171   cdk5        .
# Sequence      19 T   MIPXTATIV    0.164   PKA         .
# Sequence      19 T   MIPXTATIV    0.105   PKC         .
# Sequence      19 T   MIPXTATIV    0.091   PKB         .
# Sequence      19 T   MIPXTATIV    0.019   unsp        .
#
# Sequence      21 T   PXTATIVES    0.895   unsp        YES
# Sequence      21 T   PXTATIVES    0.469   CKII        .
# Sequence      21 T   PXTATIVES    0.455   CaM-II      .
# Sequence      21 T   PXTATIVES    0.426   GSK3        .
# Sequence      21 T   PXTATIVES    0.424   cdc2        .
# Sequence      21 T   PXTATIVES    0.362   CKI         .
# Sequence      21 T   PXTATIVES    0.345   DNAPK       .
# Sequence      21 T   PXTATIVES    0.339   p38MAPK     .
# Sequence      21 T   PXTATIVES    0.264   ATM         .
# Sequence      21 T   PXTATIVES    0.247   PKG         .
# Sequence      21 T   PXTATIVES    0.216   RSK         .
# Sequence      21 T   PXTATIVES    0.215   PKC         .
# Sequence      21 T   PXTATIVES    0.195   cdk5        .
# Sequence      21 T   PXTATIVES    0.149   PKA         .

```

# Sequence	21	T	PXTATIVES	0.111	PKB	.
#						
# Sequence	25	S	TIVESECRE	0.714	unsp	YES
# Sequence	25	S	TIVESECRE	0.640	CKII	YES
# Sequence	25	S	TIVESECRE	0.440	CKI	.
# Sequence	25	S	TIVESECRE	0.439	GSK3	.
# Sequence	25	S	TIVESECRE	0.431	CaM-II	.
# Sequence	25	S	TIVESECRE	0.352	cdc2	.
# Sequence	25	S	TIVESECRE	0.349	DNAPK	.
# Sequence	25	S	TIVESECRE	0.318	RSK	.
# Sequence	25	S	TIVESECRE	0.280	p38MAPK	.
# Sequence	25	S	TIVESECRE	0.253	ATM	.
# Sequence	25	S	TIVESECRE	0.249	PKG	.
# Sequence	25	S	TIVESECRE	0.246	PKC	.
# Sequence	25	S	TIVESECRE	0.171	cdk5	.
# Sequence	25	S	TIVESECRE	0.138	PKA	.
# Sequence	25	S	TIVESECRE	0.070	PKB	.
#						
# Sequence	30	T	ECRETEDSA	0.897	unsp	YES
# Sequence	30	T	ECRETEDSA	0.515	CKII	YES
# Sequence	30	T	ECRETEDSA	0.445	CaM-II	.
# Sequence	30	T	ECRETEDSA	0.428	GSK3	.
# Sequence	30	T	ECRETEDSA	0.411	cdc2	.
# Sequence	30	T	ECRETEDSA	0.391	CKI	.
# Sequence	30	T	ECRETEDSA	0.351	DNAPK	.
# Sequence	30	T	ECRETEDSA	0.333	p38MAPK	.
# Sequence	30	T	ECRETEDSA	0.300	RSK	.
# Sequence	30	T	ECRETEDSA	0.292	PKG	.
# Sequence	30	T	ECRETEDSA	0.261	PKA	.
# Sequence	30	T	ECRETEDSA	0.233	ATM	.
# Sequence	30	T	ECRETEDSA	0.172	cdk5	.
# Sequence	30	T	ECRETEDSA	0.112	PKC	.
# Sequence	30	T	ECRETEDSA	0.087	PKB	.
#						
# Sequence	33	S	ETEDSALIV	0.502	cdc2	YES
# Sequence	33	S	ETEDSALIV	0.446	CaM-II	.
# Sequence	33	S	ETEDSALIV	0.437	GSK3	.
# Sequence	33	S	ETEDSALIV	0.396	CKI	.
# Sequence	33	S	ETEDSALIV	0.375	CKII	.
# Sequence	33	S	ETEDSALIV	0.353	PKG	.
# Sequence	33	S	ETEDSALIV	0.342	DNAPK	.
# Sequence	33	S	ETEDSALIV	0.283	RSK	.
# Sequence	33	S	ETEDSALIV	0.275	ATM	.
# Sequence	33	S	ETEDSALIV	0.274	p38MAPK	.
# Sequence	33	S	ETEDSALIV	0.165	PKB	.
# Sequence	33	S	ETEDSALIV	0.156	PKA	.
# Sequence	33	S	ETEDSALIV	0.149	cdk5	.
# Sequence	33	S	ETEDSALIV	0.043	PKC	.
# Sequence	33	S	ETEDSALIV	0.011	unsp	.
#						
# Sequence	40	Y	IVARYTRYP	0.439	EGFR	.
# Sequence	40	Y	IVARYTRYP	0.421	INSR	.
# Sequence	40	Y	IVARYTRYP	0.326	SRC	.
# Sequence	40	Y	IVARYTRYP	0.045	unsp	.
#						
# Sequence	41	T	VARYTRYPS	0.882	unsp	YES
# Sequence	41	T	VARYTRYPS	0.452	GSK3	.
# Sequence	41	T	VARYTRYPS	0.444	PKG	.
# Sequence	41	T	VARYTRYPS	0.415	CaM-II	.
# Sequence	41	T	VARYTRYPS	0.414	cdc2	.
# Sequence	41	T	VARYTRYPS	0.382	CKI	.

# Sequence	41	T	VARYTRYPS	0.349	p38MAPK	.
# Sequence	41	T	VARYTRYPS	0.348	DNAPK	.
# Sequence	41	T	VARYTRYPS	0.328	PKA	.
# Sequence	41	T	VARYTRYPS	0.303	PKC	.
# Sequence	41	T	VARYTRYPS	0.275	CKII	.
# Sequence	41	T	VARYTRYPS	0.255	ATM	.
# Sequence	41	T	VARYTRYPS	0.213	RSK	.
# Sequence	41	T	VARYTRYPS	0.179	cdk5	.
# Sequence	41	T	VARYTRYPS	0.081	PKB	.
#						
# Sequence	43	Y	RYTRYPSIN	0.529	unsp	YES
# Sequence	43	Y	RYTRYPSIN	0.483	INSR	.
# Sequence	43	Y	RYTRYPSIN	0.425	SRC	.
# Sequence	43	Y	RYTRYPSIN	0.412	EGFR	.
#						
# Sequence	45	S	TRYPSINXS	0.967	unsp	YES
# Sequence	45	S	TRYPSINXS	0.479	CaM-II	.
# Sequence	45	S	TRYPSINXS	0.478	GSK3	.
# Sequence	45	S	TRYPSINXS	0.470	PKG	.
# Sequence	45	S	TRYPSINXS	0.456	RSK	.
# Sequence	45	S	TRYPSINXS	0.360	CKI	.
# Sequence	45	S	TRYPSINXS	0.358	cdc2	.
# Sequence	45	S	TRYPSINXS	0.354	DNAPK	.
# Sequence	45	S	TRYPSINXS	0.351	PKA	.
# Sequence	45	S	TRYPSINXS	0.319	ATM	.
# Sequence	45	S	TRYPSINXS	0.299	p38MAPK	.
# Sequence	45	S	TRYPSINXS	0.285	cdk5	.
# Sequence	45	S	TRYPSINXS	0.236	CKII	.
# Sequence	45	S	TRYPSINXS	0.198	PKB	.
# Sequence	45	S	TRYPSINXS	0.122	PKC	.
#						
# Sequence	49	S	SINXSRHxD	0.460	GSK3	.
# Sequence	49	S	SINXSRHxD	0.441	PKG	.
# Sequence	49	S	SINXSRHxD	0.421	CaM-II	.
# Sequence	49	S	SINXSRHxD	0.412	cdc2	.
# Sequence	49	S	SINXSRHxD	0.403	CKII	.
# Sequence	49	S	SINXSRHxD	0.358	CKI	.
# Sequence	49	S	SINXSRHxD	0.354	DNAPK	.
# Sequence	49	S	SINXSRHxD	0.330	PKA	.
# Sequence	49	S	SINXSRHxD	0.327	ATM	.
# Sequence	49	S	SINXSRHxD	0.280	RSK	.
# Sequence	49	S	SINXSRHxD	0.272	p38MAPK	.
# Sequence	49	S	SINXSRHxD	0.207	cdk5	.
# Sequence	49	S	SINXSRHxD	0.173	PKC	.
# Sequence	49	S	SINXSRHxD	0.089	PKB	.
# Sequence	49	S	SINXSRHxD	0.076	unsp	.
#						
# Sequence	57	S	DNIXSNEGL	0.500	CKII	YES
# Sequence	57	S	DNIXSNEGL	0.468	cdc2	.
# Sequence	57	S	DNIXSNEGL	0.464	GSK3	.
# Sequence	57	S	DNIXSNEGL	0.459	CaM-II	.
# Sequence	57	S	DNIXSNEGL	0.409	CKI	.
# Sequence	57	S	DNIXSNEGL	0.404	DNAPK	.
# Sequence	57	S	DNIXSNEGL	0.370	RSK	.
# Sequence	57	S	DNIXSNEGL	0.324	p38MAPK	.
# Sequence	57	S	DNIXSNEGL	0.304	ATM	.
# Sequence	57	S	DNIXSNEGL	0.296	unsp	.
# Sequence	57	S	DNIXSNEGL	0.225	PKG	.
# Sequence	57	S	DNIXSNEGL	0.198	cdk5	.
# Sequence	57	S	DNIXSNEGL	0.157	PKA	.
# Sequence	57	S	DNIXSNEGL	0.087	PKB	.

# Sequence	57	S	DNIXSNEGL	0.051	PKC	.
#						
# Sequence	64	T	GLECTXSXX	0.503	CKI	YES
# Sequence	64	T	GLECTXSXX	0.478	cdc2	.
# Sequence	64	T	GLECTXSXX	0.466	GSK3	.
# Sequence	64	T	GLECTXSXX	0.458	CKII	.
# Sequence	64	T	GLECTXSXX	0.446	CaM-II	.
# Sequence	64	T	GLECTXSXX	0.368	p38MAPK	.
# Sequence	64	T	GLECTXSXX	0.352	DNAPK	.
# Sequence	64	T	GLECTXSXX	0.282	PKG	.
# Sequence	64	T	GLECTXSXX	0.278	ATM	.
# Sequence	64	T	GLECTXSXX	0.254	RSK	.
# Sequence	64	T	GLECTXSXX	0.214	cdk5	.
# Sequence	64	T	GLECTXSXX	0.131	PKA	.
# Sequence	64	T	GLECTXSXX	0.126	PKC	.
# Sequence	64	T	GLECTXSXX	0.093	PKB	.
# Sequence	64	T	GLECTXSXX	0.058	unsp	.
#						
# Sequence	66	S	ECTXSXXPE	0.468	cdc2	.
# Sequence	66	S	ECTXSXXPE	0.467	CaM-II	.
# Sequence	66	S	ECTXSXXPE	0.428	GSK3	.
# Sequence	66	S	ECTXSXXPE	0.381	DNAPK	.
# Sequence	66	S	ECTXSXXPE	0.362	CKI	.
# Sequence	66	S	ECTXSXXPE	0.355	CKII	.
# Sequence	66	S	ECTXSXXPE	0.319	ATM	.
# Sequence	66	S	ECTXSXXPE	0.302	p38MAPK	.
# Sequence	66	S	ECTXSXXPE	0.293	RSK	.
# Sequence	66	S	ECTXSXXPE	0.256	PKA	.
# Sequence	66	S	ECTXSXXPE	0.253	PKG	.
# Sequence	66	S	ECTXSXXPE	0.154	PKC	.
# Sequence	66	S	ECTXSXXPE	0.152	cdk5	.
# Sequence	66	S	ECTXSXXPE	0.111	PKB	.
# Sequence	66	S	ECTXSXXPE	0.024	unsp	.
#						
# Sequence	71	S	XXPESVMRT	0.981	unsp	YES
# Sequence	71	S	XXPESVMRT	0.504	cdc2	YES
# Sequence	71	S	XXPESVMRT	0.456	GSK3	.
# Sequence	71	S	XXPESVMRT	0.427	CaM-II	.
# Sequence	71	S	XXPESVMRT	0.413	p38MAPK	.
# Sequence	71	S	XXPESVMRT	0.370	CKI	.
# Sequence	71	S	XXPESVMRT	0.365	CKII	.
# Sequence	71	S	XXPESVMRT	0.358	PKC	.
# Sequence	71	S	XXPESVMRT	0.350	DNAPK	.
# Sequence	71	S	XXPESVMRT	0.325	ATM	.
# Sequence	71	S	XXPESVMRT	0.255	RSK	.
# Sequence	71	S	XXPESVMRT	0.210	PKG	.
# Sequence	71	S	XXPESVMRT	0.200	PKA	.
# Sequence	71	S	XXPESVMRT	0.195	cdk5	.
# Sequence	71	S	XXPESVMRT	0.091	PKB	.
#						
# Sequence	75	T	SVMRTIYSI	0.447	GSK3	.
# Sequence	75	T	SVMRTIYSI	0.442	CaM-II	.
# Sequence	75	T	SVMRTIYSI	0.376	cdc2	.
# Sequence	75	T	SVMRTIYSI	0.374	CKI	.
# Sequence	75	T	SVMRTIYSI	0.354	p38MAPK	.
# Sequence	75	T	SVMRTIYSI	0.347	DNAPK	.
# Sequence	75	T	SVMRTIYSI	0.321	PKG	.
# Sequence	75	T	SVMRTIYSI	0.311	CKII	.
# Sequence	75	T	SVMRTIYSI	0.272	ATM	.
# Sequence	75	T	SVMRTIYSI	0.262	cdk5	.
# Sequence	75	T	SVMRTIYSI	0.226	RSK	.

