



UNIVERSIDADE FEDERAL DA BAHIA
FACULDADE DE MEDICINA DA BAHIA
PROGRAMA DE PÓS-GRADUAÇÃO EM
MEDICINA E SAÚDE



LUCIANA MOURA DE ASSIS

PREDIÇÃO DE EPITOPOS DE CÉLULA B EM PROTEÍNAS DE
Leishmania infantum*: UMA ANÁLISE *IN SILICO

TESE DE DOUTORADO

Salvador - Bahia

2013

LUCIANA MOURA DE ASSIS

**PREDIÇÃO DE EPITOPOS DE CÉLULA B EM PROTEÍNAS DE
Leishmania infantum: UMA ANÁLISE *IN SILICO***

Tese apresentada ao Programa de Pós-graduação em Medicina e Saúde, da Faculdade de Medicina da Bahia, Universidade Federal da Bahia, como requisito para obtenção de grau de Doutora em Medicina e Saúde.

Orientador: Prof. Dr. Edgar Marcelino de Carvalho Filho

Co-orientadora: Prof.^a Dr.^a Márcia Almeida de Melo

Salvador - Bahia

2013

FICHA CATALOGRÁFICA

A848p Assis, Luciana Moura de
Predição de epitopos de célula B em proteínas de
Leishmania infantum: uma análise *in silico* / Luciana
Moura de Assis. Bahia, 2013.
65f. : il.

Bibliografia.

Orientador: Edgar Marcelino de Carvalho Filho.

Co-orientadora: Márcia Almeida de Melo

Tese(Doutorado)Universidade Federal da Bahia,2013.

1. Biologia Molecular. 2. Bioinformática. 3. Biologia
Sintética. I. Carvalho Filho, Edgar Marcelino de. II.
Melo, Márcia Almeida de. III. Universidade Federal
da Bahia. IV. Título.

UFCG /CFP/BS

CDU – 577.2

COMISSÃO EXAMINADORA

Prof. Dr. Argemiro D'Oliveira Junior
Universidade Federal da Bahia (UFBA)

Prof.^a Dr.^a Rita de Cássia Saldanha Lucena
Universidade Federal da Bahia (UFBA)

Prof. Dr. Roque Aras Junior
Universidade Federal da Bahia (UFBA)

Prof.^a Dr.^a Renata de Souza Coelho Soares
Universidade Estadual da Paraíba (UEPB)

Prof. Dr. Suedney de Lima Silva
Universidade Federal da Paraíba (UFPB)

Dedico esta tese ao meu filho Lucas Davi, ao meu marido João Rodrigues e a minha grande família (pai, mãe, avó, irmãos, irmãs, sobrinhos e sobrinhas).

Amo vocês!

AGRADECIMENTOS

A DEUS, minha paz, minha perseverança em tudo que busco fazer nesta vida. Obrigada Pai, pela bonança, após superar as tempestades que são inerentes à vida de todos nós.

Ao Dr. Edgar Marcelino de Carvalho Filho, pela confiança no meu trabalho, disponibilidade em me orientar e propiciar as condições necessárias à realização deste trabalho acadêmico.

A Dra. Márcia Almeida de Melo que foi mais do que minha co-orientadora, uma amiga, muito obrigada pela prontidão nas orientações, pela tranquilidade em repassar as informações, pelo profissionalismo, pelo exemplo de simplicidade, de dedicação e de respeito ao próximo.

Ao Dr. Paulo Paes de Andrade pelas valiosas sugestões a este trabalho. Você é um exemplo de pessoa que reúne competência, simplicidade e serenidade. Obrigada por você e sua esposa (Márcia Melo) terem, gentilmente, acolhido a mim e ao meu marido em sua residência (em Recife), onde vivi momentos agradáveis, inclusive pela presença do hilário *Fau* que alegrava sempre!

Ao meu adorável filho Lucas Davi, minha vida, inspiração e força para continuar vencendo desafios. Obrigada filho, por alegrar meus dias, por estar presente em minha vida.

Ao meu amado esposo João Rodrigues de Sousa, pela cumplicidade, compreensão, paciência e pelo exercício diário para tornar possível a realização deste trabalho. Obrigada por fazer parte da minha vida! Eu o amo!

Aos meus pais Francisco Cavalcante de Assis (Ica) e Iracema Moura de Assis, meu alicerce, sou eternamente grata pelos cuidados e amor incondicional;

À minha amada avó pelo carinho. Aos meus irmãos José Nildo, Edileuso e minhas irmãs Zélia, Deusaide, Suênia e Sumály pelo apoio e zelo contínuo. Agradeço também a todos os outros membros da minha família que me apoiaram e torceram por mim.

À Débora, minha cunhada, obrigada pelo carinho e correções no manuscrito.

Ao meu sogro Francisco Rodrigues de Sousa (*in memoriam*) e minha sogra Maria Oliveira de Sousa, pelo carinho e atenção dispensados.

Aos colegas Deborah Rose, Marília Medeiros, Graça Loureiro, Lourdes Queiroga, Rosineide Torres, Gesira Florentino, Consuelo Padilha, Marta Barreto, Lurdes Campos, Raimunda Neves, Betânia Maria, Marilena Maria, Homero Rodrigues, Gerson Bragagnoli e Luciano Holanda, pelo harmonioso convívio durante o curso.

Em especial, agradeço aos amigos(as) Erlane Aguiar, Lúcia Oliveira, Abrão Amério e José Rómulo, pelo carinho, companheirismo e atenção dispensada durante todo o caminho do curso. E à amiga Rosália Severo, pelo valioso apoio no início do curso de doutorado.

Também sou grata a Dona Zefinha (*in memoriam*), que mesmo sem me conhecer rezou por mim quando passei momentos difíceis.

As Professoras Edineide Nunes e Veruscka Pedrosa e ao Prof. Vinícius Ximenes, pelo carinho e apoio dispensados. E Ao prof. Dr. José Cezário de Almeida e a Prof.^a Dr.^a Francisca Bezerra, pelo estímulo dado para iniciar o curso de doutorado.

Aos colegas do laboratório de Biologia Molecular do Semi-árido/Patos-PB, Aline Antas, Nedja Fernanda, Tereza Rotondano, Vanessa Lira, Gilzane Dantes, Marília Andreza, Daniele Luna, Herta, Rodrigo Mendes, Expedito Camboim e Eduardo, pela prazerosa convivência e toda ajuda oferecida.

À Universidade Federal da Bahia (UFBA) e a Universidade Federal de Campina Grande (UFCG), por tornar possível a realização deste Doutorado.

Aos alunos da pós-graduação e funcionários do Serviço de imunologia do Complexo Hospitalar Universitário Professor Edgard Santos, UFBA, pela acolhida e apoio durante o período de atividades em Salvador-Ba.

Aos professores e servidores técnico-administrativos da UFCG e UFBA que participaram da realização deste Doutorado.

Aos coordenadores do DINTER UFBA/UFCG, Dra. Helma Cotrim, Dr. Adelmir Machado, Dr. Patrício Marques e Dra. Teresa Nascimento, pela dedicação dispensada ao DINTER.

A CAPES, pelo apoio financeiro dado.

À banca examinadora pelas sugestões e correções propostas a este trabalho.

A todos que contribuíram, mesmo que indiretamente, para o caminho desta obra denominada tese de doutorado.

Porque todos aqueles que pedem recebem; aqueles que procuram acham; e a porta será aberta para quem bate.

Lucas 11, 9-10

SUMÁRIO

Lista de tabelas e figuras	9
Lista de abreviaturas e siglas	10
Resumo	11
Abstract	12
1. Introdução	13
2. Objetivos	15
2.1. Geral	15
2.2. Específicos	15
3. Revisão de literatura	16
3.1. Artigo de revisão – Ferramentas de bioinformática na predição de epitopos de célula B	17
4. Resultados	38
4.1. Artigo original - Epitopos de proteínas antigênicas de <i>Leishmania</i> <i>infantum</i> : uma análise <i>in silico</i>	39
5. Conclusões	58
6. Considerações finais	59
7. Perspectivas de estudo	60
8. Referências	61
9. Anexos	62
9.1. Anexo A – Comprovante de submissão de artigo a periódico	62
9.2. Anexo B – Instruções para autores e normas de publicação de periódico cujo artigo encontra-se em apreciação para publicação.....	65

LISTA DE TABELAS E FIGURAS

		Pag.
Tabela 1	Quantidade de epitopos lineares preditos por cada programa a partir das proteínas de <i>L. infantum</i> cisteína peptidase (CP), redutase tiol dependente1 (TDR1) e proteína de choque térmico (HSP70)	45
Tabela 2	Epitopos lineares da proteína TDR1 de <i>L. infantum</i> preditos pelo IEDB, BepiPred e BcePred.	45
Tabela 3	Epitopos lineares da proteína HSP70 de <i>L. infantum</i> preditos pelos programas IEDB, BepiPred e BcePred	46
Tabela 4	Epitopos conformacionais de célula B da proteína TDR1 (GI:146097469) de <i>L. infantum</i> , preditos pelo CBtope.	47
Tabela 5	Epitopos conformacionais de célula B da proteína HSP70 (GI:758136) de <i>L. infantum</i> , preditos pelo CBtope.	47
Tabela 6	Sinopse dos epitopos lineares e conformacionais que se sobrepõem ou não na proteína HSP70 de <i>Leishmania infantum</i> .	48
Figura 1	Frequência dos tipos de estrutura secundária predita pelo programa PHD a partir das proteínas CP, TDR1 e HSP70 de <i>L. infantum</i> .	48
Figura 2	Localização do epitopo imunodominante H17 (Quijada et al., 1996) destacado em amarelo e a região da estrutura secundária correspondente. Hh (α - hélice), Cc (randômica).	49
Figura 3	Sobreposição dos epitopos lineares preditos pelos programas IEDB, BepiPred e BcePred na estrutura secundária da TDR1 predita pelo PHD. Hh (α - hélice), Cc (randômica), Ee (fita estendida).	49
Figura 4	Sobreposição dos epitopos conformacionais preditos pelo CBTope na estrutura secundária da TDR1 predita pelo PHD. Hh (α - hélice), Cc (randômica), Ee (fita estendida).	50
Figura 5	Frequência dos resíduos de aminoácidos dos epitopos lineares das proteínas CP, TDR1 e HSP70 preditos pelos programas IEDB, BepiPred e BcePred.	50

LISTA DE ABREVIATURAS E SIGLAS

ABCpred	Predição de epítipo de célula B baseado em Rede Neural Artificial
BcePred	Predição de epítopos contínuos de células B em sequências antigênicas usando propriedades físico-químicas
BCR	Receptor de célula B
BepiPred	Predição de epítipo de célula B usando uma combinação de um modelo oculto de Markov e métodos de escala de propensão
BEST	Ferramenta de predição de epítipo de célula B usando vetores de suporte de máquinas
CBtope	Predição de epítipo de célula B Conformacional
BLAST	Basic Local Alignment Search Tool
CEP	Predição de Epítipo de célula B Conformacional e linear baseado em estrutura 3D de proteína
ELISA	Ensaio Imunoenzimático - <i>Enzyme-Linked Immunosorbent Assay</i>
GeneDb	Banco de Dados de Genes
HPLC	Cromatografia líquida de alta performance
HSP70	Proteína de Choque Térmico de 70kDa - <i>Heat Shock Proteins</i> (HSP)
IEDB	Banco de Dados de Epítopos Imune e Recursos de Análises - <i>Immune Epitope Database and Analyses Resource</i>
LV	Leishmaniose Visceral
PHD	Servidor de predição de estrutura secundária
RIFI	Reação de Imunofluorescência Indireta

RESUMO

PREDIÇÃO DE EPITOPOS DE CÉLULA B EM PROTEÍNAS DE *Leishmania infantum*: UMA ANÁLISE *IN SILICO*

A Leishmaniose visceral (LV) é uma doença crônica, endêmica em 62 países e representa um sério problema de saúde pública no Brasil. Os testes sorodiagnósticos convencionais empregam antígenos inteiros ou extratos solúveis que limitam a padronização do antígeno, e podem gerar reações cruzadas com outras doenças. Um método alternativo é o uso de peptídeos a partir de epitopos de célula B identificados através de ferramentas de bioinformática. **Objetivou-se** identificar epitopos lineares e conformacionais de célula B das proteínas de *Leishmania infantum* cisteína peptidase calpaina-like, redutase thiol dependente 1 (TDR1) e HSP70, bem como identificar sua estrutura secundária através de metodologia *in silico*; em seguida, buscou-se selecionar os epitopos lineares comuns aos diferentes métodos de predição para verificar a composição dos resíduos de aminoácidos dos mesmos. **Metodologia:** As ferramentas de bioinformática IEDB, BepiPred e BcePred foram usadas para predição de epitopos lineares de célula B e o programa CBtope para predição de epitopos conformacionais. A estrutura secundária das proteínas foi predita pelo servidor PHD. **Resultados:** As análises de predição produziram um total de 148 epitopos lineares e 164 epitopos conformacionais a partir das três proteínas, a maioria desses epitopos está localizada na mesma região. A estrutura secundária das proteínas é composta por α -hélice, fita estendida e randômica. Nas proteínas TDR1 e HSP70, os epitopos preditos estão localizados principalmente em regiões de α -hélice e randômica. **Conclusões:** Epitopos lineares e conformacionais de célula B de proteínas de *L. infantum* foram identificados *in silico* e poderão contribuir como novos antígenos com potencial aplicação no diagnóstico e controle da leishmaniose visceral. Sugere-se que vários métodos de predição de epitopos lineares sejam combinados a fim de se obter resultados mais confiáveis.

Palavras-chave: Biologia molecular, Biologia sintética, Bioinformática, Diagnóstico, Vacinas.

ABSTRACT

PREDICTION OF B-CELL EPITOPES IN PROTEINS OF *Leishmania infantum*: *IN SILICO* ANALYSIS

Visceral leishmaniasis (VL) is a chronic disease that is endemic in 62 countries and it represents a serious problem of public health in Brazil. The conventional serodiagnosis tests use entire antigens or soluble extracts that limit the standardization of the antigen, besides cross- reactions with other diseases. An alternative method is the use of peptides from B-cell epitopes identified by means of bioinformatics tools. **Objective:** It aimed at identifying linear and conformational B-cell epitopes of *Leishmania infantum* calpain-like cysteine peptidase, thiol-dependent reductase 1 (TDR1) and HSP70 proteins, as well as identifying its secondary structure through *in silico* methodology. Afterwards, it intended to select the common linear epitopes to different methods of prediction in order to verify the composition of their amino acid residues. **Methodology:** The bioinformatics tools, IEDB, BepiPred and BcePred were used for prediction of B-cell linear epitopes and the CBtope program for predicting conformational epitopes. The secondary structure of the proteins was predicted by the PHD server. **Results:** Prediction analyses produced a total of 148 linear epitopes and 164 conformational epitopes from the three proteins, being the majority of these epitopes located in the same region. The secondary structure is composed of alpha helix, extended strand and random coil proteins. In the TDR1 and HSP70 proteins, the predicted epitopes are mainly located in alpha-helix and random coil regions. **Conclusions:** Linear and conformational epitopes of B-cell of *Leishmania infantum* proteins were identified *in silico* and they can contribute as new antigens with potential application in the diagnosis and control of visceral leishmaniasis. It is suggested that several prediction methods of linear epitopes may be combined to obtain more reliable results.