# Sequence	75	T	SVMRTIYSI	0.223	PKA	.
# Sequence	75	T	SVMRTIYSI	0.154	PKC	.
# Sequence	75	T	SVMRTIYSI	0.080	PKB	.
# Sequence	75	T	SVMRTIYSI	0.049	unsp	.
#						
# Sequence	77	Y	MRTIYSIAI	0.442	INSR	.
# Sequence	77	Y	MRTIYSIAI	0.419	SRC	.
# Sequence	77	Y	MRTIYSIAI	0.373	EGFR	.
# Sequence	77	Y	MRTIYSIAI	0.072	unsp	.
#						
# Sequence	78	S	RTIYSIAIA	0.471	CaM-II	.
# Sequence	78	S	RTIYSIAIA	0.461	cdc2	.
# Sequence	78	S	RTIYSIAIA	0.440	GSK3	.
# Sequence	78	S	RTIYSIAIA	0.403	PKA	.
# Sequence	78	S	RTIYSIAIA	0.382	CKI	.
# Sequence	78	S	RTIYSIAIA	0.381	DNAPK	.
# Sequence	78	S	RTIYSIAIA	0.340	p38MAPK	.
# Sequence	78	S	RTIYSIAIA	0.332	PKC	.
# Sequence	78	S	RTIYSIAIA	0.328	PKG	.
# Sequence	78	S	RTIYSIAIA	0.301	CKII	.
# Sequence	78	S	RTIYSIAIA	0.274	RSK	.
# Sequence	78	S	RTIYSIAIA	0.270	ATM	.
# Sequence	78	S	RTIYSIAIA	0.164	cdk5	.
# Sequence	78	S	RTIYSIAIA	0.082	PKB	.
# Sequence	78	S	RTIYSIAIA	0.005	unsp	.
#						
# Sequence	94	S	VNPQSYVNI	0.475	GSK3	.
# Sequence	94	S	VNPQSYVNI	0.446	PKC	.
# Sequence	94	S	VNPQSYVNI	0.445	CaM-II	.
# Sequence	94	S	VNPQSYVNI	0.401	RSK	.
# Sequence	94	S	VNPQSYVNI	0.390	cdc2	.
# Sequence	94	S	VNPQSYVNI	0.367	DNAPK	.
# Sequence	94	S	VNPQSYVNI	0.362	p38MAPK	.
# Sequence	94	S	VNPQSYVNI	0.360	CKI	.
# Sequence	94	S	VNPQSYVNI	0.314	CKII	.
# Sequence	94	S	VNPQSYVNI	0.293	ATM	.
# Sequence	94	S	VNPQSYVNI	0.239	PKG	.
# Sequence	94	S	VNPQSYVNI	0.228	PKA	.
# Sequence	94	S	VNPQSYVNI	0.184	cdk5	.
# Sequence	94	S	VNPQSYVNI	0.106	PKB	.
# Sequence	94	S	VNPQSYVNI	0.049	unsp	.
#						
# Sequence	95	Y	NPQSYVNIK	0.955	unsp	YES
# Sequence	95	Y	NPQSYVNIK	0.433	INSR	.
# Sequence	95	Y	NPQSYVNIK	0.375	EGFR	.
# Sequence	95	Y	NPQSYVNIK	0.355	SRC	.
#						
# Sequence	100	S	VNIKSEEID	0.518	CKII	YES
# Sequence	100	S	VNIKSEEID	0.447	CKI	.
# Sequence	100	S	VNIKSEEID	0.445	GSK3	.
# Sequence	100	S	VNIKSEEID	0.427	CaM-II	.
# Sequence	100	S	VNIKSEEID	0.362	RSK	.
# Sequence	100	S	VNIKSEEID	0.360	DNAPK	.
# Sequence	100	S	VNIKSEEID	0.313	p38MAPK	.
# Sequence	100	S	VNIKSEEID	0.313	cdc2	.
# Sequence	100	S	VNIKSEEID	0.296	ATM	.
# Sequence	100	S	VNIKSEEID	0.262	PKG	.
# Sequence	100	S	VNIKSEEID	0.177	cdk5	.
# Sequence	100	S	VNIKSEEID	0.175	unsp	.
# Sequence	100	S	VNIKSEEID	0.151	PKA	.
# Sequence	100	S	VNIKSEEID	0.086	PKB	.

# Sequence	100	S	VNIKSEEID	0.059	PKC	.
#						
# Sequence	106	T	EIDLTQQEQ	0.603	DNAPK	YES
# Sequence	106	T	EIDLTQQEQ	0.561	CKII	YES
# Sequence	106	T	EIDLTQQEQ	0.490	ATM	.
# Sequence	106	T	EIDLTQQEQ	0.435	GSK3	.
# Sequence	106	T	EIDLTQQEQ	0.413	CaM-II	.
# Sequence	106	T	EIDLTQQEQ	0.390	p38MAPK	.
# Sequence	106	T	EIDLTQQEQ	0.373	CKI	.
# Sequence	106	T	EIDLTQQEQ	0.372	cdc2	.
# Sequence	106	T	EIDLTQQEQ	0.269	PKG	.
# Sequence	106	T	EIDLTQQEQ	0.227	PKC	.
# Sequence	106	T	EIDLTQQEQ	0.216	RSK	.
# Sequence	106	T	EIDLTQQEQ	0.155	cdk5	.
# Sequence	106	T	EIDLTQQEQ	0.121	unsp	.
# Sequence	106	T	EIDLTQQEQ	0.103	PKB	.
# Sequence	106	T	EIDLTQQEQ	0.089	PKA	.
#						
# Sequence	112	Y	QEQQYGEKV	0.824	unsp	YES
# Sequence	112	Y	QEQQYGEKV	0.430	SRC	.
# Sequence	112	Y	QEQQYGEKV	0.426	EGFR	.
# Sequence	112	Y	QEQQYGEKV	0.409	INSR	.
#						
# Sequence	117	T	GEKVTNCSC	0.444	CaM-II	.
# Sequence	117	T	GEKVTNCSC	0.435	GSK3	.
# Sequence	117	T	GEKVTNCSC	0.407	PKG	.
# Sequence	117	T	GEKVTNCSC	0.401	cdc2	.
# Sequence	117	T	GEKVTNCSC	0.361	CKI	.
# Sequence	117	T	GEKVTNCSC	0.348	p38MAPK	.
# Sequence	117	T	GEKVTNCSC	0.336	DNAPK	.
# Sequence	117	T	GEKVTNCSC	0.334	CKII	.
# Sequence	117	T	GEKVTNCSC	0.245	RSK	.
# Sequence	117	T	GEKVTNCSC	0.239	ATM	.
# Sequence	117	T	GEKVTNCSC	0.203	PKC	.
# Sequence	117	T	GEKVTNCSC	0.196	cdk5	.
# Sequence	117	T	GEKVTNCSC	0.149	PKA	.
# Sequence	117	T	GEKVTNCSC	0.095	PKB	.
# Sequence	117	T	GEKVTNCSC	0.035	unsp	.
#						
# Sequence	120	S	VTNCSCGWT	0.476	CKI	.
# Sequence	120	S	VTNCSCGWT	0.449	GSK3	.
# Sequence	120	S	VTNCSCGWT	0.437	PKA	.
# Sequence	120	S	VTNCSCGWT	0.412	CaM-II	.
# Sequence	120	S	VTNCSCGWT	0.378	cdc2	.
# Sequence	120	S	VTNCSCGWT	0.365	p38MAPK	.
# Sequence	120	S	VTNCSCGWT	0.363	DNAPK	.
# Sequence	120	S	VTNCSCGWT	0.313	CKII	.
# Sequence	120	S	VTNCSCGWT	0.311	PKG	.
# Sequence	120	S	VTNCSCGWT	0.307	PKC	.
# Sequence	120	S	VTNCSCGWT	0.304	ATM	.
# Sequence	120	S	VTNCSCGWT	0.273	RSK	.
# Sequence	120	S	VTNCSCGWT	0.180	cdk5	.
# Sequence	120	S	VTNCSCGWT	0.109	PKB	.
# Sequence	120	S	VTNCSCGWT	0.019	unsp	.
#						
# Sequence	124	T	SCGWTNKAR	0.771	PKC	YES
# Sequence	124	T	SCGWTNKAR	0.446	GSK3	.
# Sequence	124	T	SCGWTNKAR	0.430	CaM-II	.
# Sequence	124	T	SCGWTNKAR	0.353	CKI	.
# Sequence	124	T	SCGWTNKAR	0.336	DNAPK	.
# Sequence	124	T	SCGWTNKAR	0.333	cdc2	.

# Sequence	124	T	SCGWTNKAR	0.321	p38MAPK	.
# Sequence	124	T	SCGWTNKAR	0.317	PKG	.
# Sequence	124	T	SCGWTNKAR	0.299	CKII	.
# Sequence	124	T	SCGWTNKAR	0.249	ATM	.
# Sequence	124	T	SCGWTNKAR	0.188	cdk5	.
# Sequence	124	T	SCGWTNKAR	0.183	RSK	.
# Sequence	124	T	SCGWTNKAR	0.151	PKA	.
# Sequence	124	T	SCGWTNKAR	0.095	unsp	.
# Sequence	124	T	SCGWTNKAR	0.091	PKB	.
#						
# Sequence	135	T	GGRETLKNE	0.983	unsp	YES
# Sequence	135	T	GGRETLKNE	0.537	PKC	YES
# Sequence	135	T	GGRETLKNE	0.483	PKA	.
# Sequence	135	T	GGRETLKNE	0.441	GSK3	.
# Sequence	135	T	GGRETLKNE	0.439	CaM-II	.
# Sequence	135	T	GGRETLKNE	0.379	PKG	.
# Sequence	135	T	GGRETLKNE	0.374	DNAPK	.
# Sequence	135	T	GGRETLKNE	0.367	CKI	.
# Sequence	135	T	GGRETLKNE	0.360	cdc2	.
# Sequence	135	T	GGRETLKNE	0.352	CKII	.
# Sequence	135	T	GGRETLKNE	0.263	RSK	.
# Sequence	135	T	GGRETLKNE	0.245	p38MAPK	.
# Sequence	135	T	GGRETLKNE	0.226	ATM	.
# Sequence	135	T	GGRETLKNE	0.157	cdk5	.
# Sequence	135	T	GGRETLKNE	0.079	PKB	.
#						
# Sequence	159	T	FCGATIVTQ	0.466	CaM-II	.
# Sequence	159	T	FCGATIVTQ	0.417	GSK3	.
# Sequence	159	T	FCGATIVTQ	0.364	CKI	.
# Sequence	159	T	FCGATIVTQ	0.364	cdc2	.
# Sequence	159	T	FCGATIVTQ	0.355	p38MAPK	.
# Sequence	159	T	FCGATIVTQ	0.339	DNAPK	.
# Sequence	159	T	FCGATIVTQ	0.285	CKII	.
# Sequence	159	T	FCGATIVTQ	0.249	PKG	.
# Sequence	159	T	FCGATIVTQ	0.238	ATM	.
# Sequence	159	T	FCGATIVTQ	0.193	RSK	.
# Sequence	159	T	FCGATIVTQ	0.189	PKA	.
# Sequence	159	T	FCGATIVTQ	0.160	cdk5	.
# Sequence	159	T	FCGATIVTQ	0.092	PKB	.
# Sequence	159	T	FCGATIVTQ	0.086	PKC	.
# Sequence	159	T	FCGATIVTQ	0.017	unsp	.
#						
# Sequence	162	T	ATIVTQNHA	0.595	DNAPK	YES
# Sequence	162	T	ATIVTQNHA	0.440	CaM-II	.
# Sequence	162	T	ATIVTQNHA	0.431	cdc2	.
# Sequence	162	T	ATIVTQNHA	0.420	GSK3	.
# Sequence	162	T	ATIVTQNHA	0.390	CKI	.
# Sequence	162	T	ATIVTQNHA	0.380	ATM	.
# Sequence	162	T	ATIVTQNHA	0.365	PKG	.
# Sequence	162	T	ATIVTQNHA	0.340	CKII	.
# Sequence	162	T	ATIVTQNHA	0.272	p38MAPK	.
# Sequence	162	T	ATIVTQNHA	0.246	cdk5	.
# Sequence	162	T	ATIVTQNHA	0.193	RSK	.
# Sequence	162	T	ATIVTQNHA	0.122	PKA	.
# Sequence	162	T	ATIVTQNHA	0.110	PKC	.
# Sequence	162	T	ATIVTQNHA	0.085	PKB	.
# Sequence	162	T	ATIVTQNHA	0.065	unsp	.
#						
# Sequence	168	T	NHAITASHC	0.453	cdc2	.
# Sequence	168	T	NHAITASHC	0.449	GSK3	.
# Sequence	168	T	NHAITASHC	0.440	CaM-II	.

# Sequence	168	T	NHAI TASHC	0.409	CKI	.
# Sequence	168	T	NHAI TASHC	0.391	PKG	.
# Sequence	168	T	NHAI TASHC	0.345	CKII	.
# Sequence	168	T	NHAI TASHC	0.340	DNAPK	.
# Sequence	168	T	NHAI TASHC	0.260	PKA	.
# Sequence	168	T	NHAI TASHC	0.243	p38MAPK	.
# Sequence	168	T	NHAI TASHC	0.243	RSK	.
# Sequence	168	T	NHAI TASHC	0.238	ATM	.
# Sequence	168	T	NHAI TASHC	0.215	cdk5	.
# Sequence	168	T	NHAI TASHC	0.098	PKC	.
# Sequence	168	T	NHAI TASHC	0.084	PKB	.
# Sequence	168	T	NHAI TASHC	0.020	unsp	.
#						
# Sequence	170	S	AI TASHCTT	0.463	GSK3	.
# Sequence	170	S	AI TASHCTT	0.457	PKA	.
# Sequence	170	S	AI TASHCTT	0.449	CaM-II	.
# Sequence	170	S	AI TASHCTT	0.368	CKI	.
# Sequence	170	S	AI TASHCTT	0.368	DNAPK	.
# Sequence	170	S	AI TASHCTT	0.351	cdc2	.
# Sequence	170	S	AI TASHCTT	0.349	ATM	.
# Sequence	170	S	AI TASHCTT	0.343	p38MAPK	.
# Sequence	170	S	AI TASHCTT	0.318	CKII	.
# Sequence	170	S	AI TASHCTT	0.289	RSK	.
# Sequence	170	S	AI TASHCTT	0.255	PKG	.
# Sequence	170	S	AI TASHCTT	0.220	PKC	.
# Sequence	170	S	AI TASHCTT	0.187	cdk5	.
# Sequence	170	S	AI TASHCTT	0.100	PKB	.
# Sequence	170	S	AI TASHCTT	0.020	unsp	.
#						
# Sequence	173	T	ASHCTTPNK	0.673	PKC	YES
# Sequence	173	T	ASHCTTPNK	0.466	GSK3	.
# Sequence	173	T	ASHCTTPNK	0.423	CaM-II	.
# Sequence	173	T	ASHCTTPNK	0.395	CKI	.
# Sequence	173	T	ASHCTTPNK	0.357	cdc2	.
# Sequence	173	T	ASHCTTPNK	0.341	DNAPK	.
# Sequence	173	T	ASHCTTPNK	0.337	p38MAPK	.
# Sequence	173	T	ASHCTTPNK	0.332	CKII	.
# Sequence	173	T	ASHCTTPNK	0.284	PKG	.
# Sequence	173	T	ASHCTTPNK	0.262	RSK	.
# Sequence	173	T	ASHCTTPNK	0.257	ATM	.
# Sequence	173	T	ASHCTTPNK	0.206	cdk5	.
# Sequence	173	T	ASHCTTPNK	0.110	PKB	.
# Sequence	173	T	ASHCTTPNK	0.103	PKA	.
# Sequence	173	T	ASHCTTPNK	0.031	unsp	.
#						
# Sequence	174	T	SHCTTPNKK	0.656	cdk5	YES
# Sequence	174	T	SHCTTPNKK	0.653	unsp	YES
# Sequence	174	T	SHCTTPNKK	0.558	p38MAPK	YES
# Sequence	174	T	SHCTTPNKK	0.542	PKC	YES
# Sequence	174	T	SHCTTPNKK	0.492	GSK3	.
# Sequence	174	T	SHCTTPNKK	0.448	CKI	.
# Sequence	174	T	SHCTTPNKK	0.429	CaM-II	.
# Sequence	174	T	SHCTTPNKK	0.402	RSK	.
# Sequence	174	T	SHCTTPNKK	0.377	cdc2	.
# Sequence	174	T	SHCTTPNKK	0.341	DNAPK	.
# Sequence	174	T	SHCTTPNKK	0.259	CKII	.
# Sequence	174	T	SHCTTPNKK	0.256	PKG	.
# Sequence	174	T	SHCTTPNKK	0.243	ATM	.
# Sequence	174	T	SHCTTPNKK	0.092	PKB	.
# Sequence	174	T	SHCTTPNKK	0.061	PKA	.
#						