Keywords: Molecular biology, Synthetic biology, Bioinformatics, Diagnosis, Vaccines.

1. INTRODUÇÃO

A leishmaniose visceral (LV) é uma das doenças parasitárias reconhecidas pela Organização Mundial de Saúde (OMS), ocorrendo aproximadamente de 200 a 400 mil casos com 40.000 mortes por ano no mundo [1]. Nas Américas, a LV é uma zoonose causada pelo protozoário *Leishmania chagasi*, também conhecida como *L. infantum* [2] e transmitida ao homem por vetores flebotomíneos [3]. No Brasil, é um problema de saúde pública, em que o cão é considerado o principal reservatório doméstico [4].

No diagnóstico da LV os testes sorológicos convencionais empregam o patógeno íntegro ou extratos solúveis que induzem a reações cruzadas com outras doenças. Uma alternativa a esses antígenos é o uso de peptídeos sintéticos que mimetizam epitopos de célula B identificados por ferramentas de bioinformática, podendo ser úteis na substituição de proteínas integras como antígenos e imunógenos, e apresentam vantagens por serem seguros, padronizáveis, específicos, reprodutíveis e disponíveis em larga escala nas análises *in vitro* [5].

As análises *in silico* de predição de epitopos de célula B permitem a identificação de epitopos lineares e conformacionais, que são úteis na síntese de fragmentos sintéticos (peptídeos) de proteínas antigênicas capazes de se ligarem a anticorpos e/ou induzirem a produção de anticorpos anti-peptídeos. Dessa forma, os peptídeos sintéticos são usados para definir aqueles epitopos identificados previamente [6].

Os métodos de bioinformática para predição de epitopos, que vislumbram a utilização da tecnologia de síntese de peptídeos, é uma estratégia mais simples, rápida, e com menor custo de produção e identificação de regiões antigênicas de proteínas em organismos patogênicos, incluindo proteínas de *L. infantum*. Assim, a predição de epitopos lineares e conformacionais de célula B por metodologia *in silico* é uma importante área de pesquisa que vem a contribuir para produção de novos antígenos a serem aplicados no controle da LV.

Este estudo pretendeu, portanto, identificar epitopos lineares e conformacionais de célula B a partir de proteínas de *L. infantum* com potencial aplicação em testes imunodiagnósticos da leishmaniose visceral.

Esta tese de doutorado compõe-se por dois artigos, dos quais um é uma revisão de literatura que aborda sobre metodologias de bioinformática empregadas na predição de

epitopos de célula B; e o outro artigo (original) refere-se a uma análise *in silico* de predição de epitopo linear e conformacional de célula B, a partir de proteínas antigênicas de *L. infantum*.

2. OBJETIVOS

2.1. Geral

- Identificar *in silico* epitopos lineares e conformacionais de célula B a partir das proteínas cisteína peptidase calpaina-like, redutase thiol dependente 1 e HSP70 de *L. infantum* com aplicação em testes imunodiagnósticos da leishmaniose visceral.

2.2. Específicos

- Selecionar os epitopos lineares de célula B comuns a diferentes métodos de predição e verificar a composição dos resíduos de aminoácidos nesses epitopos;
- Comparar os epitopos conformacionais e lineares preditos a partir das proteínas cisteína peptidase calpaina-like, redutase thiol dependente 1 e HSP70 de *Leishmania infantum*.
- Verificar a estrutura secundária das proteínas cisteína peptidase calpaina-like, redutase thiol dependente 1 e HSP70 de *Leishmania infantum*.

3. REVISÃO DE LITERATURA

Artigo de Revisão:

Ferramentas de bioinformática na predição de epitopos de célula B

LUCIANA M. ASSIS, MARCIA A. MELO, EDGAR M. CARVALHO

Situação: a submeter

Artigo de Revisão:

Ferramentas de bioinformática na predição de epitopos de célula B

Bioinformatics tools to prediction of B-cell epitopes

LUCIANA M. ASSIS¹, MARCIA A. MELO², EDGAR M. CARVALHO³

¹Pos-Graduação em Medicina e Saúde, Complexo Hospitalar Universitário Professor Edgard Santos, Universidade Federal da Bahia, Salvador, Bahia, Brasil;

Unidade Acadêmica de Enfermagem, Universidade Federal de Campina Grande-UFCG, Cajazeiras, Paraíba, Brasil.

²Pos-Graduação em Medicina Veterinária; Unidade Acadêmica de Medicina Veterinária, Universidade Federal de Campina Grande-UFCG, Patos, Paraíba, Brasil

³Pos-Graduação em Medicina e Saúde; Serviço de Imunologia, Complexo Hospitalar Universitário Professor Edgard Santos, Universidade Federal da Bahia, Salvador, Bahia, Brasil

Instituto Nacional de Ciência e Tecnologia de Doenças Tropicais- INCT DT(CNPq/MCT), Salvador, Bahia, Brasil.

RESUMO

A identificação de epitopos a partir de programas de bioinformática tem sido extensivamente utilizada visando sua aplicação no imunodiagnóstico, na elaboração de vacinas e imunoterapia de várias doenças. Epitopo de célula B é um segmento de um antígeno que é reconhecido por um anticorpo secretado ou de membrana. Os epitopos de antígenos proteicos podem ser lineares, que compreendem segmentos contínuos de aminoácidos, e conformacionais, que consistem de segmentos descontínuos de aminoácidos com proximidade espacial na estrutura proteica. Atualmente, existem várias ferramentas de bioinformática para predição de epitopos que podem ser usadas para identificar peptídeos antigênicos, possibilitando o diagnóstico e controle de doenças infecciosas, não infecciosas e parasitárias. O conhecimento de metodologias *in silico* fornece uma orientação na escolha dos métodos a serem aplicados na identificação de epitopos de célula B de proteínas antigênicas. Assim, esta revisão apresenta uma visão geral das metodologias de predição de epitopos de célula B disponíveis.

Palavras-chave: análise *in silico*, anticorpos, banco de dados, proteínas.

ABSTRACT

The identification of epitopes from bioinformatics programs has been widely used with the purpose of their application in the immunodiagnosis regarding the use of vaccines and immunotherapy of several diseases. B-cell epitope is a segment of an antigen which is recognized by an secreted antibody or an membrane antibody. The epitopes of protein antigens can be linear which encompass continuous segments of amino acids and conformational epitopes that consist of discontinuous segments of amino acids with spatial proximity in the protein structure. Nowadays, there are several bioinformatics tools for predicting epitopes that can be used to identify antigenic peptides enabling the diagnosis and control of infectious, non-infectious and parasitic diseases. The knowledge of methodologies *in silico* provides guidelines concerning the choice of the methods to be applied in the identification of B-cell epitopes of antigenic proteins. Therefore, this review presents an overall view of the prediction methodologies of the available B-cell epitopes.

Keywords: *In silico* analysis, antibodies, database, proteins.

INTRODUÇÃO

A Leishmaniose visceral é uma doença crônica causada por um protozoário do gênero *Leishmania* que acomete homens e animais suscetíveis [1]. A doença é endêmica em 62 países e estima-se que a incidência seja de 202.200 a 389.100 casos novos e com 40.000 mortes por ano no mundo [2].

Nas Américas, a *Leishmania chagasi* (também conhecida como *L. infantum*) [3] é a espécie comumente responsável pelas formas clínicas da doença visceral, sendo os cães considerados os principais reservatórios domésticos desse protozoário. Em áreas endêmicas os cães infectados, embora assintomáticos, servem de reservatórios para os vetores de transmissão (flebotomíneos) aos humanos e animais suscetíveis [4].

No Brasil, a LV é um sério problema de Saúde Pública em razão de sua grande distribuição geográfica, especialmente na região Nordeste [5]. O controle da transmissão ao homem se faz por redução da população de vetores, eliminação de hospedeiros domésticos infectados e diagnóstico e tratamento precoces de casos humanos [6].

Por ser uma doença de notificação compulsória e com características clínicas de evolução grave, o diagnóstico deve ser feito de forma precisa e precoce. As técnicas recomendadas no diagnóstico da LV pelo Ministério da Saúde são a reação de imunofluorescência indireta (RIFI) e o ensaio imunoenzimático (*Enzyme Linked Immunosorbent Assay* – ELISA), ambos utilizando antígenos totais capazes de causar reações cruzadas. O ELISA é uma técnica sorológica considerada de melhor escolha por ser prática, de fácil padronização e adequada para análise em massa [7]. No entanto, a sensibilidade e especificidade dependem da qualidade do antígeno empregado. Testes sorológicos usando preparações antigênicas de extrato de *Leishmania* apresentam limitações quanto à especificidade e reprodutibilidade [8] e não diferenciam a infecção subclínica da doença ativa [9].

Nas últimas décadas houve um crescente aumento na identificação e caracterização de antígenos de *Leishmania*. Em 1990, MacFarlane et al. [10] identificaram e caracterizaram um antígeno do grupo de proteínas conservadas de *Leishmania*, denominada proteína de choque térmico (*Heat Shock Proteins* – HSP) de 70kDa. Em 1992, Andrade e Kirchhoff [11] mostraram que não houve reação cruzada entre pacientes com doença de Chagas e leishmaniose visceral quando o ensaio imunoenzimático foi realizado com HSP70 e HSP90

recombinantes. Alguns estudos [12,13,14] relataram que o uso de antígenos recombinantes ou purificados de *Leishmania* melhora a sensibilidade e especificidade do diagnóstico.

Outra forma para produção de antígenos para imunoenensaio e desenho de vacinas é através da síntese de peptídeos que mimetizam epitopos de proteínas antigênicas do patógeno. O uso desses peptídeos sintéticos aumenta a especificidade de imunoenensaio comparado aos antígenos brutos [15]. No entanto, a identificação de peptídeos através de métodos experimentais apresenta limitações como custo elevado e a demora no mapeamento de antígenos ou peptídeos/epitopos.

Um método alternativo para identificação de peptídeos com potencial antigênico e ou imunogênico é através da abordagem bioinformática de predição de epitopos de célula B.

Na atualidade, várias ferramentas computacionais de predição de epitopos de célula B estão disponíveis para identificação de peptídeos a serem aplicados em imunoenensaio, produção de vacinas e imunoterapia. Assim, o conhecimento de metodologias *in silico* fornece uma orientação na escolha dos métodos a serem aplicados na identificação de epitopos de proteínas antigênicas.

Esta revisão aborda algumas metodologias *in silico* usadas para predição de epitopos lineares e conformacionais de célula B e relata os resultados do uso dessas ferramentas na identificação de epitopos de proteínas de *Leishmania infantum*.

ASPECTOS IMUNOLÓGICOS

Quando um agente patogênico invade o organismo de um indivíduo, os linfócitos B reconhecem os antígenos do patógeno e produzem anticorpos específicos. Desse modo, os antígenos apresentam propriedades de antigenicidade e ou imunogenicidade. A antigenicidade é a propriedade do antígeno para se ligar especificamente a anticorpos complementares; enquanto que, a habilidade de um antígeno em induzir uma resposta imune é denominada de imunogenicidade [16].

O antígeno corresponde a qualquer estrutura molecular, tal como proteínas, carboidratos, lipídeos, nucleotídeos ou uma combinação dos mesmos [17]. Um epitopo ou determinante antigênico é um segmento do antígeno que é reconhecido por um anticorpo/linfócito B ou linfócito T, definido como epitopo de célula B e T, respectivamente.

A região do anticorpo ou receptor da célula B (BCR) que reconhece o epítipo é denominada de paratopo. A célula B e anticorpos geralmente reconhecem antígenos solúveis. Por essa razão, os epítipos de célula B tendem a estar em locais altamente acessíveis, expostos na superfície do antígeno.

Quando o antígeno é protéico, os epítipos de célula B podem ser lineares (contínuos) ou conformacionais (descontínuos). Epítipos lineares são pequenas sequências contínuas de resíduos de aminoácidos do antígeno; enquanto que, os epítipos conformacionais são compostos por resíduos de aminoácidos de sequências descontínuas, mas que estão juntas pelo dobramento da proteína na sua forma nativa [18].

BIOINFORMÁTICA

A bioinformática representa uma nova e crescente área da ciência que utiliza métodos computacionais para responder a questões biológicas [19]. Presentemente, os procedimentos da bioinformática são bem mais complexos do que apenas projetar e preencher bancos de dados. Ela envolve a biologia, a ciência da computação, a matemática e a estatística para analisar dados de sequências biológicas, conteúdo de genomas e para a predição de função e estruturas de macromoléculas, utilizando programas computacionais para fazer inferências a partir de dados obtidos da biologia molecular moderna, fazer conexões entre eles e derivar predições importantes e relevantes [20].

Uma das áreas de atuação da bioinformática é a predição de epítipos de célula B. Os métodos computacionais para predição de epítipos de proteínas antigênicas têm sido o foco de pesquisas [21, 22, 23]. Os estudos *in silico* vem crescendo exponencialmente devido, principalmente, a grande quantidade de dados gerados pela biologia molecular [24], e a necessidade de desenvolvimento de tecnologias adequadas ao armazenamento, manuseio e interpretação desses dados. Nas últimas décadas, vários algoritmos computacionais têm sido desenvolvidos na busca por sequências de aminoácidos que possam ser comuns a peptídeos imunogênicos, capazes de induzir uma resposta imune [25].

Banco de dados e ferramentas para análises de dados

Um banco de dados compreende um arquivo de informações, uma organização lógica dessas informações e ferramentas para se ter acesso a elas. Os bancos de dados da biologia molecular contêm sequências de ácidos nucleicos e de proteínas, estruturais e funções de macromoléculas, padrões de expressão, redes de vias metabólicas e cascatas de regulação [20].