# Sequence	191	T	AHDVTKPDE	0.492	CKII	.
# Sequence	191	T	AHDVTKPDE	0.487	cdc2	.
# Sequence	191	T	AHDVTKPDE	0.453	GSK3	.
# Sequence	191	T	AHDVTKPDE	0.425	CaM-II	.
# Sequence	191	T	AHDVTKPDE	0.363	CKI	.
# Sequence	191	T	AHDVTKPDE	0.346	DNAPK	.
# Sequence	191	T	AHDVTKPDE	0.325	PKG	.
# Sequence	191	T	AHDVTKPDE	0.260	p38MAPK	.
# Sequence	191	T	AHDVTKPDE	0.248	ATM	.
# Sequence	191	T	AHDVTKPDE	0.247	cdk5	.
# Sequence	191	T	AHDVTKPDE	0.210	RSK	.
# Sequence	191	T	AHDVTKPDE	0.165	PKC	.
# Sequence	191	T	AHDVTKPDE	0.132	unsp	.
# Sequence	191	T	AHDVTKPDE	0.123	PKA	.
# Sequence	191	T	AHDVTKPDE	0.107	PKB	.
#						
# Sequence	205	T	EIKETVEHE	0.540	CKII	YES
# Sequence	205	T	EIKETVEHE	0.448	CaM-II	.
# Sequence	205	T	EIKETVEHE	0.437	GSK3	.
# Sequence	205	T	EIKETVEHE	0.422	cdc2	.
# Sequence	205	T	EIKETVEHE	0.378	p38MAPK	.
# Sequence	205	T	EIKETVEHE	0.357	CKI	.
# Sequence	205	T	EIKETVEHE	0.343	DNAPK	.
# Sequence	205	T	EIKETVEHE	0.297	PKG	.
# Sequence	205	T	EIKETVEHE	0.251	ATM	.
# Sequence	205	T	EIKETVEHE	0.226	RSK	.
# Sequence	205	T	EIKETVEHE	0.164	PKA	.
# Sequence	205	T	EIKETVEHE	0.154	PKC	.
# Sequence	205	T	EIKETVEHE	0.145	cdk5	.
# Sequence	205	T	EIKETVEHE	0.117	unsp	.
# Sequence	205	T	EIKETVEHE	0.086	PKB	.
#						
# Sequence	211	Y	EHENYSSKS	0.856	unsp	YES
# Sequence	211	Y	EHENYSSKS	0.420	INSR	.
# Sequence	211	Y	EHENYSSKS	0.413	SRC	.
# Sequence	211	Y	EHENYSSKS	0.353	EGFR	.
#						
# Sequence	212	S	HENYSSKSY	0.857	unsp	YES
# Sequence	212	S	HENYSSKSY	0.749	PKC	YES
# Sequence	212	S	HENYSSKSY	0.499	CKI	.
# Sequence	212	S	HENYSSKSY	0.464	GSK3	.
# Sequence	212	S	HENYSSKSY	0.446	cdc2	.
# Sequence	212	S	HENYSSKSY	0.440	CaM-II	.
# Sequence	212	S	HENYSSKSY	0.344	CKII	.
# Sequence	212	S	HENYSSKSY	0.344	DNAPK	.
# Sequence	212	S	HENYSSKSY	0.305	PKG	.
# Sequence	212	S	HENYSSKSY	0.281	p38MAPK	.
# Sequence	212	S	HENYSSKSY	0.268	RSK	.
# Sequence	212	S	HENYSSKSY	0.259	ATM	.
# Sequence	212	S	HENYSSKSY	0.171	PKA	.
# Sequence	212	S	HENYSSKSY	0.160	cdk5	.
# Sequence	212	S	HENYSSKSY	0.086	PKB	.
#						
# Sequence	213	S	ENYSSKSYH	0.489	CKII	.
# Sequence	213	S	ENYSSKSYH	0.480	cdc2	.
# Sequence	213	S	ENYSSKSYH	0.444	CaM-II	.
# Sequence	213	S	ENYSSKSYH	0.443	GSK3	.
# Sequence	213	S	ENYSSKSYH	0.369	CKI	.
# Sequence	213	S	ENYSSKSYH	0.365	DNAPK	.
# Sequence	213	S	ENYSSKSYH	0.321	PKC	.
# Sequence	213	S	ENYSSKSYH	0.310	ATM	.

# Sequence	213	S	ENYSSKSYH	0.288	PKG	.
# Sequence	213	S	ENYSSKSYH	0.285	p38MAPK	.
# Sequence	213	S	ENYSSKSYH	0.263	RSK	.
# Sequence	213	S	ENYSSKSYH	0.196	PKA	.
# Sequence	213	S	ENYSSKSYH	0.171	cdk5	.
# Sequence	213	S	ENYSSKSYH	0.103	PKB	.
# Sequence	213	S	ENYSSKSYH	0.012	unsp	.
#						
# Sequence	215	S	YSSKSYHND	0.981	unsp	YES
# Sequence	215	S	YSSKSYHND	0.511	cdc2	YES
# Sequence	215	S	YSSKSYHND	0.439	GSK3	.
# Sequence	215	S	YSSKSYHND	0.426	CaM-II	.
# Sequence	215	S	YSSKSYHND	0.424	PKC	.
# Sequence	215	S	YSSKSYHND	0.383	CKII	.
# Sequence	215	S	YSSKSYHND	0.369	CKI	.
# Sequence	215	S	YSSKSYHND	0.369	RSK	.
# Sequence	215	S	YSSKSYHND	0.355	DNAPK	.
# Sequence	215	S	YSSKSYHND	0.284	p38MAPK	.
# Sequence	215	S	YSSKSYHND	0.279	ATM	.
# Sequence	215	S	YSSKSYHND	0.265	PKG	.
# Sequence	215	S	YSSKSYHND	0.249	PKA	.
# Sequence	215	S	YSSKSYHND	0.191	cdk5	.
# Sequence	215	S	YSSKSYHND	0.103	PKB	.
#						
# Sequence	216	Y	SSKSYHNDV	0.539	unsp	YES
# Sequence	216	Y	SSKSYHNDV	0.465	INSR	.
# Sequence	216	Y	SSKSYHNDV	0.332	EGFR	.
# Sequence	216	Y	SSKSYHNDV	0.331	SRC	.
#						
# Sequence	226	S	LLVLSRSIK	0.464	GSK3	.
# Sequence	226	S	LLVLSRSIK	0.446	PKG	.
# Sequence	226	S	LLVLSRSIK	0.441	cdc2	.
# Sequence	226	S	LLVLSRSIK	0.433	CaM-II	.
# Sequence	226	S	LLVLSRSIK	0.399	CKI	.
# Sequence	226	S	LLVLSRSIK	0.362	DNAPK	.
# Sequence	226	S	LLVLSRSIK	0.309	PKA	.
# Sequence	226	S	LLVLSRSIK	0.304	CKII	.
# Sequence	226	S	LLVLSRSIK	0.291	ATM	.
# Sequence	226	S	LLVLSRSIK	0.287	RSK	.
# Sequence	226	S	LLVLSRSIK	0.279	p38MAPK	.
# Sequence	226	S	LLVLSRSIK	0.150	PKC	.
# Sequence	226	S	LLVLSRSIK	0.147	cdk5	.
# Sequence	226	S	LLVLSRSIK	0.077	PKB	.
# Sequence	226	S	LLVLSRSIK	0.016	unsp	.
#						
# Sequence	228	S	VLSRSIKFT	0.877	unsp	YES
# Sequence	228	S	VLSRSIKFT	0.530	PKC	YES
# Sequence	228	S	VLSRSIKFT	0.461	cdc2	.
# Sequence	228	S	VLSRSIKFT	0.439	GSK3	.
# Sequence	228	S	VLSRSIKFT	0.434	CaM-II	.
# Sequence	228	S	VLSRSIKFT	0.385	CKI	.
# Sequence	228	S	VLSRSIKFT	0.348	DNAPK	.
# Sequence	228	S	VLSRSIKFT	0.302	p38MAPK	.
# Sequence	228	S	VLSRSIKFT	0.295	ATM	.
# Sequence	228	S	VLSRSIKFT	0.288	CKII	.
# Sequence	228	S	VLSRSIKFT	0.282	PKG	.
# Sequence	228	S	VLSRSIKFT	0.252	RSK	.
# Sequence	228	S	VLSRSIKFT	0.195	PKA	.
# Sequence	228	S	VLSRSIKFT	0.134	cdk5	.
# Sequence	228	S	VLSRSIKFT	0.084	PKB	.
#						

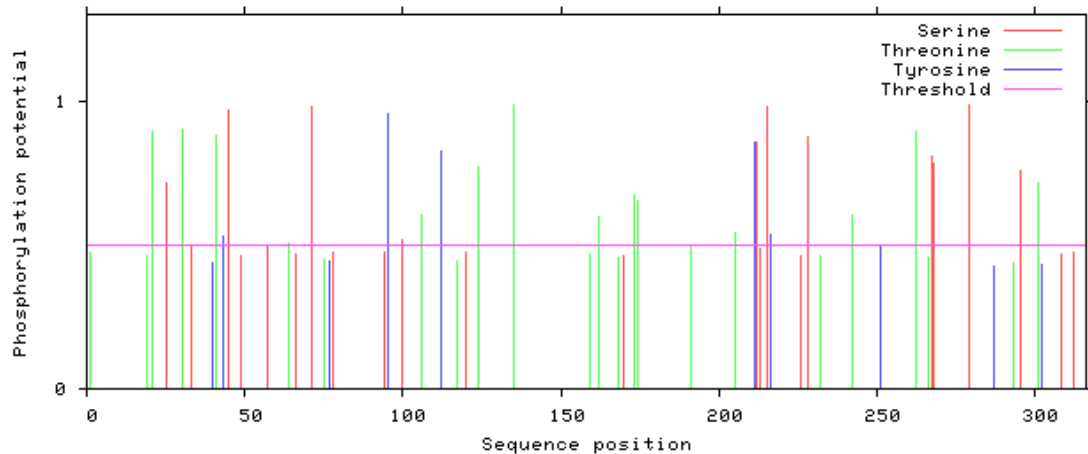
# Sequence	232	T	SIKFTQEVG	0.462	DNAPK	.
# Sequence	232	T	SIKFTQEVG	0.458	GSK3	.
# Sequence	232	T	SIKFTQEVG	0.441	ATM	.
# Sequence	232	T	SIKFTQEVG	0.417	PKC	.
# Sequence	232	T	SIKFTQEVG	0.408	cdc2	.
# Sequence	232	T	SIKFTQEVG	0.406	PKG	.
# Sequence	232	T	SIKFTQEVG	0.404	CaM-II	.
# Sequence	232	T	SIKFTQEVG	0.374	CKI	.
# Sequence	232	T	SIKFTQEVG	0.348	CKII	.
# Sequence	232	T	SIKFTQEVG	0.327	p38MAPK	.
# Sequence	232	T	SIKFTQEVG	0.251	RSK	.
# Sequence	232	T	SIKFTQEVG	0.221	PKB	.
# Sequence	232	T	SIKFTQEVG	0.208	cdk5	.
# Sequence	232	T	SIKFTQEVG	0.143	PKA	.
# Sequence	232	T	SIKFTQEVG	0.059	unsp	.
#						
# Sequence	242	T	ACLPTGKAD	0.602	PKC	YES
# Sequence	242	T	ACLPTGKAD	0.438	GSK3	.
# Sequence	242	T	ACLPTGKAD	0.428	CaM-II	.
# Sequence	242	T	ACLPTGKAD	0.398	PKG	.
# Sequence	242	T	ACLPTGKAD	0.377	cdc2	.
# Sequence	242	T	ACLPTGKAD	0.359	CKI	.
# Sequence	242	T	ACLPTGKAD	0.337	DNAPK	.
# Sequence	242	T	ACLPTGKAD	0.305	p38MAPK	.
# Sequence	242	T	ACLPTGKAD	0.281	CKII	.
# Sequence	242	T	ACLPTGKAD	0.251	ATM	.
# Sequence	242	T	ACLPTGKAD	0.200	PKA	.
# Sequence	242	T	ACLPTGKAD	0.199	RSK	.
# Sequence	242	T	ACLPTGKAD	0.163	cdk5	.
# Sequence	242	T	ACLPTGKAD	0.116	unsp	.
# Sequence	242	T	ACLPTGKAD	0.091	PKB	.
#						
# Sequence	251	Y	LVNEYIKVL	0.500	EGFR	YES
# Sequence	251	Y	LVNEYIKVL	0.467	INSR	.
# Sequence	251	Y	LVNEYIKVL	0.379	SRC	.
# Sequence	251	Y	LVNEYIKVL	0.096	unsp	.
#						
# Sequence	262	T	GRLKTGKGT	0.894	unsp	YES
# Sequence	262	T	GRLKTGKGT	0.884	PKC	YES
# Sequence	262	T	GRLKTGKGT	0.504	RSK	YES
# Sequence	262	T	GRLKTGKGT	0.472	PKB	.
# Sequence	262	T	GRLKTGKGT	0.457	CaM-II	.
# Sequence	262	T	GRLKTGKGT	0.445	GSK3	.
# Sequence	262	T	GRLKTGKGT	0.380	cdc2	.
# Sequence	262	T	GRLKTGKGT	0.373	PKG	.
# Sequence	262	T	GRLKTGKGT	0.358	CKI	.
# Sequence	262	T	GRLKTGKGT	0.357	PKA	.
# Sequence	262	T	GRLKTGKGT	0.337	DNAPK	.
# Sequence	262	T	GRLKTGKGT	0.319	p38MAPK	.
# Sequence	262	T	GRLKTGKGT	0.239	ATM	.
# Sequence	262	T	GRLKTGKGT	0.224	cdk5	.
# Sequence	262	T	GRLKTGKGT	0.194	CKII	.
#						
# Sequence	266	T	TKGKTSSVL	0.454	GSK3	.
# Sequence	266	T	TKGKTSSVL	0.432	CaM-II	.
# Sequence	266	T	TKGKTSSVL	0.429	cdc2	.
# Sequence	266	T	TKGKTSSVL	0.371	PKA	.
# Sequence	266	T	TKGKTSSVL	0.362	CKI	.
# Sequence	266	T	TKGKTSSVL	0.349	PKG	.
# Sequence	266	T	TKGKTSSVL	0.340	DNAPK	.
# Sequence	266	T	TKGKTSSVL	0.298	RSK	.