As bases de dados de epitopos consistem em uma coleção de epitopos de patógenos como vírus, bactérias, protozoários e fungos. Nos últimos anos, esses bancos de dados têm se tornado uma fonte cada vez mais utilizada na busca de antígenos para aplicação em vacinas, no diagnóstico e como imunoterápicos. Além disso, bioinformatas utilizam as bases de dados para explorar, refinar e desenvolver novas ferramentas e algoritmos [17].

IEDB (Immune Epitope Database and Analyses Resource – Base de Dados de Epitopos Imune e Recursos de Análises)

O IEDB possui dados que permitem a predição de epitopos de anticorpos e de célula T para humanos, outros primatas, roedores e outras espécies. A curagem de epitopos, peptídicos ou não, de agentes infecciosos, alérgenos, doenças autoimunes e aloantígenos é constantemente atualizada. As bases de dados se destacam pela precisão das informações armazenadas e por apresentar os epitopos juntamente com os contextos experimentais em que foram determinados [26]. Os recursos estão disponíveis gratuitamente (www.immunepitope.com) e contêm ferramentas preditivas e analíticas [17].

O desempenho das ferramentas de predição de epitopos foi avaliado rigorosamente por vários autores [27, 28, 29, 30, 31]. Parâmetros rigorosos de acurácia, sensibilidade e valores preditivos dos algoritmos foram definidos e consistentemente aplicados. Atualmente, o principal determinante do desempenho de um algoritmo parece ser a quantidade de dados para treinar e avaliar os preditores [17].

PREDIÇÃO DE EPITOPOS

A predição de peptídeos imunogênicos pode ser usada para produção de anticorpos anti-peptídeos, isto é, anticorpos usados em imunoenaios para identificar proteínas

antigênicas específicas; como também peptídeos capazes de induzir uma resposta imune humoral *in vivo*, ou seja, peptídeos capazes de provocar a produção de anticorpos que reagem cruzadamente se ligando a moléculas específicas, a partir das quais o peptídeo foi derivado [16].

Durante alguns anos várias tentativas foram feitas para descobrir peptídeos que mimetizassem epitopos de proteínas e que pudessem ser antigênicos e/ou imunogênicos. O principal objetivo da predição de epitopos é para desenhar moléculas que possam substituir antígenos em um processo de produção ou detecção de anticorpos. O uso dessas moléculas apresenta vantagens por serem de menor custo e não infecciosas, quando comparadas aos organismos patogênicos que podem ser perigosos ao pesquisador ou ao animal experimental [32].

Predição de epitopos de Células B

A identificação e caracterização de epitopos de células B desempenham um importante papel no desenho de vacinas, testes imunodiagnósticos e produção de anticorpos. Os métodos experimentais para identificação de epitopos de células B estão divididos em duas categorias – estrutural e funcional. Os métodos estruturais incluem a cristalografia de raio-X, ressonância magnética nuclear e microscopia eletrônica de complexos antígeno-anticorpo. Métodos funcionais para detecção e caracterização da ligação antígeno-anticorpo utilizam várias técnicas, como ressonância da superfície plasmática [33], a espectrometria de massa [34] e a espectroscopia de ressonância magnética nuclear, bem como imunoenaios, incluindo ELISA, ELISPOT e Western blot [35].

Os estudos estruturais identificam os epitopos estruturais, também chamados de determinantes antigênicos, definidos por um conjunto de resíduos ou átomos na proteína antigênica que estão em contato com resíduos ou átomos do anticorpo. Enquanto que, os epitopos funcionais são determinados através de ensaios funcionais, e consistem de resíduos de antígenos que contribuem significativamente para ligação ao anticorpo [36]. Epitopos funcionais em proteínas são geralmente menores do que os epitopos estruturais; apenas três a cinco resíduos do epitopo estrutural contribuem de modo significativo para a energia de ligação antígeno-anticorpo [37].

Vários métodos computacionais vêm sendo desenvolvidos e usados como ferramentas para predição de epitopos contínuos e descontínuos de proteína. Embora se acredite que a

maioria dos epitopos de célula B sejam descontínuos [38], os estudos de identificação de epitopos através dos métodos de bioinformática têm focado primariamente sobre os epitopos lineares [39], por usar na entrada do programa a sequência de aminoácidos da proteína e por ser necessário, na maioria dos métodos de predição de epitopos descontínuos de célula B, o conhecimento da estrutura tri-dimensional da proteína [16].

Predição de epitopos lineares

A razão das predições de epitopos lineares encontra-se na suposição de que peptídeos sintéticos correspondentes a segmentos lineares em proteínas podem induzir a produção de anticorpos anti-peptídeos que, por sua vez, podem reagir de forma cruzada com o antígeno da proteína nativa (os chamados anticorpos com reatividade cruzada), além de neutralizar a infectividade do patógeno abrigando o antígeno (os chamados anticorpos neutralizantes com reatividade cruzada) [32].

A maioria dos algoritmos utilizados baseia-se no pressuposto que os epitopos reconhecidos por respostas de anticorpos são expostos em superfícies de proteínas e/ou são ricos em resíduos de aminoácidos específicos [17]. De acordo com Barlow et al., (1986) as partes contínuas na superfície de uma proteína se localizam predominantemente nas alças e/ou regiões salientes da molécula [18].

O primeiro método para predição de epitopo linear foi desenvolvido por Hopp e Woods em 1981 [40], baseado na propriedade dos aminoácidos. De acordo com esse método, os epitopos são encontrados em regiões da proteína que tem alto grau de exposição ao solvente. Os autores atribuíram a cada aminoácido, em uma sequência, a escala de hidrofiliidade. A partir de então, outros trabalhos de predição de epitopos lineares usaram as propriedades físico-químicas dos aminoácidos, tais como hidrofiliidade [41], flexibilidade [42], acessibilidade ao solvente [43] e estrutura secundária [44] para prever os sítios antigênicos na proteína.

Pellequer *et al.* (1991) [45] reuniram vários métodos de predição de epitopos lineares que utilizavam escalas de propensão de aminoácidos baseados nas propriedades físico-químicas, e os avaliaram quanto à taxa de sucesso. A comparação de 12 escalas aplicadas a 85 epitopos lineares identificados em 14 proteínas demonstrou que a maioria das escalas apresentou de 50 a 62% de predições corretas [45]. Em 1993, os mesmos pesquisadores observaram que a escala beta-turn apresentou bom desempenho, predizendo corretamente

70% dos epítopos conhecidos [46]. Posteriormente, Saha e Raghava (2004), através do programa BcePred, utilizaram várias propriedades escalares combinadas e observaram uma melhora na acurácia da predição de epítopos lineares de célula B quando comparada a utilização de uma única escala [47].

A qualidade preditiva dos métodos baseados em uma única propriedade foi criticamente avaliada por Blythe e Flower (2005), que motivaram o desenvolvimento de novos métodos com abordagem mais avançada para predição de epítopos lineares [48]. Assim, vários métodos usando abordagens de aprendizagem de máquinas foram publicados, tais como os métodos BepiPred [27], ABCpred [49] e o BEST (Ferramenta de predição de epítopo de célula B usando máquina vetor de suporte) [50].

Os métodos de predição de epítopos baseados em aprendizagem de máquinas tem apresentado uma maior acurácia comparado aos que usam apenas propriedades físico-químicas. Entretanto, não há uma ferramenta preditiva perfeita, todos têm pontos fortes e fracos [51].

Predição de epítopos conformacionais

A predição de epítopos conformacionais ou descontínuos em proteína baseia-se no conhecimento da estrutura tri-dimensional (3D) da proteína. Levando-se em consideração que tais epítopos estão localizados na superfície da proteína, tornando-se favorável à ligação aos anticorpos [16].

Diferentes abordagens de predição de epítopos tornaram-se possíveis somente após a primeira estrutura de raio-X de um complexo proteína-anticorpo ter sido solucionada em 1986 [52]. Baseado nas propriedades estruturais, Jones e Thornton (2000) conseguiram prever epítopos sobre a superfície da subunidade β da gonadotrofina coriônica humana (β hCG) [53]. Por conseguinte, uma variedade de ferramentas da bioinformática para predição de epítopos conformacionais baseado na estrutura 3D de proteínas foram desenvolvidas, tais como EPIMAP [54], CEP [55], DiscoTope [56] e Mapitope [57]. Todos eles apresentam diferentes desempenhos; por exemplo, o CEP prediz com mais acurácia do que o DiscoTope. Métodos mais recentes usando modelos de aprendizagem de máquinas incluem o EPSVR e EPMeta [58].

A aplicação dos modelos computacionais para predição de epítopos baseado na estrutura é limitado pelo número ainda restrito de estruturas antigênicas disponíveis.

Atualmente, existem métodos de predição de epitopos conformacionais baseados na sequência primária de aminoácidos, tais como CBTope [30], BEST [50] e a mais recente técnica proposta por Zang et al. (2012) que reúne múltiplos modelos de aprendizagem de máquinas [59].

MÉTODOS PARA PREDIÇÃO DE EPITOPOS DE CÉLULA B

A maioria dos métodos computacionais de bioinformática para predição de epitopos de células B está disponível *on line* e compreende métodos baseados na sequência e métodos fundamentados na estrutura da proteína.

Os métodos baseados na sequência da proteína são geralmente utilizados para predição de epitopos contínuos ou lineares. Entretanto, foram propostos novos métodos baseados em sequência para predição de epitopos conformacionais [30, 50].

Geralmente, os métodos de predição são testados sob um conjunto de epitopos de referência. Segue alguns dos métodos baseados na sequência da proteína:

- Métodos baseados em escalas de aminoácidos

Esses métodos aplicam uma escala de aminoácido para calcular o escore de um resíduo i em uma dada sequência protéica. Os resíduos vizinhos $i-(n-1)/2$ em cada lado do resíduo i são usados para calcular o escore para o resíduo i em uma janela de tamanho n . O escore final do resíduo i é a média dos valores escalares para n aminoácidos na janela. O resíduo i (central) é predito positivo se seu escore for maior do que um limiar pré-determinado. Neste caso, esse escore do resíduo do centro reflete a probabilidade do resíduo fazer parte de um epitopo [16].

Em geral, o tamanho da janela de 5 a 7 aminoácidos é apropriado para encontrar regiões que podem potencialmente ser antigênicas.

Desse modo, os cálculos de predição dos métodos escalares são baseados nas escalas de propensão para cada um dos 20 aminoácidos da sequência de proteína, e esses valores escalares são obtidos de acordo com as propriedades físico-químicas dos aminoácidos tais como, hidrofobicidade, flexibilidade, acessibilidade, voltas (*turns*) e antigenicidade [29].

- ***Método de Parker de escala de hidrofiliçidade (Parker, Guo, e Hodges, 1986)***

A escala de hidrofiliçidade de Parker foi determinada experimentalmente usando cromatografia líquida de alta performance (HPLC) sobre um conjunto de 20 peptídeos sintéticos para cada um dos 20 aminoácidos. O valor hidrofílico para cada aminoácido é atribuído conforme o tempo de retenção para cada peptídeo. Nos cálculos foi usada uma janela de sete resíduos.

- ***Método de Karplus e Schulz de escala de Flexibilidade (Karplus e Schulz, 1985)***

Nesse método, a escala de flexibilidade é construída com base na mobilidade de segmentos protéicos derivado de fator B de temperatura dos átomos de Carbono α de 31 proteínas de estrutura conhecida.

- ***Método Emini de escala de acessibilidade ao solvente (Emini et al., 1985)***

Nesse método os cálculos são baseados na escala de acessibilidade à superfície, que foi determinada por Janine e Wodak (1978) [60] e reflete a probabilidade de aminoácidos estarem na superfície exposta de proteínas. Quando o score ou valor de um hexapeptídeo da sequência for maior do que o limiar, isso indica um aumento da probabilidade deste peptídeo (epitopo) ser encontrado na superfície.

- ***Método Chou e Fasman de predição B-turn (Chou e Fasman, 1978)***

O método baseia-se no cálculo da probabilidade de um trecho de resíduos fazer parte da estrutura secundária β -turn (β -volta). Cerca de um a quatro de todos os resíduos da proteína estão em β -turn [61].

Vários métodos têm sido desenvolvidos para identificação da localização de β -turn (voltas) em uma sequência de proteínas. A maioria dos métodos são empiricamente baseados na abordagem da preferência posicional, que usa as propriedades do resíduo em posições i e $i+3$ para calcular a frequência posicional e parâmetros conformacionais [52].

- ***Método de Kolaskar e Tongaonkar de escala de antigenicidade (Kolaskar e Tongaonkar, 1990)***

Os valores escalares de propensão antigênica (A_p) são calculados para cada aminoácido usando a relação $A_p = f_{Ag}/f_s$. Para isso, os autores calcularam a frequência de ocorrência de cada tipo de aminoácido (f_{Ag}) em 156 epitopos determinados experimentalmente a partir de 34 proteínas diferentes, e calcularam também a frequência de ocorrência dos aminoácidos na superfície (f_s) das proteínas, baseados em escalas de Parker [41]. Cerca de 75% dos epitopos foram corretamente preditos aplicando a escala de propensão antigênica para os epitopos para os quais a escala foi construída [62].

- **Método BcePred (Saha e Raghava, 2004)**

BcePred fornece predição de epitopos de células B, utilizando abordagem tradicional com base em várias propriedades físico-químicas combinadas. A acurácia desse método é de 58.70%. Adicionalmente, um servidor web foi desenvolvido para operacionalizar o método de predição de epitopos de célula B em uma sequência antigênica. Esse serviço está disponível *on line* (<http://www.imtech.res.in/raghava/BcePred/>).

O programa BcePred permite ao usuário selecionar as propriedades físico-químicas dos resíduos de aminoácidos, isto é, selecionar os parâmetros escalares, de forma individual ou combinada. Conforme os resultados, os resíduos de aminoácidos com valores acima do valor do limiar padrão são considerados epitopos de célula B predito [47].

Outro recurso de análise de interesse imunológico que hospeda ferramentas preditivas é o IEDB. Esse programa reúne cinco métodos escalares de predição de epitopos de célula B linear que incluem os parâmetros escalares turns, acessibilidade, flexibilidade, antigenicidade e hidrofiliidade; além de incorporar o método BepiPred [63].

- **Métodos baseados em aprendizagem de máquinas**

Aprendizagem de máquinas consiste em programas de computadores que permitem ao computador aprender, isto é, permitem ao computador melhorar seu desempenho através do uso de dados de treinamento ou experiência [64].

Destacaremos três métodos com abordagem de aprendizado de máquinas: BepiPred, ABCpred e CBtope.

- ***Método BepiPred (Lansen, Lund, e Nielsen, 2006)***

Esse método é baseado na combinação do modelo oculto de Markov e duas escalas de aminoácidos, escala de hidrofobicidade de Parker [41] e escala de estrutura secundária de Levitt [65]. Essas escalas foram consideradas serem de melhor desempenho entre cinco métodos escalares, utilizando um conjunto de dados de Pellequer [46].