# Sequence	266	T	TKGKTSSVL	0.274	CKII	.
# Sequence	266	T	TKGKTSSVL	0.254	p38MAPK	.
# Sequence	266	T	TKGKTSSVL	0.243	PKC	.
# Sequence	266	T	TKGKTSSVL	0.225	ATM	.
# Sequence	266	T	TKGKTSSVL	0.177	cdk5	.
# Sequence	266	T	TKGKTSSVL	0.142	unsp	.
# Sequence	266	T	TKGKTSSVL	0.090	PKB	.
#						
# Sequence	267	S	KGKTSSVLM	0.808	unsp	YES
# Sequence	267	S	KGKTSSVLM	0.548	PKA	YES
# Sequence	267	S	KGKTSSVLM	0.500	cdc2	YES
# Sequence	267	S	KGKTSSVLM	0.438	CaM-II	.
# Sequence	267	S	KGKTSSVLM	0.432	GSK3	.
# Sequence	267	S	KGKTSSVLM	0.390	PKG	.
# Sequence	267	S	KGKTSSVLM	0.380	CKII	.
# Sequence	267	S	KGKTSSVLM	0.361	CKI	.
# Sequence	267	S	KGKTSSVLM	0.340	DNAPK	.
# Sequence	267	S	KGKTSSVLM	0.333	p38MAPK	.
# Sequence	267	S	KGKTSSVLM	0.318	RSK	.
# Sequence	267	S	KGKTSSVLM	0.289	PKC	.
# Sequence	267	S	KGKTSSVLM	0.270	ATM	.
# Sequence	267	S	KGKTSSVLM	0.171	cdk5	.
# Sequence	267	S	KGKTSSVLM	0.080	PKB	.
#						
# Sequence	268	S	GKTSSVLMK	0.781	PKC	YES
# Sequence	268	S	GKTSSVLMK	0.499	cdc2	.
# Sequence	268	S	GKTSSVLMK	0.452	CaM-II	.
# Sequence	268	S	GKTSSVLMK	0.427	GSK3	.
# Sequence	268	S	GKTSSVLMK	0.383	PKA	.
# Sequence	268	S	GKTSSVLMK	0.375	CKI	.
# Sequence	268	S	GKTSSVLMK	0.352	DNAPK	.
# Sequence	268	S	GKTSSVLMK	0.325	RSK	.
# Sequence	268	S	GKTSSVLMK	0.306	p38MAPK	.
# Sequence	268	S	GKTSSVLMK	0.300	ATM	.
# Sequence	268	S	GKTSSVLMK	0.254	CKII	.
# Sequence	268	S	GKTSSVLMK	0.221	PKG	.
# Sequence	268	S	GKTSSVLMK	0.161	cdk5	.
# Sequence	268	S	GKTSSVLMK	0.141	unsp	.
# Sequence	268	S	GKTSSVLMK	0.097	PKB	.
#						
# Sequence	279	S	LRVISIKEC	0.987	unsp	YES
# Sequence	279	S	LRVISIKEC	0.615	PKG	YES
# Sequence	279	S	LRVISIKEC	0.512	PKA	YES
# Sequence	279	S	LRVISIKEC	0.474	CaM-II	.
# Sequence	279	S	LRVISIKEC	0.429	cdc2	.
# Sequence	279	S	LRVISIKEC	0.425	GSK3	.
# Sequence	279	S	LRVISIKEC	0.424	RSK	.
# Sequence	279	S	LRVISIKEC	0.388	CKII	.
# Sequence	279	S	LRVISIKEC	0.382	CKI	.
# Sequence	279	S	LRVISIKEC	0.342	DNAPK	.
# Sequence	279	S	LRVISIKEC	0.259	p38MAPK	.
# Sequence	279	S	LRVISIKEC	0.251	ATM	.
# Sequence	279	S	LRVISIKEC	0.177	PKB	.
# Sequence	279	S	LRVISIKEC	0.169	cdk5	.
# Sequence	279	S	LRVISIKEC	0.101	PKC	.
#						
# Sequence	287	Y	CAKNYLKKI	0.423	EGFR	.
# Sequence	287	Y	CAKNYLKKI	0.399	INSR	.
# Sequence	287	Y	CAKNYLKKI	0.291	SRC	.
# Sequence	287	Y	CAKNYLKKI	0.031	unsp	.
#						

# Sequence	293	T	KKIPTDSPN	0.434	GSK3	.
# Sequence	293	T	KKIPTDSPN	0.425	PKC	.
# Sequence	293	T	KKIPTDSPN	0.423	CaM-II	.
# Sequence	293	T	KKIPTDSPN	0.403	cdc2	.
# Sequence	293	T	KKIPTDSPN	0.378	PKG	.
# Sequence	293	T	KKIPTDSPN	0.374	CKII	.
# Sequence	293	T	KKIPTDSPN	0.355	CKI	.
# Sequence	293	T	KKIPTDSPN	0.355	DNAPK	.
# Sequence	293	T	KKIPTDSPN	0.268	p38MAPK	.
# Sequence	293	T	KKIPTDSPN	0.230	ATM	.
# Sequence	293	T	KKIPTDSPN	0.210	RSK	.
# Sequence	293	T	KKIPTDSPN	0.194	PKA	.
# Sequence	293	T	KKIPTDSPN	0.155	cdk5	.
# Sequence	293	T	KKIPTDSPN	0.088	PKB	.
# Sequence	293	T	KKIPTDSPN	0.063	unsp	.
#						
# Sequence	295	S	IPTDSPNQL	0.760	unsp	YES
# Sequence	295	S	IPTDSPNQL	0.546	cdk5	YES
# Sequence	295	S	IPTDSPNQL	0.496	GSK3	.
# Sequence	295	S	IPTDSPNQL	0.450	CaM-II	.
# Sequence	295	S	IPTDSPNQL	0.422	p38MAPK	.
# Sequence	295	S	IPTDSPNQL	0.370	CKI	.
# Sequence	295	S	IPTDSPNQL	0.358	DNAPK	.
# Sequence	295	S	IPTDSPNQL	0.348	RSK	.
# Sequence	295	S	IPTDSPNQL	0.347	PKA	.
# Sequence	295	S	IPTDSPNQL	0.327	cdc2	.
# Sequence	295	S	IPTDSPNQL	0.296	ATM	.
# Sequence	295	S	IPTDSPNQL	0.284	CKII	.
# Sequence	295	S	IPTDSPNQL	0.257	PKG	.
# Sequence	295	S	IPTDSPNQL	0.084	PKB	.
# Sequence	295	S	IPTDSPNQL	0.056	PKC	.
#						
# Sequence	301	T	NQLCTYGHE	0.715	PKC	YES
# Sequence	301	T	NQLCTYGHE	0.451	cdc2	.
# Sequence	301	T	NQLCTYGHE	0.444	GSK3	.
# Sequence	301	T	NQLCTYGHE	0.434	CaM-II	.
# Sequence	301	T	NQLCTYGHE	0.419	CKII	.
# Sequence	301	T	NQLCTYGHE	0.406	p38MAPK	.
# Sequence	301	T	NQLCTYGHE	0.363	CKI	.
# Sequence	301	T	NQLCTYGHE	0.346	DNAPK	.
# Sequence	301	T	NQLCTYGHE	0.298	PKG	.
# Sequence	301	T	NQLCTYGHE	0.243	ATM	.
# Sequence	301	T	NQLCTYGHE	0.210	RSK	.
# Sequence	301	T	NQLCTYGHE	0.174	cdk5	.
# Sequence	301	T	NQLCTYGHE	0.123	PKB	.
# Sequence	301	T	NQLCTYGHE	0.091	PKA	.
# Sequence	301	T	NQLCTYGHE	0.016	unsp	.
#						
# Sequence	302	Y	QLCTYGHEK	0.432	INSR	.
# Sequence	302	Y	QLCTYGHEK	0.424	SRC	.
# Sequence	302	Y	QLCTYGHEK	0.330	EGFR	.
# Sequence	302	Y	QLCTYGHEK	0.011	unsp	.
#						
# Sequence	308	S	HEKDSCQVS	0.469	cdc2	.
# Sequence	308	S	HEKDSCQVS	0.463	GSK3	.
# Sequence	308	S	HEKDSCQVS	0.420	CKII	.
# Sequence	308	S	HEKDSCQVS	0.419	CaM-II	.
# Sequence	308	S	HEKDSCQVS	0.374	ATM	.
# Sequence	308	S	HEKDSCQVS	0.357	CKI	.
# Sequence	308	S	HEKDSCQVS	0.355	PKA	.
# Sequence	308	S	HEKDSCQVS	0.341	DNAPK	.

# Sequence	308 S	HEKDSCQVS	0.330	PKG	.
# Sequence	308 S	HEKDSCQVS	0.324	p38MAPK	.
# Sequence	308 S	HEKDSCQVS	0.252	RSK	.
# Sequence	308 S	HEKDSCQVS	0.161	cdk5	.
# Sequence	308 S	HEKDSCQVS	0.154	unsp	.
# Sequence	308 S	HEKDSCQVS	0.143	PKC	.
# Sequence	308 S	HEKDSCQVS	0.087	PKB	.
#					
# Sequence	312 S	SCQVSIIVF	0.472	CaM-II	.
# Sequence	312 S	SCQVSIIVF	0.446	GSK3	.
# Sequence	312 S	SCQVSIIVF	0.388	CKII	.
# Sequence	312 S	SCQVSIIVF	0.371	ATM	.
# Sequence	312 S	SCQVSIIVF	0.365	cdc2	.
# Sequence	312 S	SCQVSIIVF	0.363	CKI	.
# Sequence	312 S	SCQVSIIVF	0.342	DNAPK	.
# Sequence	312 S	SCQVSIIVF	0.322	PKG	.
# Sequence	312 S	SCQVSIIVF	0.294	PKC	.
# Sequence	312 S	SCQVSIIVF	0.281	PKA	.
# Sequence	312 S	SCQVSIIVF	0.279	RSK	.
# Sequence	312 S	SCQVSIIVF	0.273	p38MAPK	.
# Sequence	312 S	SCQVSIIVF	0.202	cdk5	.
# Sequence	312 S	SCQVSIIVF	0.085	PKB	.
# Sequence	312 S	SCQVSIIVF	0.065	unsp	.
#					
TRAAPVQAAPVQHEMIPXTATIVESECRETEDSALIVARYTRYPSINXSR	#	50			
HXDNI XSNEGLECTXSXXPESVMRTIYSIAIALLLGKIWVNPQSYVNIKS	#	100			
EEIDL TQQEQKYGEKVTNCSCGWTNKARIVGGRETLKNEFPLMAGIIDME	#	150			
KKFLFCGATIVTQNHAITASHCTTPNKKKLGVLVGAHDVTKPDEKADV	#	200			
EIKETVEHENYSSKSYHNDVALLVLSRSIKFTQEVGPACLPTGKADLVNE	#	250			
YIKVLGWGRLKTKGKTSSVLMKVNLRVISIKECAKNYLKKIPTDSPNQLC	#	300			
TYGHEKDSCQVSIIVF	#	350			
%1	T...S....T..S.....T.Y.S.....	#	50		
%1	S.....T.....S.....Y....S	#	100		
%1	T.....Y.....T.....T.....	#	150		
%1	T.....TT.....	#	200		
%1	T....YS..SY.....S.....T.....	#	250		
%1 Y.....	T....SS.....S.....S.....	#	300		
%1 T.....					

NetPhos 3.1a: predicted phosphorylation sites in Sequence



Anexo 10

Resultados do software NetNGlyc para a
RnTrypsin.



NetNGlyc 1.0 Server - prediction results

Technical University of Denmark

Asn-Xaa-Ser/Thr sequons in the sequence output below are highlighted in **blue**.
 Asparagines predicted to be N-glycosylated are highlighted in **red**.

Output for 'RnTrypsin'

Name: RnTrypsin Length: 244

MRTIYSIAIALLLGKIWVNPQSYVNIKSEEIDLTQQEQKYGEKVT**NC**SCGWTNKARIVGGRETLKNEFPLMAGIIDMEKK 80

FLFCGATIVTQNHAITASHCTTPNKKKKLGLVVGADVTKPDEKADVVEIKETVEHE**NY**SSKSYHNDVALLVLSRSIKFT 160

QEVGPACLPKADLVNEYIKVLGWGRLKTKGKTSSVLMKVNLRVISIKECAKNYLKKIPTDSPNQLCTYGHEKDSCQVS 240

IIVF

.....N..... 80

.....N..... 160

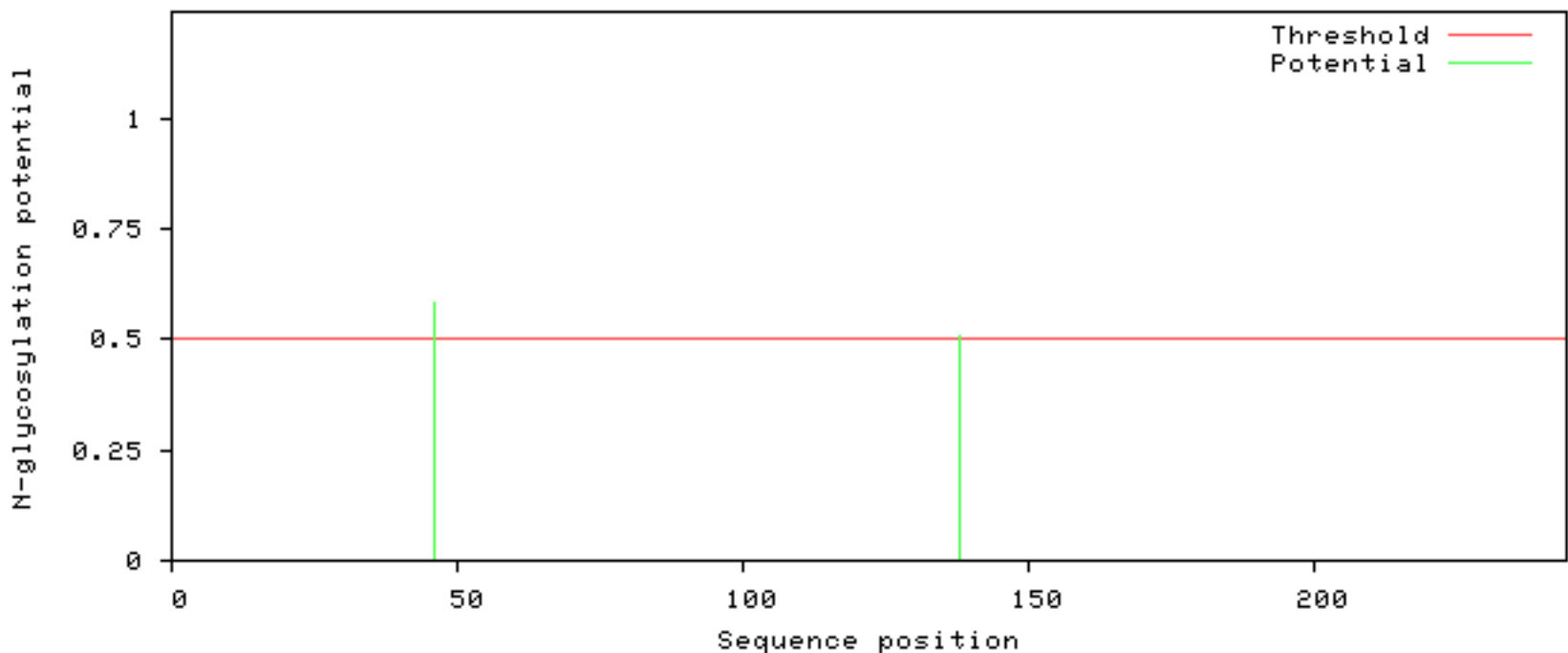
..... 240

..... 320

(Threshold=0.5)

SeqName	Position	Potential	Jury agreement	N-Glyc result
RnTrypsin	46 NCSC	0.5839	(8/9)	+
RnTrypsin	138 NYSS	0.5069	(5/9)	+

NetNGlyc 1.0: predicted N-glycosylation sites in RnTrypsin



[Graphics in PostScript](#)

[Explain](#) the output. Go [back](#).

Anexo 11

Resultados do software NetOGlyc para a
RnTrypsin.



NetOGlyc 4.0 Server - prediction results

Technical University of Denmark

##gff-version 2

##source-version NetOGlyc 4.0.0.13

##date 18-8-4

##Type Protein

#seqname	source	feature	start	end	score	strand	frame	comment
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	3	3	0.0211224	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	6	6	0.0748423	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	22	22	0.239137	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	28	28	0.533699	.	.	#POSITIVE
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	34	34	0.53761	.	.	#POSITIVE
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	45	45	0.422818	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	48	48	0.441598	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	52	52	0.65461	.	.	#POSITIVE
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	63	63	0.0727229	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	87	87	0.0130782	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	90	90	0.0254255	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	96	96	0.0379174	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	98	98	0.0952051	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	101	101	0.0286627	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	102	102	0.083288	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	119	119	0.0935974	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	133	133	0.0512188	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	140	140	0.0337503	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	141	141	0.0373591	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	143	143	0.0144719	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	154	154	0.0611338	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	156	156	0.11453	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	160	160	0.250045	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	170	170	0.089707	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	190	190	0.114732	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	194	194	0.178511	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	195	195	0.280399	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	196	196	0.257949	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	207	207	0.0340069	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	221	221	0.41746	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	223	223	0.176081	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	229	229	0.0769982	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	236	236	0.0152528	.	.	
RNTRYPSIN	netOGlyc-4.0.0.13	CARBOHYD	240	240	0.0102438	.	.	

[Explain](#) the output. Go [back](#).

Anexo 12

Resultados do software Protein Blast para
a RnTrypsin.

BLAST Results

[Questions/comments](#)

Job title: unnamed protein product (244 letters)

RID [P96XF55E014](#) (Expires on 08-05 02:15 am)

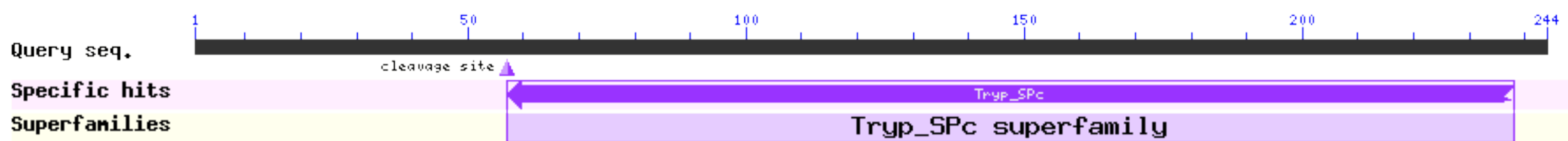
Query ID lcl|Query_125642
Description unnamed protein product
Molecule type amino acid
Query Length 244

Database Name nr
Description All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF excluding environmental samples from WGS projects
Program BLASTP 2.8.0+

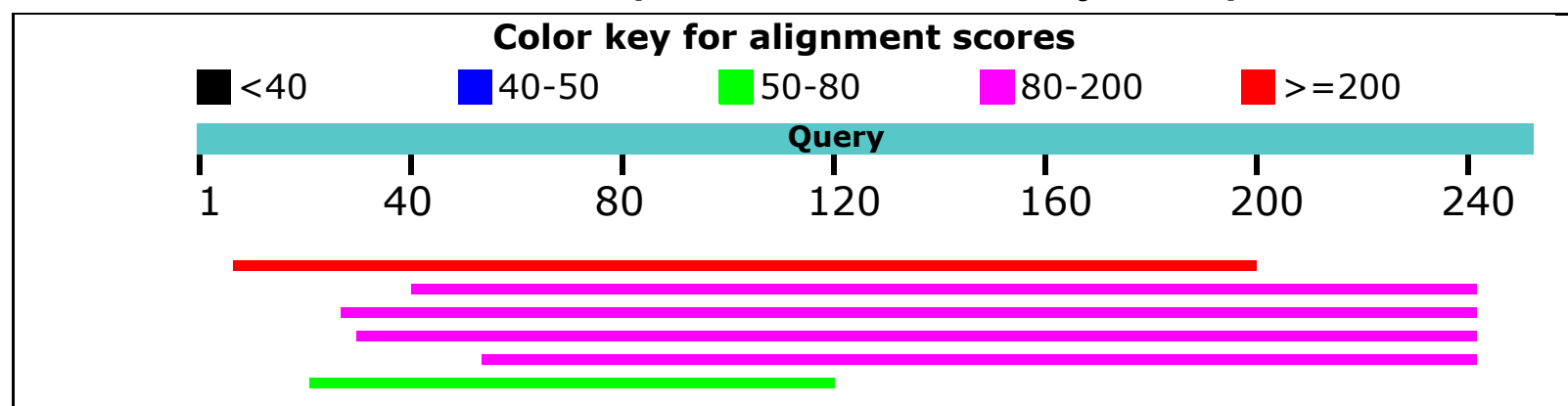
New Analyze your query with [SmartBLAST](#)

Graphic Summary

Putative conserved domains have been detected, click on the image below for detailed results.



Distribution of the top 6 Blast Hits on 6 subject sequences



Descriptions

Sequences producing significant alignments:

Description	Max score	Total score	Query cover	E value	Ident	Accession
secreted salivary trypsin [Triatoma brasiliensis]	308	308	77%	4e-109	75%	ABH09442.1
salivary trypsin [Triatoma infestans]	180	180	81%	3e-57	45%	ABR27829.1
Pc97, similar to serine protease [Panstrongylus chinai]	141	141	86%	5e-42	33%	BBA30652.1
serine protease [Panstrongylus megistus]	134	134	85%	1e-39	36%	ADF97756.1
serine proteinase [Rhodnius prolixus]	87.0	87.0	75%	2e-21	31%	ACF70480.1
unnamed protein product [Triatoma dimidiata]	62.8	62.8	39%	4e-14	38%	BAI53115.1

Alignments

secreted salivary trypsin [Triatoma brasiliensis]

Sequence ID: **ABH09442.1** Length: 197 Number of Matches: 1
Range 1: 8 to 197

Score	Expect	Method	Identities	Positives	Gaps	Frame
308 bits(789)	4e-109()	Compositional matrix adjust.	143/190(75%)	162/190(85%)	0/190(0%)	
Features:						
Query	8	AIALLLGKIWVNPQSYVNIKSEEIDLTOQEQKYGEKVTNCSCGWTNKARIVGGRETLKNE				67
Sbjct	8	AI LLLG W NPQSY+ +KSE D + QK G K TNCSCGWTNKARIVGGRETLKNE				67
Query	68	FPLMAGIIDMEKKFLFCGATIVTQNHAITASHCTTPNKKKLLGLVVG AHDVTKPDEKADV				127
Sbjct	68	FPLMAGI++MEKK LFCGATIVT NHA+TASHCT P K KLGLV+GAHDV+KPDEKAD+				127
Query	128	VEIKETVEHENYSSKSYHNDVALLVLSRSIKFTQEVGPACLP TGKADLVNEYIKVLGWGR				187
Sbjct	128	+EIKET+EHENY+ K YHNDVALL+LSRSIKFTQ VGPACLP TG++D+VNE+IK+LGWGR				187
Query	188	LKTGKTSSV 197				
Sbjct	188	L KTS + LVITRKTSSL 197				

salivary trypsin [Triatoma infestans]

Sequence ID: **ABR27829.1** Length: 308 Number of Matches: 1
Range 1: 41 to 243

Score	Expect	Method	Identities	Positives	Gaps	Frame
180 bits(456)	3e-57()	Compositional matrix adjust.	92/204(45%)	127/204(62%)	7/204(3%)	
Features:						
Query	41	GEKVTNCSCGWTNKA--RIVGGRETLKNEFPLMAGIIDMEKKFLFCGATIVTQNHAITAS				98
Sbjct	41	G+K TNC CGW NK RI+GG ET NE+P+MAG+ K+ LFCG +I+TQ H +TA+ GDKTTNCD CGWANKEDKRIIGGEETNVNEYPMAGLFYKPKELLFCGGSIIITQYHILTAA				100
Query	99	HCTTPNKK--KKLGLVVG AHDVTKPDEKADV--EIKETVEHENYSSKSYHNDVALLVLS				154
Sbjct	101	HCT P ++ + + +V G HD K DE + V ++ V HE Y + +D+A+++L HCTQPFEEDVRDIQIVSGEHDQDKVDESSSTVYIDVLFVPHGYYLIGHRHDAIILLK				160
Query	155	RSIKFTQEVGPACLP TGKADLVNEYIKVLGWGR LKTKGKTSSVLMKVNLRVISIKECAKN				214
Sbjct	161	I +T VGPAC+PT D++ IKVLGWG+L G +S VLMKV L V+ ++EC N DKIVYTNIVGPACMPTEPIDMIGHKIKVLGWGKLSANGPSSKVLKVVYLNVPLEEC--NN				219
Query	215	YLKKIPTDSPNQLCTYGHEKDSCQ 238				
Sbjct	220	+ I Q+CTY KDSCQ THEYIGLVERRQICTYHPTKDSCQ 243				

Pc97, similar to serine protease [Panstrongylus chinai]

Sequence ID: **BBA30652.1** Length: 330 Number of Matches: 1
Range 1: 51 to 264

Score	Expect	Method	Identities	Positives	Gaps	Frame
141 bits(355)	5e-42()	Compositional matrix adjust.	72/216(33%)	126/216(58%)	7/216(3%)	
Features:						
Query	28	SEEIDLTOQEQKYGEKVTNCSCGWTNK--ARIVGGRETLKNEFPLMAGI-IDMEKKFLFC				84
Sbjct	51	S EID ++ G+ TNC+CG N+ ARI+GGR+ E+P +A I + + + C SPEIDSSEHGVTTPGKMETNCTCGRANRGGARIIGRQFPFEYPPFLALISLSTDPERPHC				110
Query	85	GATIVTQNHAITASHCTTPNKKKLLGLV--GAHDVTKPDEKADVVEIKETVEHENYSSK				142
Sbjct	111	G+TI++ H +TA+HCT+ +GL+V H++ K+ VE++ ++HENY+ + GSTIISDRHLLTAAHCTS--KAGGNIGLLVTVAVHEIGN--SPKSYTVEVERLIQHENYNER				168

Query 143 SYHNDVALLVLSRSIKFTQEVGPACLP TGKADLVNEYIKVLGWGRLKTKGKTSSVLMKVN 202
 + ND++LL L ++ T+ V PAC+P ++ +Y++V+GWG + G + + KV+
 Sbjct 169 NLLNDISLLYLKLNPLQLTEVVSPACMPLPGMSVIGKYVVRVIGWGATEAYGPMTVLPQKVD 228

Query 203 LRVISIKECAKNYLKKIPTDSPNQLCTYGHEKDSCQ 238
 L VIS ++C + + ++ TD Q+C K +C
 Sbjct 229 LEVISTEKCHEYWQLQVATDPVTQICRMWENKSACH 264

serine protease [Panstrongylus megistus]

Sequence ID: **ADF97756.1** Length: 309 Number of Matches: 1
 Range 1: 22 to 240

Score	Expect	Method	Identities	Positives	Gaps	Frame
134 bits(338)	1e-39()	Compositional matrix adjust.	80/224(36%)	122/224(54%)	21/224(9%)	
Features:						
Query 31	IDLTQEQKYG----	EKVTNCSGWTK--	ARIVGGRETLKNEFPLMAGIIDMEKKFLFC			84
Sbjct 22	IDEGEDSSEYGVVEGDVGTNCTCGIANKEDQRIVGGEEAKKNEYPMIAEL--	SVQGYHMC				79
Query 85	GATIVTQNHAITASHCTT-----	PNKKKKLGLVVGADHV--	TKPDEKADVVEIKETVEH			136
Sbjct 80	GTTIITKKHVVTAAHCASFSTLDPLPKDMLGVYVGIHEVKGARAFEGGQYFAIEEFITH					139
Query 137	ENYSSKSYHNDVALLVLSRSIKF-TQEVGPACLP TGKADLVNEYIKVLGWGRLKTKGKTS					195
Sbjct 140	PNYHYKNKDHVIAVILLDDEIDFNSHNVPACLPVGNEDIIGDKLKVGLWPAYRADAG					199
Query 196	S-VLMKVNLRVISIKECAKNYLKKIPTDSPNQLCTYGHEKDSCQ			238		
Sbjct 200	TQVLMKVNIAVTHNECAPIYKY---	YNVKTNICMFTQKKDACQ		240		

serine proteinase [Rhodnius prolixus]

Sequence ID: **ACF70480.1** Length: 383 Number of Matches: 1
 Range 1: 132 to 329

Score	Expect	Method	Identities	Positives	Gaps	Frame
87.0 bits(214)	2e-21()	Compositional matrix adjust.	61/200(31%)	99/200(49%)	17/200(8%)	
Features:						
Query 54	KARIVGGRETLKNEFPLMAGIIDMEKKF----	LFCGATIVTQNHAITASHCTTPNK----				105
Sbjct 132	KIRIVGGQPSDLGAWPWLAVLGYRSNRNPTTQWLCGGALVTSRHIVTAAHCTRHP SLSLF					191
Query 106	KKKLGLVVGADHVTKPDEKADVVEIKETVEHENYSSKSYHNDVALLVLSRSIKFTQEVGP					165
Sbjct 192	KVRLGELDLNNDVNDGANPIDV-NIERTIVHPSYNPQKYTDDIAVLKLQNEVPF SRNIQP					250
Query 166	ACLPTG----	KADLVNEYIKVLGWGRLKTKGKTSSVLMKVNLRVISIKECAKNYLKKIPT				221
Sbjct 251	ICLPTTSELREMSLTKKFPFVAGWGSVQFKGPSLTALQEVQVPVVENEECRRAYKAKGAD					310
Query 222	DSPNQLCTYGHE---	KDSCQ		238		
Sbjct 311	IISRQLCA-GFALGGKDACQ			329		

unnamed protein product [Triatoma dimidiata]

Sequence ID: **BAI53115.1** Length: 134 Number of Matches: 1
 Range 1: 18 to 120

Score	Expect	Method	Identities	Positives	Gaps	Frame
62.8 bits(151)	4e-14()	Compositional matrix adjust.	40/105(38%)	58/105(55%)	10/105(9%)	
Features:						
Query 22	SYVNIKSEEIDLTOEQKYGEKVTNCSGWTKA--	RIVGGRETLKNEFPLMAGIIDMEK				79
Sbjct 18	SCANIGDDE-DSSEYGVVDGVDGTNCTCGIANKEDERIVGGEEAKKNEYPMIAELSVYGE					76
Query 80	KFLFCGATIVTQNHAITASHCT-----	TPNKKKKLGLVVGADHV		118		
Sbjct 77	SHM-CGGTIIITKRHVVTAAHCAFYPKSLDPLPKDIIDVHVG IHEV			120		

BLAST is a registered trademark of the National Library of Medicine



[Support center Mailing list](#) [YouTube](#)



[National Library Of Medicine](#)



[National Institutes Of Health](#)



- [U.S. Department of Health & Human Services](#)



- [USA.gov](#)

NCBI

National Center for Biotechnology Information, [U.S. National Library of Medicine](#) 8600 Rockville Pike, Bethesda MD, 20894 USA
[Policies and Guidelines](#) | [Contact](#)

Anexo 13

Resultados do software SignalP para a RnKazalPI

Sequência FASTA - Uniprot

```
>tr|A0A0P4W3X1|A0A0P4W3X1_9HEMI Putative salivary secreted kazaltype
proteinase inhibitor OS=Rhodnius neglectus OX=72488 PE=2 SV=1
MKLSSSVLSVVVIAVFLCMLDKHVFVGDEVECENTDCGDKYEPVCAQPSDEEKFETFKNLCE
LDKAMKCNDEKTWKVFVSEGETD
```

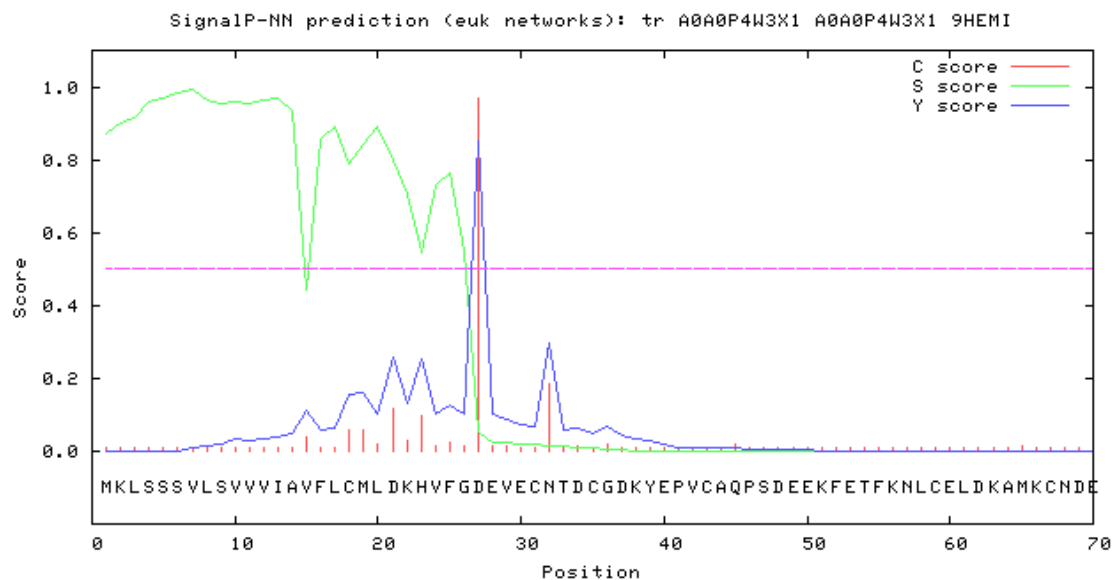
SignalP 3.0 Server - prediction results

Technical University of Denmark

Using neural networks (NN) and hidden Markov models (HMM) trained on
eukaryotes

```
>tr_A0A0P4W3X1_A0A0P4W3X1_9HEMI Putative salivary secreted kazaltype
proteinase inhibitor OS_Rhodnius neglectus OX_72488 PE_2 SV_1
```

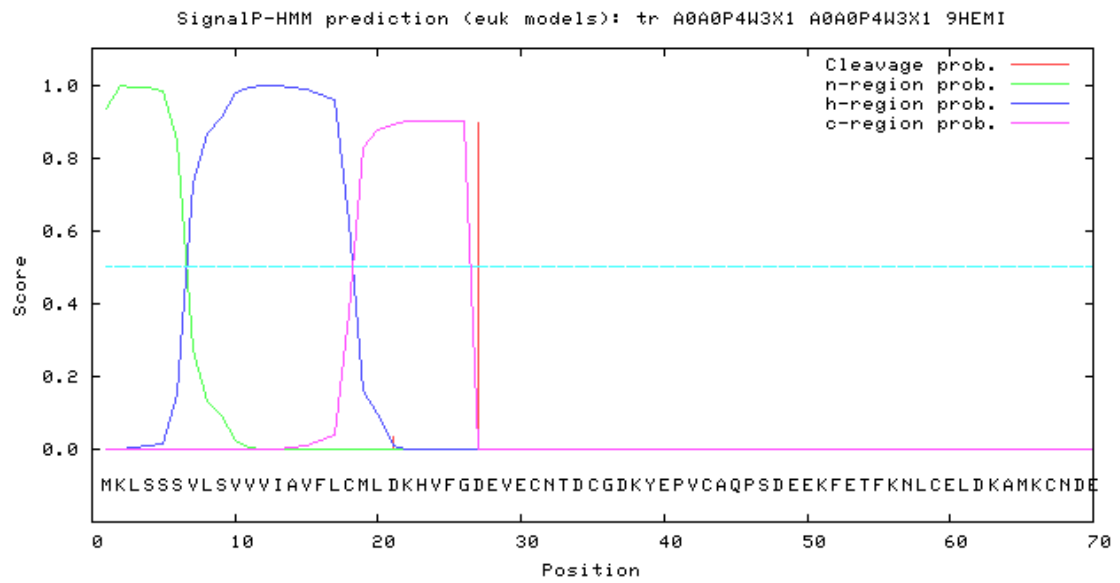
SignalP-NN result:



data