- ***Método ABCpred (Saha e Raghava, 2006)***

O método ABCpred aplica rede neural artificial para predição de epitopos contínuos na sequência de proteínas. Esse método foi treinado sob um conjunto não redundante de 700 epitopos da base de dados Bcipep [66] e 700 peptídeos derivados aleatoriamente da base de dados do Swiss-Prot. O método alcançou uma acurácia de aproximadamente 66% [49].

- ***Método CBtope (Ansari et al., 2010)***

O CBtope é um método para predição de epitopos conformacionais de célula B a partir de sequência primária antigênica e teve sua maior acurácia (86,59%) quando usou uma janela de 19 resíduos de aminoácidos. Esse preditor é baseado no modelo de vetores de aprendizagem de máquinas que predizem os resíduos de aminoácidos na proteína que interage com anticorpos. Foi observado que esse modelo baseado em sequência é tão bom quanto os baseados em estrutura [30].

ESTUDOS DE APLICAÇÃO DA PREDIÇÃO DE EPITOPOS PARA O DIAGNÓSTICO E O DESENHO DE VACINAS

Métodos computacionais para predição de epitopos vêm sendo usados na busca por peptídeos para aplicação em imunodiagnósticos e vacinas. Quando usado em testes diagnósticos, o peptídeo deve ser antigênico, isto é capaz de se ligar a um anticorpo específico. Hsueh *et al.*, (2004) encontrou 100% de especificidade no diagnóstico da síndrome

respiratória aguda severa usando um ELISA (enzyme-linked immunosorbent assay) constituído de um peptídeo sintético formado por epitopos contínuos [67].

Em um estudo com *L. chagasi*, objetivando identificar novos antígenos, 41 proteínas foram mapeadas para epitopos de célula B usando o programa BepiPred. Dos 180 peptídeos preditos pelo programa, sintetizados e testados em imunoensaio, 25 desses demonstraram ser de interesse para uso no sorodiagnóstico da LV [68].

Faria et al. (2011) mapearam epitopos de célula B de proteínas imunogênicas de *L. infantum* usando os programas ABCPred, BCPred e BepiPred; os peptídeos preditos foram sintetizados e testados em imunoensaio. A predição de epitopos de célula B usando ferramentas de bioinformática demonstrou grande eficiência em razão da maioria dos peptídeos/epitopos terem sido imunogênicos no imunoensaio. Ademais, dos 360 peptídeos testados 10 foram os mais reativos (específicos) no sorodiagnóstico da leishmaniose visceral canina (LVC), e foram os epitopos preditos simultaneamente pelos programas ABCPred, BCPred, e destes 10, nove eram peptídeos comuns aos três programas de predição. Estes peptídeos demonstraram boa reatividade como antígenos e foram capazes de reconhecer anticorpos específicos em ELISA, demonstrando, assim, serem de interesse para uso no sorodiagnóstico da LVC [69].

Em outros estudos têm sido observado resultados promissores com o uso de epitopos de célula B preditos por técnicas de predição de bioinformática. Sharma et al. (2013) identificaram epitopos de célula B de proteína alergênica usando várias ferramentas de bioinformática, modificaram as regiões dos epitopos com propriedades alergênicas e observaram uma redução na especificidade destas regiões através do ELISA [70].

Outra aplicação da predição de epitopos é no desenho de vacinas de peptídeos que imitam epitopos imunogênicos. Peptídeos da região N-terminal de proteínas virais (VP2) acoplados a um carreador induziram resposta imune protetora contra a parvovirose em cães e vison [71]. Para um peptídeo ser considerado candidato a vacina é preciso primeiramente demonstrar antigenicidade. Além disso, sua possível imunogenicidade precisa ser testada experimentalmente e, para alcançar imunogenicidade suficiente, a vacina sintética deve incluir epitopos de célula T [72].

CONSIDERAÇÕES FINAIS

Esta revisão de literatura relata o papel da bioinformática como ferramenta útil na identificação de epítopos lineares e conformacionais de célula B e sua aplicação no diagnóstico e desenho de vacinas. A sequência do epítopo pode ser usada para produzir um peptídeo sintético que pode sofrer mudanças conformacionais produzindo um antígeno adequado que imita a estrutura nativa da região antigênica da proteína, sendo capaz de se ligar e/ou estimular a produção de anticorpos específicos. A predição de epítopos ainda têm limitações quanto à acurácia dos métodos, ao banco de epítopos utilizado e ao número restrito de estrutura terciária (antigênica) disponível; sendo necessários estudos que melhorem a acurácia dos métodos de predição; isso pode ser obtido melhorando a qualidade dos bancos de dados de epítopos que são usados para avaliar a eficiência dos métodos. Apesar das limitações, alguns estudos usando a bioinformática demonstraram que vários epítopos preditos *in silico* submetidos à síntese de peptídios e testados em imunoenaios apresentaram resultados satisfatórios, demonstrando a eficiência dos métodos de predição, sua utilidade na identificação de novos antígenos, sobretudo de *L. infantum*, e sua aplicação no desenvolvimento de testes diagnósticos mais precisos e desenho de vacinas. Ademais, foi observado que os métodos de identificação de epítopos *in silico* apresentam vantagens por serem rápidos, de menor custo, reprodutíveis e mais simples do que os métodos convencionais.

REFERÊNCIAS

1. Amato VS. Leishmaniose tegumentar americana. In: Lopes, AC. Tratado de Clínica Médica. VIII. São Paulo: roca, 2006. cap. 407, p.4114 – 4121.
2. Alvar J, Vélez ID, Bern C, Herrero M, Desjeux P, Cano J, Jannin J, Boer M den. Leishmaniasis Worldwide and Global Estimates of Its Incidence. PLoS One. 2012;7(5): 1-12. e35671.
3. Maurício IL, Stohard JR, Miles MA. The strange case of Leishmania chagasi. Parasitol Today. 2000; 16: 188-189.

4. Michalsky EM, Rocha MF, da Rocha Lima AC, França-Silva JC, Pires MQ, et al. Infectivity of seropositive dogs, showing different clinical forms of leishmaniasis, to *Lutzomia longipalpis* phlebotomine sand flies. *Vet Parasitol* 2007; 147: 67–76.
5. Urias EV, Carvalho SF, Oliviera CL, Carvalho ML; Teles LF; Rodrigues MC, Maia CN. Prevalência de adultos infectados por *Leishmania leishmania chagasi* entre doadores de sangue do Hemocentro Regional de Montes Claros, Minas Gerais, Brasil. *Revista Brasileira de Hematologia e Hemoterapia*. 2009; 31(5): 348-354.
6. Tesh R. Control of zoonotic visceral leishmaniasis: is it time to change strategies? *Am J Trop Med Hyg*. 1995; 52: 287–292.
7. Mohapatra TM, Singh DP, Sen MR, Bharti K; Sundar S. Comparative evaluation of rK9, rK26 and rK39 antigens in the serodiagnosis of Indian visceral leishmaniasis. *J Infect Dev Ctries*. 2010; 4(2): 114-117.
8. Badaro R, Reed SG, Carvalho EM. Immunofluorescent antibody test in American visceral leishmaniasis: sensitivity and specificity of different morphological forms of two *Leishmania* species. *Am J Trop Med Hyg*. 1983; 32: 480-484.
9. Carvalho EM, Badaro R, Reed SG, Jones TC, Johnson JR. Absence of gamma interferon and interleukin 2 production during active visceral leishmaniasis. *J. Clin. Investig*. 1985; 76: 2066–2069.
10. Macfalane JML, Blaxter RP, Bishop MA, KELLY JM. Identification and characterisation of a *Leishmania donovani* antigen belonging to the 70-kDa heat-shock protein family. *Eur. J. Biochem*. 1990; 190: 377–384.
11. Andrade CR, Kirchoff LV. Recombinant *Leishmania* Hsp90 and Hsp70 are recognized by sera from visceral leishmaniasis patients but not Chagas' Disease Patients. *J Clin Microbiol*. 1992 feb; 30(2): 330-335.
12. Shreffler WG, Burns JM, Badaró R, Ghalib HW, Button LL, McMaster WL, Reed SG. Antibody responses of visceral leishmaniasis patients to gp63, a major surface glycoprotein of *Leishmania* species. *J Infect Dis*. 1993; 167: 426–430.
13. Lohman KL, Langer PJ, Mcmahon-Pratt D. Molecular cloning and characterization of the immunologically protective surface glycoprotein GP46/M-2 of *Leishmania amazonensis*. *Proc Natl Acad Sci* 1990; 87: 8393–8397.

14. Badaró R, Benson D, Eulálio MC, Freire S, Cunha EM, Ntto D, Pedral-Sampaio, Madureira C, Burns JM, Houghton RL, David JR, Reed SG. rK39: a cloned antigen of *Leishmania chagasi* that predicts active visceral leishmaniasis. *J Infect Dis.* 1996; 173: 758–761.
15. Ferrer E, Benitez L, Foster-Cuevas M, Bryce D, Wamae LW, et al. Taenia saginata derived synthetic peptides with potential for the diagnosis of bovine cysticercosis. *Vet Parasitol.* 2003; 111: 83–94.
16. Ponomarenko JV, Van Regenmortel MHV. B cell epitope prediction. In *Structural Bioinformatics*, 2nd ed, eds. Gu J, Bourne PE, Hoboken, NJ: John Wiley; 2009; 849-879.
17. Salimi N, Fleri W, Peters B, Sette A. Design and utilization of epitope-based databases and predictive tools. *Immunogenetics.* 2010; 62: 185–196.
18. Barlow DJ, Edwards MS, Thornton JM, et al. Continuous and discontinuous protein antigenic determinants. *Nature.* 1986; 322: 747–748.
19. Baxevanis AD, Ouellette BFF. *Bioinformatics a practical guide to the analysis of Genes and Proteins.* 2.ed. John Wiley & Sons, Inc. 2001.
20. Lesk AM. *Introdução à Bioinformática.* Tradução Ardalá Elisa Breda Andrade, 2.ed. Porto Alegre. Artmed, 2008.
21. Alix AJ (1999): Predictive estimation of protein linear epitopes by using the program PEOPLE. *Vaccine.* 1999; 18(3–4):311–314.
22. Eigenmann PA. Do we have suitable in-vitro diagnostic tests for the diagnosis of food allergy? *Curr Opin Allergy Clin Immunol.* 2004; 4(3): 211–213.
23. Mahdavi M, Mohabaktar M, Keyhanfar M, Dehkordi AJ, Rabbani M. Linear and conformational B Cell epitope Prediction of the HER 2 ECD-Subdomain III in silico Methods. *Asian Pacific J Cancer Prev.* 2012; 13: 3053-3059.
24. Baldi P, Brunak S. *Bioinformatics: The Machine Learning approach.* 2.ed. Cambridge: MIT Press, 2001.
25. Luchese A, Stenovic S, Sinha AA, Mittelman A, Kanduc D. Role of MHC II affinity and molecular mimicry anti-HER-2/neu MAb-3 linear peptide epitope. *Peptides.* 2003; 24(2): 193-197.

26. Vita R, Zarebski L, Greenbaum JA, Emami H, et al. The Immune Epitope Database 2.0. *Nucleic Acids Research*. 2009; 38:854-862.
27. Larsen JE, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. *Immunome Res*. 2006; 2(2).
28. Ponomarenko JV, Bourne PE. Antibody-protein interactions: benchmark databases and prediction tools evaluation. *BMC Structural Biology* 2007; 7:64
29. El-Manzalawy Y, Dobbs D, Honavar V. Predicting linear B-cell epitopes using string kernels. *J Mol Recognit*. 2008; 21: 243–255.
30. Ansari HR, Gajendra PS, Raghava GP. Identification of conformational B-cell epitopes in an antigen from its primary sequence. *Immunol Res*. 2010; 6(6).
31. Elbashir MK, Sheng Y, Jianxin W, Wu FX, Li M. Predicting β -Turns in Proteins Using Kernel Logistic Regression. BioMed Research Corporation 2013; 2013: doi:10.1155/2013/870372
32. Van Regenmortel MHV. Immunoinformatics may lead to a reappraisal of the nature of B cell epitopes and of the feasibility of synthetic peptide vaccines. *J Mol Recognit*. 2006; 19(3): 183–187.
33. Pattnaik P. Surface plasmon resonance: applications in understanding receptor–ligand interaction. *Appl Biochem Biotechnol*. 2005; 126(2): 79–92.
34. Hager-Braun C, Tomer KB. Determination of protein-derived epitopes by mass spectrometry. *Expert Rev Proteomics*. 2005; 2(5): 745–756.
35. Johnson MA, Pinto BM. NMR spectroscopic and molecular modeling studies of protein–carbohydrate and protein–peptide interactions. *Carbohydr Res*. 2004; 339(5): 907–928.
36. Sundberg EJ, Mariuzza RA (2004): Antibody structure and recognition of antigen. In: Honjo AFT, Neuberger M, editor. *Molecular Biology of B Cells*. Elsevier Science (USA). 2004. p. 491–509.
37. Cunningham BC, Wells JA (1993): Comparison of a structural and a functional epitope. *J Mol Biol*. 1993; 234(3): 554–563.
38. Walter G. Production and use of antibodies against synthetic peptides. *J. Immunol. Methods* 1986; 88: 149–161.

39. Flower DR. Immunoinformatics: Predicting immunogenicity in silico. 1st ed. Humana, Totowa NJ. 2007.
40. Hopp TP, Woods KR. Prediction of Protein Antigenic Determinants from Amino Acid Sequences. *Proc Natl Acad Sci USA*. 1981; 78: 3824–3828.
41. Parker JM, Guo D, Hodges RS. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry*. 1986; 25(19): 5425–5432.
42. Karplus PA, Schulz GE. Prediction of chain flexibility in proteins. A tool for the selection of peptide antigens. *Naturwissenschaften*. 1985; 72:S.212.
43. Emini EA, Hughes JV, Perlow DS, Boger J. Induction of hepatitis A virus: neutralizing antibody by a virus-specific synthetic peptide. *J Virol*. 1985; 55(3): 836–839.
44. Chou PY, Fasman GD: Prediction of the secondary structure of proteins from their amino acid sequence. *Adv Enzymol Relat Areas e Fasman, Mol Biol*. 1978; 47: 45–148.
45. Pellequer JL, Westhof E, Van Regenmortel MH. Predicting location of continuous epitopes in proteins from their primary structures. *Methods Enzymol*. 1991; 203: 176–201.
46. Pellequer JL, Westhof E, Van Regenmortel MH): Correlation between the location of antigenic sites and the prediction of turns in proteins. *Immunol Lett*. 1993; 36(1): 83–99.
47. Saha and GP Raghava. BcePred: Prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties. *Artificial Immune Systems, Third International Conference (ICARIS 2004), LNCS, 2004; 3239: 197–204*.
48. Blythe MJ, Flower DR. Benchmarking Bcell epitope prediction: underperformance of existing methods. *Protein Sci*. 2005; 14(1): 246–248.
49. Saha S, Raghava GP. Prediction of continuous B-cell epitopes in na antigen using recurrent neural network. *Proteins*. 2006; 65: 40–44.
50. Gao J, Faraggi E, Zhou Y, Ruan J, Kurgan L. BEST: Improved Prediction of B-Cell Epitopes from Antigen Sequences. *PloS One* 2012; 7: e40104.

51. Chen P, Rayner S, Hu KH. Advances of Bioinformatics tools applied in virus epitopes prediction. *Virologica sinica*. 2006; 26:1-7.
52. Amit AG, Mariuzza RA, Phillips SE, Poljak RJ. Three-dimensional structure of an antigen–antibody complex at 2.8Å resolution. *Science*. 1986; 233(4765):747–753.
53. Jones S, Thornton JM. Analysis and classification of protein–protein interactions from a structural perspective. In: Kleanthous C, editor. *Protein–Protein Recognition*. New York: Oxford University Press. 2000; 33–59.
54. Mumei BM, Bailey BW, Kirkpatrick B, Jesaitis AJ, Angel T, Dratz EA. A new method for mapping discontinuous antibody epitopes to reveal structural features of proteins. *J Comput Biol*. 2003; 10(3–4):555–567.
55. Kulkarni-Kale U, Bhosle S, Kolaskar AS. CEP: a conformational epitope prediction server. *Nucleic Acids Res*. 2005; 33: 168-171
56. Haste Andersen P, Nielsen M, Lund O. Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. *Protein Sci*. 2006 15(11):2558–2567
57. Bublil EM, Freund NT, Mayrose I, Penn O, Roitburd-Berman A, Rubinstein ND, Pupko T, Gershoni JM. Stepwise prediction of conformational discontinuous B-cell epitopes using the mapitope algorithm. *Proteins*. 2007; 68(1):294–304.
58. Liang S, Zheng D, Standley DM, Yao B, Zacharias M, et al. EPSVR and EPMeta: prediction of antigenic epitopes using support vector regression and multiple server results. *BMC Bioinformatics*. 2010; 11: 381.
59. Zang W, , Niu Y, Xiong Y, Zhao M, Rongwei Y, Liu J. Computational Prediction of Conformational B-Cell Epitopes from Antigen Primary Structures by Ensemble Learning. *PLoS One*. 2012; 7(8).
60. Janin, J.; Wodak S. Conformation of amino acid side-chains in proteins. *J Mol Biol*. 1978; 125: 357–386.
61. Harpreet K; Gajendra P, Raghava S. Prediction of β -turns in proteins from multiple alignment using neural network. *For the Record. Protein Science*, n.12. p.627–634. Published by Cold Spring Harbor Laboratory Press. Copyright © 2003 The Protein Society. 2003.
62. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett*. 1990; 276(1–2):172–174.

63. Peters B, Sidney J, Bourne P, Bui HH, Buus S, Doh G, Fleri W, Kronenberg M, Kubo R, Lund O, Nemazee D, Ponomarenko JV, Sathiamurthy M, Schoenberger SP, Stewart S, Surko P, Way S, Wilson S, Sette A: The design and implementation of the immune epitope database and analysis resource. *Immunogenetics*. 2005; 57(5):326-336.
64. Larranaga P, Calvo B, Santana R, Bielza C, Galdiano J, Inza I, Lozano JA, Armananzas R, Santafe G, Perez A, Robles V. Machine learning in bioinformatics. 2005; 7(1):86-112.
65. Levitt M. Conformational preferences of amino acids in globular proteins. *Biochemistry*. 1978; 17(20):4277–4285.
66. Saha S, Bhasin M, Raghava GP. Bcipep: a database of B-cell epitopes. *BMC Genomics*. 2005; 6(1):79.
67. Hsueh PR, Kao CL, Lee CN, Chen LK, Ho MS, Sia C, Fang XD, Lynn S, Chang TY, Liu SK, Walfield AM, Wang CY. SARS antibody test for serosurveillance. *Emerg Infect Dis*. 2004; 10(9):1558–1562.
68. Costa MM, Andrade HM, Bartholomeu DC, Freitas L, Pires SF, et al. Analysis of *Leishmania chagasi* by 2-D Difference Gel Electrophoresis (2-D DIGE) and Immunoproteomic: identification of novel candidate antigens for diagnostic tests and vaccine. *J Proteome Res*. 2011; 10: 2172–184.
69. Faria AR, Costa MM, GiustaMS, Grimaldi G Jr., Penido ML, Gazzinelli RT, et al. High-Throughput Analysis of Synthetic Peptides for the Immunodiagnosis of Canine Visceral Leishmaniasis. *PLoS Negl Trop Dis*. 2011; 5: e1310.
70. Sharma P, Gaur SN, Arora N. In silico Identification of IgE-Binding Epitopes of Osmotin Protein. *Plos One*. 2013; 8(1).
71. Casal JI, Langeveld JP, Cortes E, Schaaper WW, van Dijk E, Vela C, Kamstrup S, Melen R H. Peptide vaccine against canine parvovirus: identification of two neutralization subsites in the N terminus of VP2 and optimization of the amino acid sequence. *J Virol*. 1995; 69(11): 7274–7277.
72. Noya O, Patarroyo ME, Guzman F, Alarcon de Noya B. Immunodiagnosis of parasitic diseases with synthetic peptides. *Curr Protein Pept Sci*. 2003; 4(4): 299–308.

4. RESULTADOS

Artigo Original

Epitopos de proteínas antigênicas de *Leishmania infantum*: uma análise *in silico*

Epitopes of antigenic proteins in *Leishmania infantum*: an *in silico* analysis

LUCIANA M. ASSIS, JOÃO R. SOUSA, ABRÃO A. SILVA, ANTÔNIO FERNANDO DE MELO VAZ, PAULO P. ANDRADE, EDGAR M. CARVALHO, MARCIA A. MELO

Revista: **Parasite Immunology**

Situação: **submetido**

Artigo Original

Epitopos de proteínas antigênicas de *Leishmania infantum*: uma análise *in silico*

LUCIANA M. ASSIS¹, JOÃO R. SOUSA², ABRÃO A. SILVA³, ANTÔNIO FERNANDO M. VAZ⁴, PAULO P. ANDRADE⁵, EDGAR M. CARVALHO⁶, MARCIA A. MELO⁷

¹*Pos-Graduação em Medicina e Saúde, Complexo Hospitalar Universitário Professor Edgard Santos, Universidade Federal da Bahia, Salvador, Bahia, Brasil.*

Unidade Acadêmica de Enfermagem, Universidade Federal de Campina Grande-UFCG, Cajazeiras, Paraíba, Brasil.

²*Faculdade Santa Maria, Cajazeiras, Paraíba, Brasil.*

³*Unidade Acadêmica de Ciências Exatas e Naturais, Universidade Federal de Campina Grande, Cajazeiras, Paraíba, Brasil.*

⁴*Unidade Acadêmica de Medicina Veterinária, Universidade Federal de Campina Grande-UFCG, Patos, Paraíba, Brasil.*

⁵*Departamento de Genética, Universidade Federal de Pernambuco-UFPE, Recife, Pernambuco, Brasil.*

⁶*Pos-Graduação em Medicina e Saúde; Serviço de Imunologia, Complexo Hospitalar Universitário Professor Edgard Santos, Universidade Federal da Bahia, Salvador, Bahia, Brasil*

Instituto Nacional de Ciência e Tecnologia de Doenças Tropicais- INCT DT(CNPq/MCT), Salvador, Bahia, Brasil.

⁷*Pos-Graduação em Medicina Veterinária; Unidade Acadêmica de Medicina Veterinária, Universidade Federal de Campina Grande-UFCG, Patos, Paraíba, Brasil.*

RESUMO

Epitopos de célula B são sequências de aminoácidos de um antígeno reconhecidas por anticorpos solúveis e de membrana de célula B. A identificação de epitopos de célula B em proteínas é importante na aplicação de imunodiagnóstico e controle de doenças por vacinas. O objetivo deste estudo foi identificar *in silico* epitopos lineares e conformacionais de célula B a partir das proteínas cisteína peptidase calpaína-like, redutase tiol dependente 1 (TDR1) e HSP70 de *Leishmania infantum*; selecionar os epitopos lineares comuns aos métodos de predição e verificar a estrutura secundária dessas proteínas. A identificação de epitopos lineares foi realizada pelos programas de predição IEDB, BepiPred e BcePred. A sobreposição dos epitopos preditos por esses três programas gerou um total de 148 epitopos lineares a partir das três proteínas. Os epitopos conformacionais foram identificados pelo servidor CBtope. Foram preditos 164 epitopos conformacionais, a maioria localizada na região dos epitopos lineares. A predição da estrutura secundária foi realizada pelo PHD. Nas proteínas TDR1 e HSP70, os epitopos preditos estão localizados em regiões de alfa hélices e randômica. Novos epitopos lineares e conformacionais de célula B de proteínas de *L. infantum* foram identificados *in silico* e, baseado no conhecimento prévio da imunogenicidade da HSP70, a predição por vários programas garante uma maior acurácia dos resultados.

Palavras-chave: Bioinformática, Kalazar, Biologia molecular, Biologia sintética, Vacina.

ABSTRACT

B-cell epitopes are amino acid sequences of an antigen recognised by soluble antibodies and B-cell membranes. Identifying B-cell epitopes in proteins is important for applying immunodiagnostics and for disease control by vaccines. This study aimed to identify *in silico* linear and conformational B-cell epitopes from the calpain-like cysteine peptidase (CP), thiol-dependent reductase 1 (TDR1) and HSP70 proteins of *Leishmania infantum*, to select the common linear epitopes to different methods of prediction and to verify the secondary structure of these antigen proteins. The identification of linear epitopes were performed by Immune Epitope Database and Analysis Resource (IEDB), BepiPred and BcePred prediction programs. The overlap of predicted epitopes by these three programs generated a total of 148 linear epitopes from the three proteins. The conformational epitopes were identified using the CBTope server. In total, 164 conformational epitopes were predicted, mostly located in the linear epitope region. The predicted epitopes are located in α helix and random coil regions in the TDR1 and HSP70 proteins. New linear and conformational B-cell epitopes of *L. infantum* proteins were identified *in silico*, and the prediction using various programs ensures greater accuracy of the results based on prior knowledge of the immunogenicity of HSP70.

Keywords: Bioinformatics, Kalazar, Molecular biology, Synthetic biology, Vaccine.

Introdução

A identificação de regiões antigênicas/imunogênicas em antígenos protéicos tem uma importante aplicação no diagnóstico e no controle de doenças infecciosas através de vacinas. Por outro lado, a produção e uso de antígenos nativos para testes diagnósticos e ensaios vacinais é dificultada pela variação das espécies, além de implicar em custos elevados para a purificação, sobretudo na produção de vacinas [1]. Uma alternativa para resolver essas limitações é o uso de peptídeos sintéticos construídos a partir de epitopos identificados com ferramentas computacionais para predição de peptídeos antigênicos e/ou imunogênicos [2,3,4].

Epitopos de células B são partes de um antígeno proteico reconhecido por anticorpos e receptores de membrana de células B específicos [5]. Os epitopos de proteínas antigênicas podem ser lineares (contínuos) ou conformacionais (descontínuos). Os epitopos lineares são pequenas sequências contínuas de resíduos de aminoácidos do antígeno, enquanto que os epitopos conformacionais são compostos por resíduos de aminoácidos de sequências descontínuas que se encontram juntos no antígeno nativo pelo dobramento da proteína [6].

O primeiro método (Hopp e Woods, 1981) para predição de epitopo linear foi baseado nas propriedades dos aminoácidos [7], pressupondo que regiões hidrofílicas das proteínas são predominantemente localizadas na superfície, sendo potencialmente antigênicas. Posteriormente, Parker et al. (1986)[8] usaram a escala hidrofílica modificada e verificaram uma melhora na predição de epitopos quando comparado ao método de Hopp e Woods [7]. Outros trabalhos de predição de epitopos de célula B foram publicados baseados nas propriedades físico-químicas dos aminoácidos tais como flexibilidade [9] e acessibilidade ao solvente [10]. Kolaskar e Tongaonkar (1990) [11] usaram o parâmetro antigenicidade para prever sítios antigênicos em proteínas e observaram uma acurácia do método em torno de 75%. Posteriormente, Pellequer et al. (1993) [12] verificaram uma correlação entre sítios antigênicos e predição de voltas (turns) na proteína.

Os vários métodos de predição de epitopos lineares de célula B que utilizam escalas de propensão de aminoácidos, baseadas nas propriedades físico-químicas, foram avaliados quanto à taxa de sucesso. Uma comparação de 12 escalas aplicadas a 85 epitopos lineares identificados em 14 proteínas demonstraram que a maioria das escalas apresentou de 50 a

62% de predições corretas [13]; a escala *B-turn* teve o melhor desempenho, predizendo corretamente 70% dos epitopos conhecidos [12].

O programa BcePred é um dos programas para predição de epitopos lineares de célula B [14]. Este programa reúne sete diferentes escalas físico-químicas dos métodos de predição escalares existentes que podem ser usadas individualmente ou combinadas. As análises do BcePred usando as propriedades escalares combinadas, tais como hidrofobicidade [8], flexibilidade [9], polaridade [15] e superfície exposta [16], demonstraram uma melhor acurácia (58,7%) quando comparado a outras combinações e com o uso de escala única [14].

Na última década, métodos mais sofisticados baseados em modelos de aprendizagem de máquinas foram desenvolvidos na tentativa de melhorar a acurácia de predição de epitopos [17]. O programa BepiPred [18] é baseado na combinação do modelo oculto de Markov com dois métodos escalares: escala de hidrofobicidade de Parker [8] e escala de estrutura secundária de Levitt [19]. Esse programa demonstrou que o desempenho de escalas combinadas é melhor do que o desempenho de apenas uma única escala. Atualmente, o método BepiPred está incorporado no programa IEDB (Base de Dados de Epitopos Imunes e Ferramentas de Análise) [20] que é uma técnica em fase crescente de utilização [21].