```
>tr_A0A0P4W3X1_A0A0P4 length = 70
# Measure Position Value Cutoff signal peptide?
max. C 27 0.968 0.32 YES
max. Y 27 0.852 0.33 YES
max. S 7 0.991 0.87 YES
mean S 1-26 0.850 0.48 YES
D 1-26 0.851 0.43 YES
# Most likely cleavage site between pos. 26 and 27: VFG-DE
```

SignalP-HMM result:



[data](#)

```
>tr_A0A0P4W3X1_A0A0P4W3X1_9HEMI
```

Prediction: Signal peptide

Signal peptide probability: 0.936

Signal anchor probability: 0.063

Max cleavage site probability: 0.897 between pos. 26 and 27

[gnuplot script](#)

for making the plot(s)

Anexo 14

Resultados do software TargetP para a
RnKazalPI.

TargetP 1.1 Server - prediction results

Technical University of Denmark

```
### targetp v1.1 prediction results #####  
Number of query sequences: 1  
Cleavage site predictions not included.  
Using NON-PLANT networks.
```

Name	Len	mTP	SP	other	Loc	RC
tr_A0A0P4W3X1_A0A0P4	83	0.018	0.950	0.073	S	1
cutoff		0.000	0.000	0.000		

Anexo 15

Resultados do software ProtParam para a
RnKazalPI.

ProtParam

User-provided sequence:

```
      10      20      30      40      50      60
MKLSSSVLSV VVIAVFLCML DKHVFGEVE CNTDCGDKYE PVCAQPSDEE KFETFKNLCE

      70      80
LDKAMKCNDK KTWKVFSEGE CTD
```

Number of amino acids: 83

Molecular weight: 9381.68

Theoretical pI: 4.39

Amino acid composition: [CSV format](#)

Ala (A)	3	3.6%
Arg (R)	0	0.0%
Asn (N)	3	3.6%
Asp (D)	8	9.6%
Cys (C)	7	8.4%
Gln (Q)	1	1.2%
Glu (E)	10	12.0%
Gly (G)	3	3.6%
His (H)	1	1.2%
Ile (I)	1	1.2%
Leu (L)	6	7.2%
Lys (K)	9	10.8%
Met (M)	3	3.6%
Phe (F)	5	6.0%
Pro (P)	2	2.4%
Ser (S)	6	7.2%
Thr (T)	4	4.8%
Trp (W)	1	1.2%
Tyr (Y)	1	1.2%
Val (V)	9	10.8%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 18

Total number of positively charged residues (Arg + Lys): 9

Atomic composition:

Carbon	C	406
Hydrogen	H	633
Nitrogen	N	99
Oxygen	O	135
Sulfur	S	10

Formula: C₄₀₆H₆₃₃N₉₉O₁₃₅S₁₀
Total number of atoms: 1283

Extinction coefficients:

Extinction coefficients are in units of M⁻¹ cm⁻¹, at 280 nm measured in water.

Ext. coefficient 7365
Abs 0.1% (=1 g/l) 0.785, assuming all pairs of Cys residues form cystines

Ext. coefficient 6990
Abs 0.1% (=1 g/l) 0.745, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
 >20 hours (yeast, in vivo).
 >10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 39.43
This classifies the protein as stable.

Aliphatic index: 67.95

Grand average of hydropathicity (GRAVY): -0.263

Anexo 16

Resultados do software NetPhos para a
RnKazaPI.

NetPhos 3.1 Server - prediction results

Technical University of Denmark

```
>Sequence      176 amino acids
#
# netphos-3.1b prediction results
#
# Sequence      # x   Context      Score   Kinase      Answer
# -----
# Sequence      1 T   ----TRAAP    0.500   PKC         YES
# Sequence      1 T   ----TRAAP    0.453   GSK3        .
# Sequence      1 T   ----TRAAP    0.434   CaM-II      .
# Sequence      1 T   ----TRAAP    0.410   cdc2        .
# Sequence      1 T   ----TRAAP    0.376   CKI         .
# Sequence      1 T   ----TRAAP    0.349   DNAPK       .
# Sequence      1 T   ----TRAAP    0.329   PKG         .
# Sequence      1 T   ----TRAAP    0.308   CKII        .
# Sequence      1 T   ----TRAAP    0.297   p38MAPK     .
# Sequence      1 T   ----TRAAP    0.252   ATM         .
# Sequence      1 T   ----TRAAP    0.199   PKA         .
# Sequence      1 T   ----TRAAP    0.190   RSK         .
# Sequence      1 T   ----TRAAP    0.189   cdk5        .
# Sequence      1 T   ----TRAAP    0.101   PKB         .
# Sequence      1 T   ----TRAAP    0.039   unsp        .
#
# Sequence      19 T   MIPXTATIV    0.463   GSK3        .
# Sequence      19 T   MIPXTATIV    0.424   CaM-II      .
# Sequence      19 T   MIPXTATIV    0.400   cdc2        .
# Sequence      19 T   MIPXTATIV    0.376   CKI         .
# Sequence      19 T   MIPXTATIV    0.375   p38MAPK     .
# Sequence      19 T   MIPXTATIV    0.344   DNAPK       .
# Sequence      19 T   MIPXTATIV    0.341   CKII        .
# Sequence      19 T   MIPXTATIV    0.274   PKG         .
# Sequence      19 T   MIPXTATIV    0.252   ATM         .
# Sequence      19 T   MIPXTATIV    0.206   RSK         .
# Sequence      19 T   MIPXTATIV    0.195   cdk5        .
# Sequence      19 T   MIPXTATIV    0.173   PKA         .
# Sequence      19 T   MIPXTATIV    0.105   PKC         .
# Sequence      19 T   MIPXTATIV    0.092   PKB         .
# Sequence      19 T   MIPXTATIV    0.019   unsp        .
#
# Sequence      21 T   PXTATIVES    0.895   unsp        YES
# Sequence      21 T   PXTATIVES    0.456   CaM-II      .
# Sequence      21 T   PXTATIVES    0.425   cdc2        .
# Sequence      21 T   PXTATIVES    0.420   GSK3        .
# Sequence      21 T   PXTATIVES    0.416   CKII        .
# Sequence      21 T   PXTATIVES    0.362   CKI         .
# Sequence      21 T   PXTATIVES    0.345   DNAPK       .
# Sequence      21 T   PXTATIVES    0.322   p38MAPK     .
# Sequence      21 T   PXTATIVES    0.286   PKG         .
# Sequence      21 T   PXTATIVES    0.242   RSK         .
# Sequence      21 T   PXTATIVES    0.242   ATM         .
# Sequence      21 T   PXTATIVES    0.172   cdk5        .
# Sequence      21 T   PXTATIVES    0.159   PKA         .
# Sequence      21 T   PXTATIVES    0.142   PKC         .
```

# Sequence	21	T	PXTATIVES	0.095	PKB	.
#						
# Sequence	25	S	TIVESALIV	0.457	CaM-II	.
# Sequence	25	S	TIVESALIV	0.433	GSK3	.
# Sequence	25	S	TIVESALIV	0.419	PKG	.
# Sequence	25	S	TIVESALIV	0.382	cdc2	.
# Sequence	25	S	TIVESALIV	0.380	CKI	.
# Sequence	25	S	TIVESALIV	0.376	CKII	.
# Sequence	25	S	TIVESALIV	0.349	DNAPK	.
# Sequence	25	S	TIVESALIV	0.291	p38MAPK	.
# Sequence	25	S	TIVESALIV	0.268	RSK	.
# Sequence	25	S	TIVESALIV	0.265	ATM	.
# Sequence	25	S	TIVESALIV	0.176	cdk5	.
# Sequence	25	S	TIVESALIV	0.149	PKA	.
# Sequence	25	S	TIVESALIV	0.074	PKC	.
# Sequence	25	S	TIVESALIV	0.070	PKB	.
# Sequence	25	S	TIVESALIV	0.003	unsp	.
#						
# Sequence	32	Y	IVARYSECR	0.489	EGFR	.
# Sequence	32	Y	IVARYSECR	0.380	INSR	.
# Sequence	32	Y	IVARYSECR	0.332	SRC	.
# Sequence	32	Y	IVARYSECR	0.014	unsp	.
#						
# Sequence	33	S	VARYSECRE	0.993	unsp	YES
# Sequence	33	S	VARYSECRE	0.543	CKII	YES
# Sequence	33	S	VARYSECRE	0.446	CKI	.
# Sequence	33	S	VARYSECRE	0.433	GSK3	.
# Sequence	33	S	VARYSECRE	0.428	CaM-II	.
# Sequence	33	S	VARYSECRE	0.417	PKA	.
# Sequence	33	S	VARYSECRE	0.411	cdc2	.
# Sequence	33	S	VARYSECRE	0.387	RSK	.
# Sequence	33	S	VARYSECRE	0.343	DNAPK	.
# Sequence	33	S	VARYSECRE	0.286	p38MAPK	.
# Sequence	33	S	VARYSECRE	0.261	ATM	.
# Sequence	33	S	VARYSECRE	0.250	PKG	.
# Sequence	33	S	VARYSECRE	0.245	PKC	.
# Sequence	33	S	VARYSECRE	0.151	cdk5	.
# Sequence	33	S	VARYSECRE	0.074	PKB	.
#						
# Sequence	38	T	ECRETEDKA	0.937	unsp	YES
# Sequence	38	T	ECRETEDKA	0.510	CKII	YES
# Sequence	38	T	ECRETEDKA	0.467	cdc2	.
# Sequence	38	T	ECRETEDKA	0.431	GSK3	.
# Sequence	38	T	ECRETEDKA	0.428	CaM-II	.
# Sequence	38	T	ECRETEDKA	0.388	RSK	.
# Sequence	38	T	ECRETEDKA	0.363	CKI	.
# Sequence	38	T	ECRETEDKA	0.351	DNAPK	.
# Sequence	38	T	ECRETEDKA	0.326	p38MAPK	.
# Sequence	38	T	ECRETEDKA	0.300	PKG	.
# Sequence	38	T	ECRETEDKA	0.223	ATM	.
# Sequence	38	T	ECRETEDKA	0.200	PKC	.
# Sequence	38	T	ECRETEDKA	0.197	cdk5	.
# Sequence	38	T	ECRETEDKA	0.168	PKA	.
# Sequence	38	T	ECRETEDKA	0.086	PKB	.
#						
# Sequence	46	T	AXALYPEP	0.529	CKII	YES
# Sequence	46	T	AXALYPEP	0.468	GSK3	.
# Sequence	46	T	AXALYPEP	0.459	cdc2	.
# Sequence	46	T	AXALYPEP	0.440	PKG	.
# Sequence	46	T	AXALYPEP	0.427	CaM-II	.
# Sequence	46	T	AXALYPEP	0.389	CKI	.