A ferramenta preditiva ABCpred utiliza rede neural artificial e quando aplicado obteve uma acurácia máxima de 66% [22]. De forma geral, a acurácia dos preditores de epitopos lineares, que usam múltiplas escalas de aminoácidos ou sofisticadas técnicas de aprendizagem, é de aproximadamente 60 a 66% [23].

Novas metodologias *in silico* de predição de epitopos de célula B também foram desenvolvidas usando modelos vetores de suporte de máquinas (VSM). Os VSM são uma classe de métodos de aprendizagem de máquinas supervisionada usadas para classificação e regressão. As ferramentas que empregam essas metodologias incluem BCPred [24], COBEpro [25] e os mais recentes métodos baseados em sequência, CBtope [26] e BEST [27]. O preditor CBtope foi proposto para predição de epitopos conformacionais, e utiliza múltiplas propensões, com acurácia de 86,59%.

Vários estudos vêm sendo realizados usando peptídeos sintéticos que imitam epitopos no diagnóstico de doenças infecciosas e parasitárias [28], alergias [29], doenças autoimunes e cânceres [30,31]. O primeiro passo no desenho de peptídeos sintéticos para aplicação em vacinas e no imunodiagnóstico das doenças é identificar potenciais epitopos. A combinação de métodos e uma análise dos resultados podem reduzir o número de epitopos potencialmente

úteis e conseqüentemente otimizar recursos e esforços para a identificação *in vitro* e *in vivo* de epitopos efetivamente úteis.

Este trabalho descreve e analisa a predição de epitopos de células B de proteínas antigênicas de *L. infantum* através de ferramentas de bioinformática.

Metodologia

Origem das seqüências proteicas

Os genes correspondentes às proteínas analisadas neste estudo foram identificados por immunoscreening de uma biblioteca de DNA complementar (cDNA) de *Leishmania chagasi*, empregando soro de humanos com leishmaniose visceral, realizado em um trabalho anterior. As seqüências nucleotídicas parciais foram obtidas pelo sequenciamento do cDNA correspondente.

Identificação dos genes em *L. infantum*, correspondentes aos cDNAs seqüenciados

As seqüências parciais dos cDNAs foram submetidas à ferramenta BLAST (Basic Local Alignment Search Tool - http://www.genedb.org/blast/submitblast/GeneDB_Linfantum) no GeneDB (www.genedb.org) para identificação das seqüências gênicas completas no genoma de *L. infantum* JPCM5. As seqüências de aminoácidos deduzidas de cada gene foram armazenadas no formato FASTA para posterior utilização nos programas de predição de epitopos.

Foram analisadas três proteínas de *L. infantum* denominadas: cisteína peptidase (putative calpain-like cysteine peptidase, GI:146090707), redutase tiol dependente 1 (thiol-dependent reductase 1, GI:146097469) e proteína de choque térmico de 70 kDa - HSP70 (heat-shock protein, GI:758136.1), com 6168, 450 e 653 aminoácidos (aa), respectivamente.

A proteína HSP70 foi utilizada como um controle da predição pela presença de epitopos imunodominantes identificados experimentalmente (Quijada et al., 1996)[32].

Predição de epitopos lineares e conformacionais de célula B

Para determinação de epitopos lineares, as sequências proteicas de *L. infantum* foram submetidas aos servidores Web IEDB[20] BepiPred [18] e BcePred[14]. O IEDB e o BcePred são baseados em métodos de escala de propensão de acordo com as propriedades físico-químicas dos aminoácidos. O método BepiPred, também uma das ferramentas do IEDB, é uma combinação do modelo oculto de Markov com dois métodos escalares: hidrofiliçidade de Parker[8] e estrutura secundária de Levit[19].

Os métodos de predição de epitopos de célula B, que são comuns aos programas IEDB e BcePred utilizados neste estudo, foram baseados nas escalas: hidrofiliçidade[8], flexibilidade[9] e acessibilidade ao solvente[10]. Os epitopos de célula B lineares que foram comuns, parcialmente ou totalmente, aos métodos escalares de hidrofiliçidade, flexibilidade e acessibilidade do IEDB e do BcePred foram comparados com os resultados do BepiPred. Os epitopos lineares comuns aos três programas foram selecionados para verificar a frequência dos resíduos de aminoácidos.

Os métodos escalares que fazem parte dos programas IEDB e BcePred utilizam nos cálculos de predição uma janela móvel composta de 6-7 aminoácidos. Segundo Manzalawy et al. (2008), esses métodos realizam uma varredura da sequência de aminoácidos usando uma janela móvel e atribui um escore ao aminoácido central da janela que corresponde a média dos valores escalares dos aminoácidos da janela [33]. O resíduo de aminoácido central é predito positivo se seu escore for maior do que um limiar predeterminado e esse escore reflete a probabilidade do aminoácido fazer parte de um epitopo.

O programa CBtope foi usado para predição de epitopos conformacionais de célula B [26]. Esse programa utiliza uma janela de 19 resíduos de aminoácidos para analisar a região do epitopo conformacional [26].

As configurações padrão foram aplicadas a todas as ferramentas usadas.

Predição da estrutura secundária

A estrutura secundária das proteínas foi predita através do servidor PHD (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_phd.html).

Resultados

Predição de epítopos de célula B

Os programas IEDB, BcePred e BepiPred identificaram 130, 103 e 437 epítopos lineares, respectivamente, obtidos a partir das três proteínas, cisteína peptidase (CP), redutase tiol dependente 1 (TDR1) e proteína de choque térmico (HSP70). A quantidade de epítopos e o tamanho das proteínas estão mostrados na tabela 1.

Tabela 1- Quantidade de epítopos lineares preditos por cada programa a partir das proteínas de *L. infantum* cisteína peptidase (CP), redutase tiol dependente1 (TDR1) e proteína de choque térmico (HSP70)

Proteínas	Total AA	IEDB*		BcePred*		BepiPred	
		n	%	n	%	n	%
CP	6168	114	87,7%	84	81,6%	382	87,4%
TDR1	450	6	4,6%	5	4,8%	19	4,4%
HSP70	653	10	7,7%	14	13,6%	36	8,2%
TOTAL	7271	130	100%	103	100%	437	100%

*Apenas os epítopos comuns aos métodos escalares: hidrofobicidade, flexibilidade e acessibilidade.

Através da sobreposição dos programas do IEDB, BepiPred e BcePred, foram preditos 357 epítopos lineares para a proteína CP, dos quais 130 foram comuns, parcialmente ou totalmente, aos três programas. Dos 19 epítopos da proteína TDR1 preditos pelos três métodos (Tabela 1), apenas quatro foram comuns a todos os programas, conforme demonstrado na tabela 2.

Tabela 2 - Epítopos lineares da proteína TDR1 de *L. infantum* preditos pelo IEDB, BepiPred e BcePred.

PROGRAMAS	POSICÃO	SEQUÊNCIA	EPÍTOPO
1;2;3 ^a	39-54	REEMPQWYKQINPRET	Pt1
1;2;3	55-70	VPTLEVGNADKRFMFE	Pt2
1;2;3	195-214	AAQRASVRETSPTAAQCIEN	Pt3
1;2;3	310-324	ALVPRGDAEKEYEYEVG	Pt4

^aIEDB (1), BepiPred (2), BcePred (3).

Para a HSP70, dos 37 epítopos preditos, 14 foram comuns aos três programas (Tabela 3).

Tabela 3 - Epitopos lineares da proteína HSP70 de *L. infantum* preditos pelos programas IEDB, BepiPred e BcePred

PROGRAMAS	POSIÇÃO	SEQUÊNCIA	EPITOPO
1;2;3 ^a	28-46	IIANDQGNRTTPSYVAFTD	Ph1
1;2;3	105-121	SVQYRGEEKTFTPEEIS	Ph2
1;2;3	149-165	AYFNDSQRQATKDAGTI	Ph3
1;2;3	183-197	YGLDKGDDGKERNVL	Ph4
1;2;3	241-261	FFTEEFKRKNKGKNLASSHRA*	Ph5
1;2;3	323-335	LQDAKMDKRSVHD	Ph6
1;2;3	380-393	FILTGGKSKQTEGL	Ph7
1;2;3	488-506	LNVSAAEEKGTGKRNQITIT	Ph8
1;2;3	504-524	TITNDKGRLSKDEIERMVNDA**	Ph9
1;2;3	525-545	MKYEEDDKAQRDRVEAKNGLE***	Ph10
1;2;3	546-564	NYAYSMKNTLSDSNVSGKL	Ph11
1;2;3	558-576	SNVSGKLVSGKLESDKATLNKEI	Ph12
1;2;3	581-599	EWLSSNQEAAKEEYEHKQK	Ph13
1;2;3	587-605	QEAAKEEYEHKQKELESVC	Ph14

^a IEDB (1), BepiPred (2), BcePred (3); *Epitopo imunodominante e espécie específico H17; resíduos de aminoácidos presentes nos epitopos H35** e H36*** (QUIJADA et al., 1996).

O epitopo Ph5 (aa 241-260) possui todos os resíduos de aminoácidos do epitopo H17 (FFTEEFKRKNKGKNLASSHR) identificado como espécie específico por Quijada et al., (1996). Os epitopos Ph9 e Ph10 possuem a sequência parcial dos resíduos de aminoácidos dos epitopos H35 e H36 que também foram imunogênicos e identificados pelo mesmo autor. Os epitopos Ph2, Ph4, Ph7 e Ph8 não foram indicados como imunogênicos no ensaio do Quijada et al., (1996) (tabela 3).

No programa CBtope foram preditos 143, 11 e 10 regiões de epitopos conformacionais para as proteínas CP, TDR1 (tabela 4) e HSP70 (tabela 5). Dos 143 epitopos conformacionais da CP, 67 contêm, de forma total ou parcial, resíduos de aminoácidos de epitopos lineares. Já para a proteína TDR1, três epitopos lineares (Pt1, Pt3 e Pt4) estão incluídos, total ou parcialmente, entre os onze epitopos conformacionais preditos (tabela 4). Para a proteína HSP70, dois epitopos são exclusivamente conformacionais e oito contêm resíduos de aminoácidos de epitopos lineares, parcialmente ou totalmente (tabela 5).

Tabela 4 - Epitopos conformacionais de célula B da proteína TDR1 (GI:146097469) de *L. infantum* preditos pelo CBtope.

Nº de resíduos epitopos ^a	Posição	Sequência	Epitopo Conformacional (Pect)
36 (43)	16-57	FCHRVEIVAREKQVS YDRVAVGL REEMPQWYK QINPRETVPT*	Pect1
21 (24)	117-140	LRDPLSGEKRKAMDDNAAAYVDGLL	Pect2
6 (09)	161-169	ALVPFLVRL	Pect3
10 (12)	208-219	AAQCIENYRHLV	Pect4
28 (44)	238-281	LFCPFVDRARLASELRKFQM HIVEVPLHPQPEWY KYINPRDTPV	Pect5
5 (09)	293-301	ESQLIVQYI	Pect6
9 (11)	311-321	LVPRGDAEKEY	Pect7
4 (04)	374-377	FGGK	Pect8
9 (10)	389-399	LVRAKAFMPE	Pect9
11 (18)	413-430	LNGLAEAGMATPEAKSVF	Pect10
6 (06)	445-450	RRAQSG	Pect11

^atamanho da região do epitopo conformacional está entre parêntese; * os aminoácidos em negrito fazem parte dos epitopos lineares da mesma região.

Tabela 5 - Epitopos conformacionais de célula B da proteína HSP70 (GI:758136) de *L. infantum*, preditos pelo CBtope.

Número de resíduos epitopos ^a	Posição	Sequência	Epitopo Conformacional (Pec)
17 (20)	21-40	WQNERVDIIAND QGNRTTPS*	Pech1
22 (35)	48-82	ERLIGDAAKNQVAMNPHNTVFD AKRLIGRKFN DSV	Pech2
15 (29)	136-164	LGKQVKKAVVTVPAYF NSQRQATKDAGT	Pech3
27 (38)	321-359	RVLQDAKMDKRSVHDVVLVGGSTRIPKVQSL VSDFFGGK	Pech4
6 (06)	381-386	ILTGGK	Pech5
21 (23)	427-449	SQIFSTYADNQPGVHIQVFEGE R**	Pech6
21 (32)	476-507	IEVTFDL DANGILNVS AEEKGTGKR NQITITN	Pech7
4 (04)	539-542	EAKN	Pech8
11 (15)	561-575	SGKLESDKATLNKE	Pech9
5 (07)	603-609	SVCNPIM	Pech10

^atamanho da região do epitopo conformacional está entre parêntese; * os aminoácidos em negrito fazem parte dos epitopos lineares da mesma região; ** A sequência peptídica NQPGVHIQVFEGE corresponde ao epitopo imunodominante H30 (Quijada et al. 1996).

Comparando os epítomos lineares e conformacionais, sobrepostos ou não, percebe-se que a análise da HSP70 encontrou sete epítomos descritos como imunogênicos por Quijada et al. (1996)[32]. Outros sete estavam no limite do cut off estipulado pelos autores e apenas dois foram preditos erroneamente pelos programas (Tabela 6).

Tabela 6 – Sinopse dos epítomos lineares e conformacionais que se sobrepõem ou não na proteína HSP70 de *Leishmania infantum*.

	Tipo de epítomo			Total
	Linear/conformacional	Linear	conformacional	
Positivo	Ph1/Pech1; Ph3/Pech3 Ph14/Pech10	Ph5; Ph9	Pech2; Pech6	7
Limite	Ph6/Pech4; Ph8/Pech7; Ph10/Pech8; Ph12/Pech9	Ph4, Ph11; ph13		7
Negativo	Ph7/Pech5	Ph2		2
Total	8	6	2	16

Predição da estrutura secundária

As estruturas preditas pelo programa PHD foram alfa hélice (α -hélice), fita estendida e randômica (Figura 1). Para as três proteínas não houve predição de folha beta.

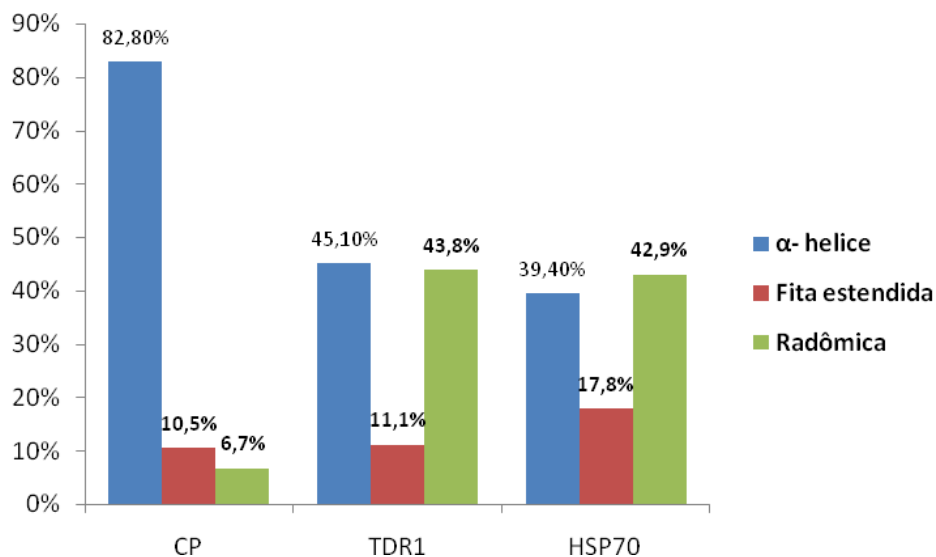


Figura 1. Frequência dos tipos de estrutura secundária predita pelo programa PHD a partir das proteínas CP, TDR1 e HSP70 de *L. infantum*.

Os resíduos do epítipo imunodominante H17 da HSP70 (Quijada et al., 1996)[32], que corresponde ao epítipo linear Ph5, estão localizados na região da estrutura secundária α -hélice (241-249) e randômica (250-260) dessa proteína, conforme ilustrado na figura 2.

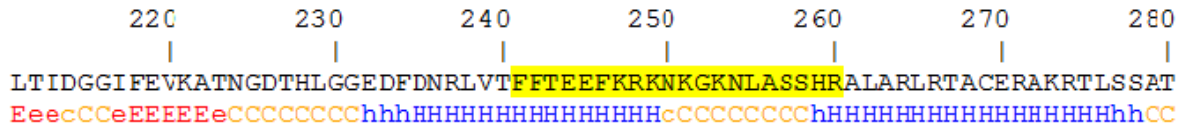


Figura 2- Localização do epítipo imunodominante H17 (Quijada et al., 1996) destacado em amarelo e a região da estrutura secundária correspondente. Hh (α - hélice), Cc (randômica).

De acordo com a figura 1, a TDR1 tem uma frequência semelhante em quantidades de α - hélices (45,1%) e randômica (43,8%). Avaliando a localização dos epítipos lineares e conformacionais, percebe-se que estão localizados principalmente em α - hélices e randômica (Figuras 3 e 4).



Figura 3- Sobreposição dos epítipos lineares preditos pelos programas IEDB, BepiPred e BcePred na estrutura secundária da TDR1 predita pelo PHD. Hh (α - hélice), Cc (randômica), Ee (fita estendida)

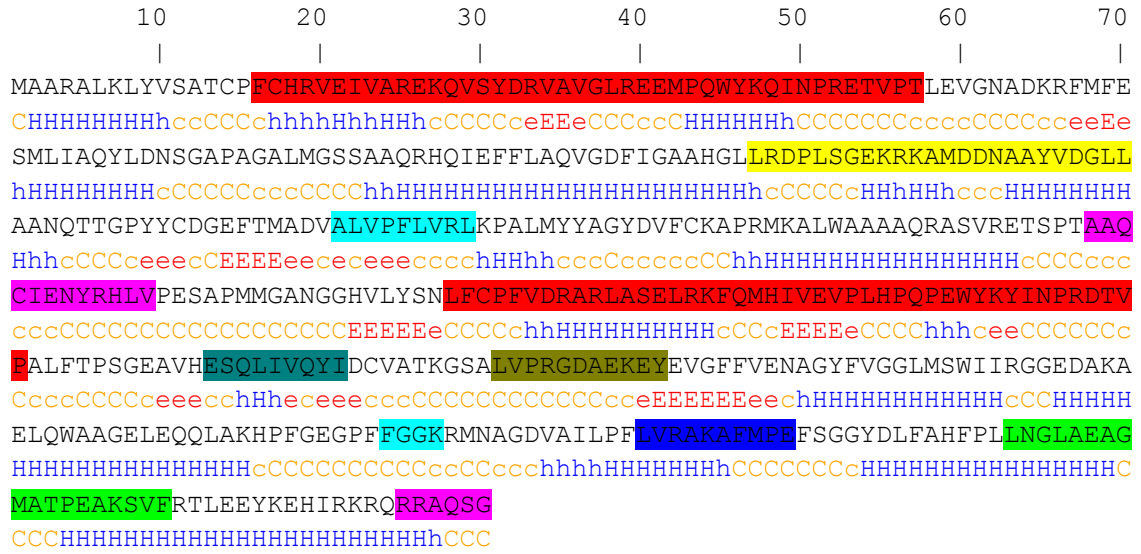


Figura 4- Sobreposição dos epítopos conformacionais preditos pelo CBTope na estrutura secundária da TDR1 predita pelo PHD. Hh (α- hélice), Cc (randômica), Ee (fita estendida).

A frequência dos resíduos de aminoácidos nos epítopos lineares comuns aos programas IEDB, BepiPred e BcePred das proteínas CP, TDR1 e HSP70 está ilustrada na Figura 5. Observa-se que os epítopos lineares são constituídos principalmente pelos aminoácidos E, A, K, R, D, L, e T.

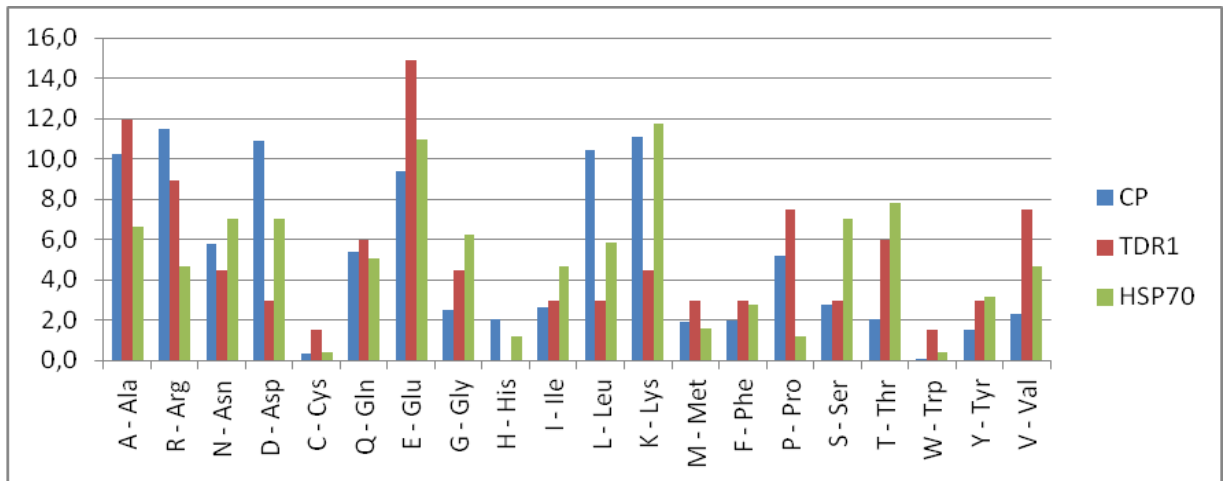


Figura 5 - Frequência dos resíduos de aminoácidos dos epítopos lineares das proteínas CP, TDR1 e HSP70 preditos pelos programas IEDB, BepiPred e BcePred.

Discussão

A identificação de epitopos a partir de programas de bioinformática tem sido utilizada em imunodiagnóstico [34], desenho de vacinas [35] e imunoterapia [36]. Entretanto, até o momento, não havia sido realizada a predição de epitopos de célula B presentes nas proteínas de *L. infantum*, aqui estudadas, empregando-se os programas de predição IEDB, BepiPred, BcePred e CBtope.

Através do IEDB, BepiPred e BcePred, 14, 04 e 130 epitopos lineares foram preditos pelos três programas para as proteínas HSP70, TDR1 e CP, respectivamente.

Neste estudo, a proteína HSP70 foi utilizada como um controle de predição, uma vez que seus epitopos já foram testados experimentalmente por Wallace et al. (1992)[37] e Quijada et al. (1996)[32]. Os três programas foram capazes de prever toda a região do epitopo H17 (241-260) da HSP70 de *L. infantum*. De acordo com Quijada et al. (1996), o peptídeo H17 é um epitopo de célula B imunodominante e espécie-específico, pois é divergente da região equivalente na HSP70 de *Trypanosoma cruzi* [32]. Adicionalmente, foi predito pelos programas um epitopo localizado na região carboxi-terminal (Met-525 a Glu-545) identificado aqui como Ph10, que inclui alguns resíduos de aminoácidos da sequência imunodominante (EADDRA) descrita por Wallace et al. (1992)[37]. Entre os 14 epitopos lineares preditos, cinco estão acima do cut off estabelecido por Quijada et al. (1996)[32]. Adicionalmente, sete estão no limite ou discretamente abaixo do cut off. Apenas dois epitopos preditos não foram equivalentes aos peptídeos imunogênicos identificados pelos autores citados acima (Tabela 6).

Os epitopos Pech2 e Pech6, exclusivamente conformacionais, coincidem com mais cinco peptídeos descritos por Quijada et al. (1996)[32]. O Pech6 contém 14 aa (NQPGVHIQVFEGGER) dos 20 que fazem parte do peptídeo imunodominante H30, descrito pelos autores citados (Tabela 6).

Associando-se os resultados de predição de epitopos lineares e conformacionais, 7 epitopos são claramente imunogênicos, 7 estão no limite do cut off determinado por Quijada et al., (1996)[32] e apenas dois epitopos foram falsamente preditos. Considerando os epitopos lineares e conformacionais, nota-se claramente que esses estão concentrados na região carboxi-terminal da HSP70, que é reconhecidamente a região mais divergente da proteína (ANDRADE et al. 1992)[38]. Esses resultados indicam que as predições podem ser mais

fidedignas quando as ferramentas são usadas de forma combinada na identificação de epitopos.

Utilizando a mesma metodologia, entre os epitopos conformacionais preditos, 46,8% e 27,3% contêm resíduos de aminoácidos de epitopos lineares das proteínas CP e TDR1, respectivamente.

Os aminoácidos arginina (R), ácido glutâmico (E) e lisina (K) estão entre os mais frequentes nas proteínas CP, TDR1 e HSP70, respectivamente. Estes aminoácidos contêm cadeias laterais com grupos funcionais básicos (arginina, lisina) e ácidos (ácido glutâmico) que permitem formar ligações iônicas (ligações eletrostáticas) e pontes de hidrogênio [39], interações que também ocorrem quando há o reconhecimento do epitopo pelo anticorpo. Esses aminoácidos também são os resíduos com maiores valores escalares de acessibilidade ao solvente [10]. A acessibilidade dos resíduos de epitopos ao solvente é um fator importante para a imunogenicidade de um antígeno ou peptídeo sintético [40]. Esses achados indicam que os epitopos lineares de célula B preditos neste estudo, simultaneamente pelos três programas, são bons candidatos a serem testados em ensaios de diagnósticos e vacinais da leishmaniose visceral.

O conhecimento da estrutura secundária é necessário para a predição da estrutura terciária. Neste estudo, através do programa PHD, pode se verificar que as proteínas analisadas são constituídas de α -hélices, fitas estendidas e randômica, com uma predominância da primeira para as proteínas CP e TDR1. Há uma forte tendência de determinados aminoácidos em formar um tipo específico de estrutura secundária; nesse caso, os resíduos glutamato (E), alanina (A), leucina (L), metionina (M), glutamina (Q), lisina (K), arginina (R) e histidina (H) são preferencialmente encontrados em alfa hélices [41]. De acordo com a figura 1, a HSP70 e a TDR1 são formadas por 39,4% e 45,1% de α -hélice, respectivamente, que está de acordo com a composição de aminoácidos (Figura 5).

Para a proteína TDR1, a maioria dos resíduos de aminoácidos que compõem os epitopos lineares está localizada na estrutura secundária de α -hélices e randômica (*random coil*). Na HSP70, o epitopo imunodominante H17 está localizado em uma região de α -hélice e randômica.

Em nenhuma das proteínas foi predita as estruturas secundárias do tipo alças (*loop*) e voltas (*turn*). As estruturas secundárias alças e voltas, presentes na superfície das proteínas, podem participar de interações com outras moléculas. Conforme observado por Rose et al.

(1985) a ocorrência de voltas sobre a superfície exposta ao solvente torna essa estrutura acessível ao reconhecimento molecular e a interações entre peptídeos e receptores [43]. Entretanto, como já discutido por Smith e Pease (1980), a presença de voltas não torna a proteína mais antigênica [44], que é corroborado com a predição da HSP70. Sabe-se que essa proteína é altamente antigênica e imunogênica e composta por α hélice, fita estendida e randômica.

A identificação de epitopos através da bioinformática ainda apresenta limitações, sendo necessários estudos que melhorem a acurácia dos métodos de predição de epitopos de célula B. Segundo Greenbaum et al. (2007), uma maior acurácia pode ser obtida melhorando a qualidade dos atuais bancos de dados, que contêm epitopos delineados erroneamente [45]. Isso é importante, uma vez que os métodos de predição utilizam bancos de dados de epitopos para avaliar a eficiência do método. Como os resultados de predição produzidos por vários métodos podem ser distintos [40] é interessante usar múltiplas ferramentas para conseguir resultados mais coerentes e precisos, que foi a abordagem empregada neste trabalho. Posteriormente, os dados de bioinformática precisam ser corroborados por ensaios laboratoriais para que as informações obtidas alimentem os bancos de dados de forma mais fidedigna.

A maioria dos epitopos de célula B nas proteínas são conformacionais ou descontínuos [45], entretanto, a maioria dos métodos de predição estão disponíveis para identificação de epitopos lineares ou contínuos. O principal entrave é a necessidade do conhecimento da estrutura do antígeno para a predição de epitopos conformacionais e os bancos de dados de modelos de proteínas ainda são limitados na quantidade de estruturas terciárias e quaternárias disponíveis. Apesar disso, a identificação de epitopos lineares tem demonstrado resultados promissores em pesquisas na identificação de antígenos vacinais e de alérgenos, e também no imunodiagnóstico da leishmaniose [32, 34].

Conclusão

Novos epitopos lineares e conformacionais de célula B de proteínas de *L. infantum* foram identificados *in silico* e, baseado no conhecimento prévio da imunogenicidade da HSP70, a predição por vários programas garante uma maior acurácia dos resultados. E a associação desses resultados a avaliações de estrutura secundária e terciária permitem a identificação de epitopos em antígenos com elevada acurácia.

Conflito de interesse

Os autores declaram que não há conflito de interesse.