# Sequence	46	T	AXALTYPEP	0.378	PKC	.
# Sequence	46	T	AXALTYPEP	0.362	DNAPK	.
# Sequence	46	T	AXALTYPEP	0.353	p38MAPK	.
# Sequence	46	T	AXALTYPEP	0.318	ATM	.
# Sequence	46	T	AXALTYPEP	0.284	unsp	.
# Sequence	46	T	AXALTYPEP	0.247	RSK	.
# Sequence	46	T	AXALTYPEP	0.211	cdk5	.
# Sequence	46	T	AXALTYPEP	0.161	PKA	.
# Sequence	46	T	AXALTYPEP	0.127	PKB	.
#						
# Sequence	47	Y	XALTYPEPR	0.398	EGFR	.
# Sequence	47	Y	XALTYPEPR	0.381	INSR	.
# Sequence	47	Y	XALTYPEPR	0.364	SRC	.
# Sequence	47	Y	XALTYPEPR	0.183	unsp	.
#						
# Sequence	53	T	EPRXTEINA	0.786	unsp	YES
# Sequence	53	T	EPRXTEINA	0.468	PKG	.
# Sequence	53	T	EPRXTEINA	0.448	GSK3	.
# Sequence	53	T	EPRXTEINA	0.438	CaM-II	.
# Sequence	53	T	EPRXTEINA	0.419	CKII	.
# Sequence	53	T	EPRXTEINA	0.409	cdc2	.
# Sequence	53	T	EPRXTEINA	0.388	CKI	.
# Sequence	53	T	EPRXTEINA	0.339	DNAPK	.
# Sequence	53	T	EPRXTEINA	0.315	RSK	.
# Sequence	53	T	EPRXTEINA	0.274	p38MAPK	.
# Sequence	53	T	EPRXTEINA	0.264	PKA	.
# Sequence	53	T	EPRXTEINA	0.256	ATM	.
# Sequence	53	T	EPRXTEINA	0.186	cdk5	.
# Sequence	53	T	EPRXTEINA	0.105	PKC	.
# Sequence	53	T	EPRXTEINA	0.083	PKB	.
#						
# Sequence	58	S	EINASEINH	0.903	unsp	YES
# Sequence	58	S	EINASEINH	0.486	CKII	.
# Sequence	58	S	EINASEINH	0.457	cdc2	.
# Sequence	58	S	EINASEINH	0.426	GSK3	.
# Sequence	58	S	EINASEINH	0.422	CaM-II	.
# Sequence	58	S	EINASEINH	0.402	CKI	.
# Sequence	58	S	EINASEINH	0.359	PKA	.
# Sequence	58	S	EINASEINH	0.349	DNAPK	.
# Sequence	58	S	EINASEINH	0.318	p38MAPK	.
# Sequence	58	S	EINASEINH	0.287	ATM	.
# Sequence	58	S	EINASEINH	0.286	RSK	.
# Sequence	58	S	EINASEINH	0.256	PKG	.
# Sequence	58	S	EINASEINH	0.165	cdk5	.
# Sequence	58	S	EINASEINH	0.147	PKC	.
# Sequence	58	S	EINASEINH	0.092	PKB	.
#						
# Sequence	66	T	HIXITXRXS	0.855	PKC	YES
# Sequence	66	T	HIXITXRXS	0.602	unsp	YES
# Sequence	66	T	HIXITXRXS	0.476	GSK3	.
# Sequence	66	T	HIXITXRXS	0.438	CaM-II	.
# Sequence	66	T	HIXITXRXS	0.399	cdc2	.
# Sequence	66	T	HIXITXRXS	0.360	CKI	.
# Sequence	66	T	HIXITXRXS	0.339	DNAPK	.
# Sequence	66	T	HIXITXRXS	0.316	p38MAPK	.
# Sequence	66	T	HIXITXRXS	0.306	PKG	.
# Sequence	66	T	HIXITXRXS	0.286	CKII	.
# Sequence	66	T	HIXITXRXS	0.276	ATM	.
# Sequence	66	T	HIXITXRXS	0.256	RSK	.
# Sequence	66	T	HIXITXRXS	0.204	cdk5	.
# Sequence	66	T	HIXITXRXS	0.110	PKA	.

# Sequence	66	T	HIXITXRXS	0.094	PKB	.
#						
# Sequence	70	S	TXRXSRHXD	0.978	unsp	YES
# Sequence	70	S	TXRXSRHXD	0.486	PKA	.
# Sequence	70	S	TXRXSRHXD	0.449	GSK3	.
# Sequence	70	S	TXRXSRHXD	0.432	CaM-II	.
# Sequence	70	S	TXRXSRHXD	0.418	PKG	.
# Sequence	70	S	TXRXSRHXD	0.392	cdc2	.
# Sequence	70	S	TXRXSRHXD	0.364	CKI	.
# Sequence	70	S	TXRXSRHXD	0.363	CKII	.
# Sequence	70	S	TXRXSRHXD	0.351	DNAPK	.
# Sequence	70	S	TXRXSRHXD	0.340	RSK	.
# Sequence	70	S	TXRXSRHXD	0.323	ATM	.
# Sequence	70	S	TXRXSRHXD	0.299	p38MAPK	.
# Sequence	70	S	TXRXSRHXD	0.195	PKC	.
# Sequence	70	S	TXRXSRHXD	0.189	cdk5	.
# Sequence	70	S	TXRXSRHXD	0.090	PKB	.
#						
# Sequence	78	S	DNIXSNEGL	0.500	CKII	YES
# Sequence	78	S	DNIXSNEGL	0.468	cdc2	.
# Sequence	78	S	DNIXSNEGL	0.464	GSK3	.
# Sequence	78	S	DNIXSNEGL	0.459	CaM-II	.
# Sequence	78	S	DNIXSNEGL	0.409	CKI	.
# Sequence	78	S	DNIXSNEGL	0.404	DNAPK	.
# Sequence	78	S	DNIXSNEGL	0.370	RSK	.
# Sequence	78	S	DNIXSNEGL	0.324	p38MAPK	.
# Sequence	78	S	DNIXSNEGL	0.304	ATM	.
# Sequence	78	S	DNIXSNEGL	0.296	unsp	.
# Sequence	78	S	DNIXSNEGL	0.225	PKG	.
# Sequence	78	S	DNIXSNEGL	0.198	cdk5	.
# Sequence	78	S	DNIXSNEGL	0.157	PKA	.
# Sequence	78	S	DNIXSNEGL	0.087	PKB	.
# Sequence	78	S	DNIXSNEGL	0.051	PKC	.
#						
# Sequence	85	T	GLECTXSXX	0.503	CKI	YES
# Sequence	85	T	GLECTXSXX	0.478	cdc2	.
# Sequence	85	T	GLECTXSXX	0.466	GSK3	.
# Sequence	85	T	GLECTXSXX	0.458	CKII	.
# Sequence	85	T	GLECTXSXX	0.446	CaM-II	.
# Sequence	85	T	GLECTXSXX	0.368	p38MAPK	.
# Sequence	85	T	GLECTXSXX	0.352	DNAPK	.
# Sequence	85	T	GLECTXSXX	0.282	PKG	.
# Sequence	85	T	GLECTXSXX	0.278	ATM	.
# Sequence	85	T	GLECTXSXX	0.254	RSK	.
# Sequence	85	T	GLECTXSXX	0.214	cdk5	.
# Sequence	85	T	GLECTXSXX	0.131	PKA	.
# Sequence	85	T	GLECTXSXX	0.126	PKC	.
# Sequence	85	T	GLECTXSXX	0.093	PKB	.
# Sequence	85	T	GLECTXSXX	0.058	unsp	.
#						
# Sequence	87	S	ECTXSXXPE	0.469	cdc2	.
# Sequence	87	S	ECTXSXXPE	0.467	CaM-II	.
# Sequence	87	S	ECTXSXXPE	0.428	GSK3	.
# Sequence	87	S	ECTXSXXPE	0.381	DNAPK	.
# Sequence	87	S	ECTXSXXPE	0.362	CKI	.
# Sequence	87	S	ECTXSXXPE	0.355	CKII	.
# Sequence	87	S	ECTXSXXPE	0.313	ATM	.
# Sequence	87	S	ECTXSXXPE	0.302	p38MAPK	.
# Sequence	87	S	ECTXSXXPE	0.295	RSK	.
# Sequence	87	S	ECTXSXXPE	0.253	PKG	.
# Sequence	87	S	ECTXSXXPE	0.239	PKA	.

# Sequence	87	S	ECTXSXXPE	0.154	PKC	.
# Sequence	87	S	ECTXSXXPE	0.153	cdk5	.
# Sequence	87	S	ECTXSXXPE	0.111	PKB	.
# Sequence	87	S	ECTXSXXPE	0.024	unsp	.
#						
# Sequence	92	S	XXPESVMKL	0.733	unsp	YES
# Sequence	92	S	XXPESVMKL	0.477	cdc2	.
# Sequence	92	S	XXPESVMKL	0.463	GSK3	.
# Sequence	92	S	XXPESVMKL	0.449	CaM-II	.
# Sequence	92	S	XXPESVMKL	0.445	RSK	.
# Sequence	92	S	XXPESVMKL	0.412	PKC	.
# Sequence	92	S	XXPESVMKL	0.377	p38MAPK	.
# Sequence	92	S	XXPESVMKL	0.363	CKI	.
# Sequence	92	S	XXPESVMKL	0.350	DNAPK	.
# Sequence	92	S	XXPESVMKL	0.327	CKII	.
# Sequence	92	S	XXPESVMKL	0.321	ATM	.
# Sequence	92	S	XXPESVMKL	0.228	cdk5	.
# Sequence	92	S	XXPESVMKL	0.208	PKG	.
# Sequence	92	S	XXPESVMKL	0.188	PKA	.
# Sequence	92	S	XXPESVMKL	0.089	PKB	.
#						
# Sequence	97	S	VMKLSSSVL	0.658	PKA	YES
# Sequence	97	S	VMKLSSSVL	0.450	cdc2	.
# Sequence	97	S	VMKLSSSVL	0.438	GSK3	.
# Sequence	97	S	VMKLSSSVL	0.424	CaM-II	.
# Sequence	97	S	VMKLSSSVL	0.407	RSK	.
# Sequence	97	S	VMKLSSSVL	0.400	PKG	.
# Sequence	97	S	VMKLSSSVL	0.363	CKI	.
# Sequence	97	S	VMKLSSSVL	0.348	DNAPK	.
# Sequence	97	S	VMKLSSSVL	0.326	CKII	.
# Sequence	97	S	VMKLSSSVL	0.304	p38MAPK	.
# Sequence	97	S	VMKLSSSVL	0.304	ATM	.
# Sequence	97	S	VMKLSSSVL	0.266	PKC	.
# Sequence	97	S	VMKLSSSVL	0.176	unsp	.
# Sequence	97	S	VMKLSSSVL	0.150	cdk5	.
# Sequence	97	S	VMKLSSSVL	0.082	PKB	.
#						
# Sequence	98	S	MKLSSSVLS	0.451	GSK3	.
# Sequence	98	S	MKLSSSVLS	0.425	CaM-II	.
# Sequence	98	S	MKLSSSVLS	0.417	cdc2	.
# Sequence	98	S	MKLSSSVLS	0.364	CKI	.
# Sequence	98	S	MKLSSSVLS	0.361	p38MAPK	.
# Sequence	98	S	MKLSSSVLS	0.343	DNAPK	.
# Sequence	98	S	MKLSSSVLS	0.308	PKA	.
# Sequence	98	S	MKLSSSVLS	0.290	PKC	.
# Sequence	98	S	MKLSSSVLS	0.285	CKII	.
# Sequence	98	S	MKLSSSVLS	0.280	ATM	.
# Sequence	98	S	MKLSSSVLS	0.276	RSK	.
# Sequence	98	S	MKLSSSVLS	0.265	PKG	.
# Sequence	98	S	MKLSSSVLS	0.188	cdk5	.
# Sequence	98	S	MKLSSSVLS	0.101	PKB	.
# Sequence	98	S	MKLSSSVLS	0.028	unsp	.
#						
# Sequence	99	S	KLSSSVLSV	0.576	cdc2	YES
# Sequence	99	S	KLSSSVLSV	0.569	PKC	YES
# Sequence	99	S	KLSSSVLSV	0.469	CaM-II	.
# Sequence	99	S	KLSSSVLSV	0.441	GSK3	.
# Sequence	99	S	KLSSSVLSV	0.368	CKI	.
# Sequence	99	S	KLSSSVLSV	0.346	DNAPK	.
# Sequence	99	S	KLSSSVLSV	0.311	p38MAPK	.
# Sequence	99	S	KLSSSVLSV	0.298	RSK	.