Agradecimentos

Os autores agradecem ao Instituto Nacional de Ciência e Tecnologia de Doenças Tropicais (INCT-DT), MCT/CNPQ, que foi a fonte financiadora de parte dos estudos.

Referências

1. List C, Qi W, Maag E, Gottstein B, Müller N, Felger I. Serodiagnosis of *Echinococcus* spp. infection: explorative selection of diagnostic antigens by peptide microarray. *PLoS Negl Trop Dis* 2010; **4**: e771.
2. Van Regenmortel MH. Immunoinformatics may lead to a reappraisal of the nature of B cell epitopes and of the feasibility of synthetic peptide vaccines. *J Mol Recognit* 2006; **19**: 183-187.
3. Gomara MJ, Haro I. Synthetic peptides for the immunodiagnosis of human diseases. *Curr Med Chem* 2007; **14**: 531-546.
4. Chen P, Rayner S, Hu KH. Advances of bioinformatics tools applied in virus epitopes prediction. *Virol Sin* 2011; **26**: 1-7.
5. Peters B, Sidney J, Bourne P, *et al.* The design and implementation of the immune epitope database and analysis resource. *Immunogenetics* 2005; **57**: 326-336.
6. Barlow DJ, Edwards MS, Thornton JM, *et al.* Continuous and discontinuous protein antigenic determinants. *Nature* 1986; **322**: 747-748.
7. Hopp TP, Woods KR. Prediction of protein antigenic determinants from amino acid sequences. *Proc Natl Acad Sci USA* 1981; **78**: 3824-3828.
8. Parker JM, Guo D, Hodges RS. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry* 1986; **25**: 5425-5432.

9. Karplus PA, Schulz GE. Prediction of chain flexibility in proteins. A tool for the selection of peptide antigens. *Naturwissenschaften* 1985; **72**: 212-213.
10. Emini EA, Hughes JV, Perlow DS, Boger J. Induction of hepatitis A virus: neutralizing antibody by a virus-specific synthetic peptide. *J Virol* 1985; **55**: 836-839.
11. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett* 1990; **276**: 172-174.
12. Pellequer JL, Westhof E, Van Regenmortel MH. Correlation between the location of antigenic sites and the prediction of turns in proteins. *Immunol Lett* 1993; **36**: 83-99.
13. Pellequer JL, Westhof E, Van Regenmortel MH. Predicting location of continuous epitopes in proteins from their primary structures. *Methods Enzymol* 1991; **203**: 176-201.
14. Saha S, Raghava GP. BcePred: prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties. *Artificial Immune Systems, Third International Conference (ICARIS 2004), LNCS*, 2004; **3239**: 197-204.
15. Ponnuswamy PK, Prabhakaran M, Manavalan P. Hydrophobic packing and spatial arrangements of amino acid residues in globular proteins. *Biochim Biophys Acta* 1980; **623**: 301-316.
16. Janin J, Wodak S. Conformation of amino acid side-chains in proteins. *J Mol Biol* 1978; **125**: 357-386.
17. Söllner J, Mayer B. Machine learning approaches for prediction of linear B-cell epitopes on proteins. *J Mol Recognit* 2006; **19**: 200-208.
18. Larsen JE, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. *Immunome Res* 2006; **2**: 2.
19. Levitt M. Conformational preferences of amino acids in globular proteins. *Biochemistry* 1978; **17**: 4277-4285.
20. Peters B, Sidney J, Bourne P, *et al.* The immune epitope database and analysis resource: from vision to blueprint. *PLoS Biol* 2005; **3**: e91.
21. Salimi N, Fleri W, Peters B, Sette A. Design and utilization of epitope-based databases and predictive tools. *Immunogenetics* 2010; **62**: 185-196.
22. Saha S, Raghava GP. Prediction of continuous B-cell epitopes in the antigen using recurrent neural network. *Proteins* 2006; **65**: 40-48.

23. Ponomarenko JV, Van Regenmortel MHV. B cell epitope prediction. In *Structural Bioinformatics*, 2nd ed, eds. Gu J, Bourne PE, Hoboken, NJ: John Wiley; 2009; 849-879.
24. El-Manzalawy Y, Dobbs D, Honavar V. Predicting linear B-cell epitopes using string kernels. *J Mol Recognit* 2008; **21**: 243-255.
25. Sweredoski MJ, Baldi P. COBEpro: a novel system for predicting continuous B-cell epitopes. *Protein Eng Des Sel* 2009; **22**: 113-120.
26. Ansari HR, Gajendra PS, Raghava GP. Identification of conformational B-cell epitopes in an antigen from its primary sequence. *Immunol Res* 2010; **6**: 6.
27. Gao J, Faraggi E, Zhou Y, Ruan J, Kurgan L. BEST: improved prediction of B-cell epitopes from antigen sequences. *PloS One* 2012; **7**: e40104.
28. Noya O, Patarroyo ME, Guzman F, Alarcon de Noya B. Immunodiagnosis of parasitic diseases with synthetic peptides. *Curr Protein Pept Sci* 2003; **4**: 299-308.
29. Eigenmann PA. Do we have suitable in-vitro diagnostic tests for the diagnosis of food allergy? *Curr Opin Allergy Clin Immunol* 2004; **4**: 211-213.
30. Fournel S, Muller S. Synthetic peptides in the diagnosis of systemic autoimmune diseases. *Curr Protein Pept Sci* 2003; **4**: 261-274.
31. Sundaram R, Dakappagari NK, Kaumaya PT. Synthetic peptides as cancer vaccines. *Biopolymers* 2002; **66**: 200-216.
32. Quijada L, Requema MJ, Soto M, *et al.* Mapping of the linear antigenic determinants of the *Leishmania infantum* Hsp70 recognized by leishmaniasis sera. *Immunol Lett* 1996; **52**: 73-79.
33. El-Manzalawy Y, Dobbs D, Honavar V. Predicting flexible length linear B-cell epitopes. *Comput Syst Bioinformatics Conf* 2008; **7**: 121-132.
34. Faria AR, Costa MM, Giusta MS, *et al.* High-throughput analysis of synthetic peptides for the immunodiagnosis of canine visceral leishmaniasis. *PLoS Negl Trop Dis* 2011; **5**: e1310. doi: 10.1371/journal.pntd.0001310.
35. Sirskyj S, Diaz-Mitoma F, Golshani A, Kumar A, Azizi A. Innovative bioinformatic approaches for developing peptide-based vaccines against hypervariable viruses. *Immunol Cell Biol* 2011; **89**: 81-89. doi:10.1038/icb.2010.65.

36. Beck A, Klinguer-Hamour C, Bussat MC, *et al.* Peptides as tools and drugs for immunotherapies. *J Pept Sci* 2007; **13**: 588-602.
37. Wallace GR, Ball AE, MacFarlane J, El Safi SH, Miles MA, Kelly JM. Mapping of a visceral leishmaniasis-specific immunodominant B-cell epitope of *Leishmania donovani* Hsp70. *Infect Immun* 1992; **60**: 2688-2693.
38. de Andrade CR, Kirchhoff LV, Donelson JE, Otsu K. Recombinant *Leishmania* Hsp90 and Hsp70 are recognized by sera from visceral leishmaniasis patients but not Chagas' disease patients. *J Clin Microbiol* 1992; **30**: 330-335.
39. Smith C, Marks AD, Liebermam M. Bioquímica médica básica de Marks [Basic medical biochemistry by Marks], 2nd ed. Porto Alegre: Artmed; 2007
40. Mahdavi M, Mohabaktar M, Keyhanfar M, Dehkordi AJ, Rabbani M. Linear and conformacional B Cell epitope prediction of the HER 2 ECD-subdomain III *in silico* methods. *Asian Pac J Cancer Prev* 2012; **13**: 3053-3059.
41. Berg JM, Tymoczko JL, Stryer L. Bioquímica [Biochemistry], 6th ed. Rio de Janeiro: Guanabara Koogan; 2008.
42. Rose GD, Gierasch LM, Smith JA. Turns in peptides and proteins. *Adv Protein Chem* 1985; **37**: 100-109.
43. Smith JA, Pease LG. Reverse turns in peptides and proteins. *Crit. Rev. Biochem* 1980;**8**
44. Greenbaum JA, Andersen PH, Blythe M, *et al.* Towards a consensus on datasets and evaluation metrics for developing B-cell epitope prediction tools. *J Mol Recognit* 2007; **20**: 75-82.
45. Novotn J, Handschumacher M, Haber E, *et al.* Antigenic determinants in proteins coincide with surface regions accessible to large probes (antibody domains). *Proc Natl Acad Sci USA* 1986; **83**: 226-230.

5. CONCLUSÕES

- Novos epitopos lineares e conformacionais de célula B foram identificados pela combinação de ferramentas de bioinformática a partir das proteínas cisteína peptidase calpaina-like, redutase thiol dependente 1 e HSP70 de *L. infantum*.
- Na predição de epitopos, vários métodos devem ser combinados para maior fidelidade dos resultados.
- As análises *in silico* das proteínas estudadas definiram epitopos lineares de célula B compostos com resíduos de aminoácidos com alta acessibilidade ao solvente e, portanto, com grande chance de fazerem parte do sítio de interação antígeno-anticorpo.
- A estrutura secundária das proteínas CP, TDR1 e HSP70 é composta por α -hélice, fita estendida e randômica. Os epitopos preditos a partir da TDR1 e HSP70, estão localizados em regiões da estrutura secundária alfa hélice e randômica.
- Os epitopos lineares de célula B preditos pelo IEDB, BepiPred e BcePred e os epitopos conformacionais preditos pelo CBtope são potenciais alvos no desenvolvimento do imunodiagnóstico da leishmaniose visceral e produção de anticorpos anti-peptídeos.

6. CONSIDERAÇÕES FINAIS

O estudo *in silico* de predição de epitopos contribui para a descoberta de novos antígenos na aplicação de testes diagnósticos mais sensíveis e específicos, desenho de vacinas e na imunoterapia. Contudo, ainda existem limitações quando a ferramenta de predição é baseada em sequências de aminoácidos; nesse caso, quando se deseja ter uma definição mais apurada da localização e acessibilidade dos resíduos de aminoácidos dos epitopos lineares e conformacionais, necessita-se conhecer a estrutura tridimensional das proteínas. Por outro lado, o passo inicial para identificar novos epitopos de forma mais confiável a partir de proteínas de estrutura desconhecida é através da combinação de mais de um método de predição de epitopos baseada em sequência, sobretudo, os métodos de predição de epitopos lineares baseados nas propriedades físico-químicas dos aminoácidos, como demonstrado neste estudo *in silico* de predição de epitopos de célula B de proteínas de *L.infantum*.

7. PERSPECTIVAS E PROPOSTAS DE ESTUDO

- Predizer a estrutura terciária das proteínas aqui estudadas para localizar os peptídeos preditos e verificar a acessibilidade ao solvente de todos os resíduos desses epitopos, bem como usar ferramentas computacionais que visualizem a estrutura tridimensional.
- Após definir a estrutura terciária das proteínas usar ferramentas de predição de epitopos conformacionais, baseados na estrutura, e comparar os resultados com os obtidos neste estudo.
- Realizar a predição de epitopos de célula T a partir das proteínas estudadas e comparar com os epitopos de célula B linear e conformacional e, por conseguinte, selecionar os possíveis epitopos antigênicos e imunogênicos que poderão ser usados na produção de anticorpos protetores em indivíduos expostos a *L. infantum*.
- Os epitopos preditos neste trabalho podem ser usados para síntese de peptídeos a serem testados *in vitro* e em estudos experimentais *in vivo* da leishmaniose visceral.

8. REFERÊNCIAS (INTRODUÇÃO)

1. Alvar J, Iván D. Vélez, Caryn Bern, Mercé Herrero, Philippe Desjeux, Jorge Cano, Jean Jannin, Margriet den Boer. Leishmaniasis Worldwide and Global Estimates of Its Incidence. *PLoS One*. 2012; 7(5):e35671.
2. Maurício II, Stohard Jr, Miles Ma. The strange case of *Leishmania chagasi*. *Parasitol Today*. 2000; 16: 188-189.
3. Michalsky EM, Rocha MF, da Rocha Lima AC, Francisca-Silva JC, Pires MQ, et al. Infectivity of seropositive dogs, showing different clinical forms of leishmaniasis, to *Lutzomia longipalpis* phlebotomine sand flies. *Vet Parasitol*. 2007; 147: 67–76.
4. Alvar J, Cañavate C, Molina R, Moreno J, Nieto J. Canine leishmaniasis. *Adv Parasitol*. 2004; 57:1-88.
5. Ponomarenko JV, Van Regenmortel MHV. B cell epitope prediction. In *Structural Bioinformatics*, 2nd ed, eds. Gu J, Bourne PE, Hoboken, NJ: John Wiley; 2009; 849-879.
6. Van Regenmortel MH. Immunoinformatics may lead to a reappraisal of the nature of B cell epitopes and of the feasibility of synthetic peptide vaccines. *J Mol Recognit* 2006; 19: 183-187.

9. ANEXOS

9.1. ANEXO A. Comprovante de submissão de artigo a periódico

21/07/13

Parasite Immunology

ScholarOne Manuscripts

Edit Account | Instructions & Forms | Log Out | [Get Help Now](#)

SCHOLARONE[™] Manuscripts

[Main Menu](#) → [Author Dashboard](#) → Submission Confirmation

You are logged in as Marcia Melo

Submission Confirmation

Thank you for submitting your manuscript to *Parasite Immunology*.

Manuscript ID: PIM-2013-0078

Title: B- cell epitopes of antigenic proteins in *Leishmania infantum*: an in silico analysis

Assis, Luciana
Sousa, João
Silva, Abrão

Authors: Vaz, Antônio
Andrade, Paulo
Carvalho, Edgar
Melo, Marcia

Date Submitted: 21-Jul-2013

 Print  Return to Dashboard

ScholarOne Manuscripts[™] v4.12 (patent #7,257,767 and #7,263,655). © ScholarOne, Inc., 2013. All Rights Reserved.
ScholarOne Manuscripts is a trademark of ScholarOne, Inc. ScholarOne is a registered trademark of ScholarOne, Inc.

 Follow ScholarOne on Twitter

[Terms and Conditions of Use](#) - [ScholarOne Privacy Policy](#) - [Get Help Now](#)

9.2. ANEXO B. Instruções para autores e normas de publicação de periódico cujo artigo encontra-se em apreciação para publicação (link)

1- Parasite Immunology

[http://onlinelibrary.wiley.com/journal/10.1111/\(ISSN\)1365-3024/homepage/ForAuthors.html](http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1365-3024/homepage/ForAuthors.html)