# Sequence	99	S	KLSSSVLSV	0.283	ATM	.
# Sequence	99	S	KLSSSVLSV	0.279	PKA	.
# Sequence	99	S	KLSSSVLSV	0.258	CKII	.
# Sequence	99	S	KLSSSVLSV	0.233	unsp	.
# Sequence	99	S	KLSSSVLSV	0.220	PKG	.
# Sequence	99	S	KLSSSVLSV	0.177	cdk5	.
# Sequence	99	S	KLSSSVLSV	0.082	PKB	.
#						
# Sequence	102	S	SSVLSVVVI	0.564	unsp	YES
# Sequence	102	S	SSVLSVVVI	0.445	CaM-II	.
# Sequence	102	S	SSVLSVVVI	0.436	GSK3	.
# Sequence	102	S	SSVLSVVVI	0.422	cdc2	.
# Sequence	102	S	SSVLSVVVI	0.391	p38MAPK	.
# Sequence	102	S	SSVLSVVVI	0.367	CKI	.
# Sequence	102	S	SSVLSVVVI	0.346	DNAPK	.
# Sequence	102	S	SSVLSVVVI	0.338	RSK	.
# Sequence	102	S	SSVLSVVVI	0.308	CKII	.
# Sequence	102	S	SSVLSVVVI	0.292	ATM	.
# Sequence	102	S	SSVLSVVVI	0.287	PKC	.
# Sequence	102	S	SSVLSVVVI	0.243	PKG	.
# Sequence	102	S	SSVLSVVVI	0.195	cdk5	.
# Sequence	102	S	SSVLSVVVI	0.193	PKA	.
# Sequence	102	S	SSVLSVVVI	0.073	PKB	.
#						
# Sequence	126	T	VECNTDCGD	0.609	CKII	YES
# Sequence	126	T	VECNTDCGD	0.438	GSK3	.
# Sequence	126	T	VECNTDCGD	0.437	CKI	.
# Sequence	126	T	VECNTDCGD	0.425	CaM-II	.
# Sequence	126	T	VECNTDCGD	0.387	cdc2	.
# Sequence	126	T	VECNTDCGD	0.386	p38MAPK	.
# Sequence	126	T	VECNTDCGD	0.364	unsp	.
# Sequence	126	T	VECNTDCGD	0.341	DNAPK	.
# Sequence	126	T	VECNTDCGD	0.254	ATM	.
# Sequence	126	T	VECNTDCGD	0.227	PKG	.
# Sequence	126	T	VECNTDCGD	0.214	RSK	.
# Sequence	126	T	VECNTDCGD	0.160	cdk5	.
# Sequence	126	T	VECNTDCGD	0.128	PKC	.
# Sequence	126	T	VECNTDCGD	0.095	PKA	.
# Sequence	126	T	VECNTDCGD	0.087	PKB	.
#						
# Sequence	132	Y	CGDKYEPVC	0.963	unsp	YES
# Sequence	132	Y	CGDKYEPVC	0.534	INSR	YES
# Sequence	132	Y	CGDKYEPVC	0.462	SRC	.
# Sequence	132	Y	CGDKYEPVC	0.345	EGFR	.
#						
# Sequence	140	S	CAQPSDEEK	0.972	unsp	YES
# Sequence	140	S	CAQPSDEEK	0.663	CKII	YES
# Sequence	140	S	CAQPSDEEK	0.479	PKG	.
# Sequence	140	S	CAQPSDEEK	0.436	GSK3	.
# Sequence	140	S	CAQPSDEEK	0.429	CaM-II	.
# Sequence	140	S	CAQPSDEEK	0.408	CKI	.
# Sequence	140	S	CAQPSDEEK	0.393	cdc2	.
# Sequence	140	S	CAQPSDEEK	0.383	ATM	.
# Sequence	140	S	CAQPSDEEK	0.353	DNAPK	.
# Sequence	140	S	CAQPSDEEK	0.321	p38MAPK	.
# Sequence	140	S	CAQPSDEEK	0.273	RSK	.
# Sequence	140	S	CAQPSDEEK	0.159	cdk5	.
# Sequence	140	S	CAQPSDEEK	0.116	PKA	.
# Sequence	140	S	CAQPSDEEK	0.088	PKC	.
# Sequence	140	S	CAQPSDEEK	0.080	PKB	.
#						

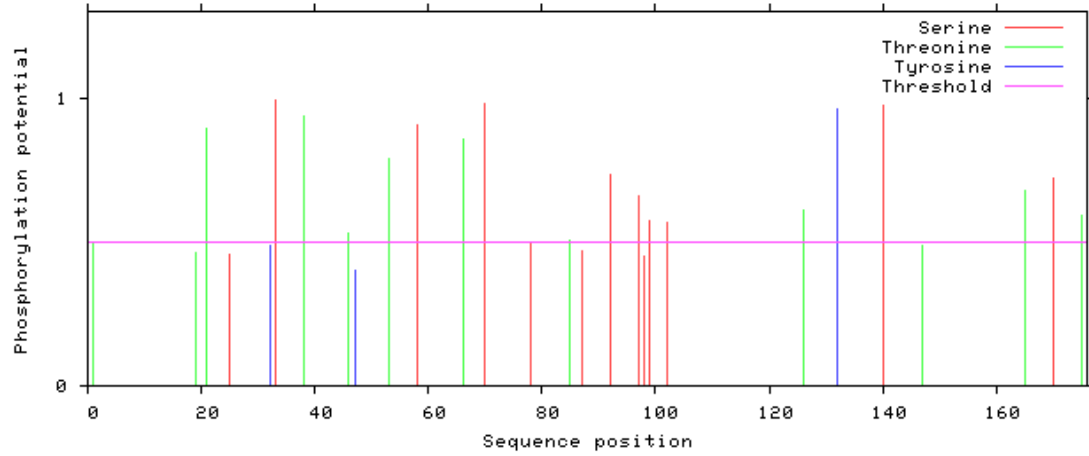
# Sequence	147	T	EKFETFKNL	0.490	CKI	.
# Sequence	147	T	EKFETFKNL	0.456	CKII	.
# Sequence	147	T	EKFETFKNL	0.439	CaM-II	.
# Sequence	147	T	EKFETFKNL	0.430	GSK3	.
# Sequence	147	T	EKFETFKNL	0.420	unsp	.
# Sequence	147	T	EKFETFKNL	0.382	cdc2	.
# Sequence	147	T	EKFETFKNL	0.347	p38MAPK	.
# Sequence	147	T	EKFETFKNL	0.342	DNAPK	.
# Sequence	147	T	EKFETFKNL	0.321	PKC	.
# Sequence	147	T	EKFETFKNL	0.277	RSK	.
# Sequence	147	T	EKFETFKNL	0.253	PKG	.
# Sequence	147	T	EKFETFKNL	0.246	ATM	.
# Sequence	147	T	EKFETFKNL	0.169	PKA	.
# Sequence	147	T	EKFETFKNL	0.161	cdk5	.
# Sequence	147	T	EKFETFKNL	0.103	PKB	.
#						
# Sequence	165	T	NDEKTWKFV	0.679	PKC	YES
# Sequence	165	T	NDEKTWKFV	0.512	unsp	YES
# Sequence	165	T	NDEKTWKFV	0.434	GSK3	.
# Sequence	165	T	NDEKTWKFV	0.428	CaM-II	.
# Sequence	165	T	NDEKTWKFV	0.413	CKII	.
# Sequence	165	T	NDEKTWKFV	0.390	CKI	.
# Sequence	165	T	NDEKTWKFV	0.340	DNAPK	.
# Sequence	165	T	NDEKTWKFV	0.311	cdc2	.
# Sequence	165	T	NDEKTWKFV	0.296	p38MAPK	.
# Sequence	165	T	NDEKTWKFV	0.263	PKG	.
# Sequence	165	T	NDEKTWKFV	0.243	ATM	.
# Sequence	165	T	NDEKTWKFV	0.201	RSK	.
# Sequence	165	T	NDEKTWKFV	0.160	cdk5	.
# Sequence	165	T	NDEKTWKFV	0.100	PKB	.
# Sequence	165	T	NDEKTWKFV	0.083	PKA	.
#						
# Sequence	170	S	WKVSEGEC	0.719	unsp	YES
# Sequence	170	S	WKVSEGEC	0.625	CKII	YES
# Sequence	170	S	WKVSEGEC	0.468	CKI	.
# Sequence	170	S	WKVSEGEC	0.439	CaM-II	.
# Sequence	170	S	WKVSEGEC	0.432	GSK3	.
# Sequence	170	S	WKVSEGEC	0.419	PKA	.
# Sequence	170	S	WKVSEGEC	0.353	cdc2	.
# Sequence	170	S	WKVSEGEC	0.341	DNAPK	.
# Sequence	170	S	WKVSEGEC	0.302	RSK	.
# Sequence	170	S	WKVSEGEC	0.292	ATM	.
# Sequence	170	S	WKVSEGEC	0.275	PKG	.
# Sequence	170	S	WKVSEGEC	0.272	p38MAPK	.
# Sequence	170	S	WKVSEGEC	0.156	cdk5	.
# Sequence	170	S	WKVSEGEC	0.138	PKC	.
# Sequence	170	S	WKVSEGEC	0.091	PKB	.
#						
# Sequence	175	T	EGECTD---	0.595	CKII	YES
# Sequence	175	T	EGECTD---	0.448	GSK3	.
# Sequence	175	T	EGECTD---	0.415	CaM-II	.
# Sequence	175	T	EGECTD---	0.372	p38MAPK	.
# Sequence	175	T	EGECTD---	0.370	cdc2	.
# Sequence	175	T	EGECTD---	0.364	CKI	.
# Sequence	175	T	EGECTD---	0.341	DNAPK	.
# Sequence	175	T	EGECTD---	0.255	ATM	.
# Sequence	175	T	EGECTD---	0.253	PKG	.
# Sequence	175	T	EGECTD---	0.220	RSK	.
# Sequence	175	T	EGECTD---	0.161	cdk5	.
# Sequence	175	T	EGECTD---	0.126	PKA	.
# Sequence	175	T	EGECTD---	0.087	PKB	.


```

# Sequence          175 T   EGETD---   0.069   PKC       .
# Sequence          175 T   EGETD---   0.063   unsp      .
#
TRAAPWXAAPWXHEMIPXTATIVESALIVARYSECRETEDKAXALTYPEP # 50
RXTEINASEINHIKITXRXSRLHXDNIXSNEGLECTXSXXPESVMKLSSSV # 100
LSVVVIAVFLCMLDKHVFVGDEVECNTPDCGDKYEPVCAQPSDEEKFETFKN # 150
LCELDKAMKCNDEKTKWKFVSEGETD # 200
%1 T.....T.....S.....T..... # 50
%1 ..T...S...T..S.....T.....S...S.S. # 100
%1 .S.....T.....Y.....S..... # 150
%1 .....T.....S.....T.

```

NetPhos 3.1a: predicted phosphorylation sites in Sequence



Anexo 17

Resultados do software NetNGlyc para a
RnKazalPI.

Anexo 18

Resultados do software NetOGlyc para a
RnKazalPI.



NetOGlyc 4.0 Server - prediction results

Technical University of Denmark

```
##gff-version 2
##source-version NetOGlyc 4.0.0.13
##date 18-8-6
##Type Protein
#seqname      source  feature start  end    score  strand  frame  comment
RNKAZALPI    netOGlyc-4.0.0.13  CARBOHYD  4      4      0.0613366  .      .
RNKAZALPI    netOGlyc-4.0.0.13  CARBOHYD  5      5      0.0370443  .      .
RNKAZALPI    netOGlyc-4.0.0.13  CARBOHYD  6      6      0.0419188  .      .
RNKAZALPI    netOGlyc-4.0.0.13  CARBOHYD  9      9      0.0437116  .      .
RNKAZALPI    netOGlyc-4.0.0.13  CARBOHYD  33     33     0.224702   .      .
RNKAZALPI    netOGlyc-4.0.0.13  CARBOHYD  47     47     0.303156   .      .
RNKAZALPI    netOGlyc-4.0.0.13  CARBOHYD  54     54     0.194413   .      .
RNKAZALPI    netOGlyc-4.0.0.13  CARBOHYD  72     72     0.0816562  .      .
RNKAZALPI    netOGlyc-4.0.0.13  CARBOHYD  77     77     0.155466   .      .
RNKAZALPI    netOGlyc-4.0.0.13  CARBOHYD  82     82     0.00981173 .      .
```

[Explain](#) the output. Go [back](#).

Anexo 19

Resultados do software Protein Blast para
a RnKazalPI

BLAST Results

[Questions/comments](#)

Job title: RnKazalPI

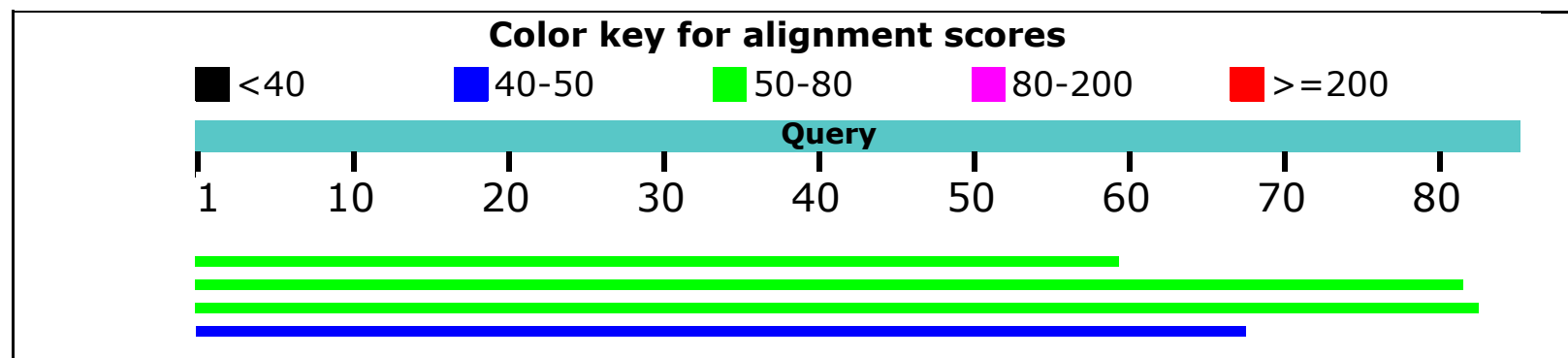
RID [PF0GTKWA014](#) (Expires on 08-07 07:02 am)

Query ID lcl|Query_111758
Description RnKazalPI
Molecule type amino acid
Query Length 83

Database Name nr
Description All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF excluding environmental samples from WGS projects
Program BLASTP 2.8.0+

New Analyze your query with [SmartBLAST](#) **Graphic Summary**

No putative conserved domains have been detected

Distribution of the top 4 Blast Hits on 4 subject sequences

Descriptions

Sequences producing significant alignments:

Description	Max score	Total score	Query cover	E value	Ident	Accession
Pc266, similar to salivary kazal-type proteinase inhibitor [Panstrongylus chinai]	58.2	58.2	71%	1e-14	48%	BBA30668.1
salivary secreted kazal-type proteinase inhibitor [Triatoma infestans]	56.2	56.2	97%	7e-14	46%	ABR27937.1
salivary kazal-type proteinase inhibitor [Triatoma infestans]	53.1	53.1	98%	1e-12	42%	ABR27896.1
Pc77, similar to salivary kazal-type proteinase inhibitor [Panstrongylus chinai]	47.8	47.8	80%	1e-10	47%	BBA30649.1

Alignments

Pc266, similar to salivary kazal-type proteinase inhibitor, partial [Panstrongylus chinai]

Sequence ID: **BBA30668.1** Length: 73 Number of Matches: 1

Range 1: 15 to 73

Score	Expect	Method	Identities	Positives	Gaps	Frame
58.2 bits(139)	1e-14()	Compositional matrix adjust.	29/61(48%)	42/61(68%)	4/61(6%)	

Features:

Query	1	MKLSVVLSVAVF--LCMLDKHVFGEVEECNTDCGDKYEPVCAQPSDEEKFETFKNL	58
Sbjct	15	MTLTTALS SVVIAVAV +CMLDK V+ +++C C +Y+PVCAQ +E F+ F N+	72
Query	59	C	59
Sbjct	73	C	73

salivary secreted kazal-type proteinase inhibitor [Triatoma infestans]

Sequence ID: **ABR27937.1** Length: 80 Number of Matches: 1

Range 1: 4 to 77

Score	Expect	Method	Identities	Positives	Gaps	Frame
56.2 bits(134)	7e-14()	Compositional matrix adjust.	37/81(46%)	46/81(56%)	7/81(8%)	

Features:

Query	1	MKLSVVLSVAVFLCMLDKHVFGEVEECNTDCGDKYEPVCAQPSDEEKFETFKNLCE	60
Sbjct	4	MKLS+SVL VVVIAVFL M+D FG E CN C + VC Q E++ TF N C+	57
Query	61	LDKAMKCNDEKTWKFVSEGE	81
Sbjct	58	IEKENKCGG-GGWTIRNRGHC	77

salivary kazal-type proteinase inhibitor [Triatoma infestans]

Sequence ID: **ABR27896.1** Length: 81 Number of Matches: 1

Range 1: 3 to 80

Score	Expect	Method	Identities	Positives	Gaps	Frame
53.1 bits(126)	1e-12()	Compositional matrix adjust.	35/84(42%)	44/84(52%)	8/84(9%)	

Features:

Query	1	MKLSVVLSVAVFL--CMLDKHVFGEVEECNTDCGDKYEPVCAQPSDEEKFETFKNL	58
Sbjct	3	MKLS+S SVVIAVFL CM D FG+ C C + VC Q +E TF +L	57
Query	59	CELDKAMKCNDEKTWKFVSEGE	82
Sbjct	58	CEMEKENLCGSGG-WTKLKNHGCS	80

Pc77, similar to salivary kazal-type proteinase inhibitor [Panstrongylus chinai]

Sequence ID: **BBA30649.1** Length: 78 Number of Matches: 1

Range 1: 1 to 65

Score	Expect	Method	Identities	Positives	Gaps	Frame
47.8 bits(112)	1e-10()	Compositional matrix adjust.	33/70(47%)	41/70(58%)	8/70(11%)	

Features:

Query	1	MKLSSSVLSVVVIAVFL--CMLDKHVFGEVEECNTDC-GDKYEPVCAQPSDEEKFFETFKN	57
		MKLS+SVLSV+VIAVFL CM D F + + C+ C PVC Q S E+K TF +	
Sbjct	1	MKLSTSVLSVLVIAVFLTVCMFD---FSEAI SCSEQCPWGAGVPVCGQRSMEKK--TFTS	55
Query	58	LCELDKAMKC 67	
		C D +C	
Sbjct	56	KCHFDYYNQC 65	

BLAST is a registered trademark of the National Library of Medicine



[Support center Mailing list](#)

[YouTube](#)



• [National Library Of Medicine](#)



• [National Institutes Of Health](#)



• [U.S. Department of Health & Human Services](#)



• [USA.gov](#)

NCBI

[National Center for Biotechnology Information](#), [U.S. National Library of Medicine](#) 8600 Rockville Pike, Bethesda MD, 20894 USA
[Policies and Guidelines](#) | [Contact](#)