

ANÁLISE IN SILICO DO DOMÍNIO C-TYPE EM LECTINAS DE LEPTOSPIRA SPP.

**MONIZE N. PROVVISOR SANTOS¹; FREDERICO SCHMITT KREMER²;
MARCUS REDÜ ESLABÃO², LUCIANO DA SILVA PINTO³**

¹Universidade Federal de Pelotas, Centro de Desenvolvimento Tecnológico, Núcleo de Biotecnologia, Laboratório de Bioinformática – moprovisor@hotmail.com

²Universidade Federal de Pelotas, Centro de Desenvolvimento Tecnológico, Núcleo de Biotecnologia, Laboratório de Bioinformática – fred.s.kremer@gmail.com

²Universidade Federal de Pelotas, Centro de Desenvolvimento Tecnológico, Núcleo de Biotecnologia, Laboratório de Bioinformática – marcus.eslabao@yahoo.com.br

³Universidade Federal de Pelotas, Centro de Desenvolvimento Tecnológico, Núcleo de Biotecnologia, Laboratório de Biotecnologia Vegetal – ls_pinto@gmail.com

1. INTRODUÇÃO

Lectinas são glicoproteínas de origem não imunológica presentes em diversos tipos de organismos, que se caracterizam pela capacidade de se ligarem a vários tipos de carboidratos presentes em membranas celulares, podendo exercer diversos papéis biológicos, sendo um deles a capacidade de estimular células do sistema imune. (FREIRE, 2003). Com relação à estrutura e a sua capacidade de ligação a diferentes tipos de carboidratos, existem diferentes classificações aplicadas às lectinas, sendo que uma delas as descreve como lectinas do tipo-C ou CTLs (do inglês: C-type lectins). Estas podem ser encontradas no soro, matriz extracelular ou em membranas celulares (SHARON; LIS, 2004) e já foram preditas para algumas espécies de micro-organismos patogênicos. No entanto, ainda existem poucas informações conhecidas sobre esse tema.

O gênero *Leptospira* é composto por bactérias pertencentes à família *Leptospiraceae* da ordem *Spirochaetales*, que se caracterizam por possuir diversos sorovares patogênicos tanto para humanos quanto para animais, sendo a espécie *L. interrogans* considerada extremamente infectante, principalmente por possuir um elevado grau de variação antigênica. Estas bactérias são responsáveis por causar a leptospirose, uma infecção aguda antroponozoonótica contemporânea, de ampla distribuição geográfica, responsável por gerar perdas econômicas que estão relacionadas principalmente com a criação de bovinos, tanto para a produção de leite quanto de corte. (ARAÚJO *et al.*, 2005). Segundo Cullen (CULLEN *et al.*, 2004), as proteínas de membrana externa de espiroquetas interagem com o ambiente e com o sistema do hospedeiro, contribuindo para a patogenia da doença. Durante a infecção, as *Leptospiras* devem ser capazes de driblar o sistema imune do hospedeiro para efetuar uma apropriada colonização, sendo que moléculas associadas à membrana devem mediar essa interação entre o hospedeiro e o microrganismo.

Tendo em vista estes fatos, pesquisas relacionadas com a detecção, purificação e caracterização de lectinas podem ser consideradas fundamentais para a busca de novos mecanismos de combate a infecções por microrganismos patológicos. Considerando a atual falta de informações disponíveis sobre lectinas no gênero *Leptospira* e a potencial aplicação biotecnológica destas proteínas, o presente trabalho teve como objetivo

analisar a proteína hipotética C-type-lectin de *Leptospira spp.* através de diferentes abordagens bioinformáticas.

2. METODOLOGIA

A sequência da lectina LIC_12359 (Proteína hipotética), extraída da sequência de *Leptospira interrogans* serogroup *Icterohaemorrhagiae* serovar *copenhageni* (strain Fiocruz L1-130), foi retirada do Banco de dados Uniprot (Q72PV8). A análise de seus domínios foi realizada através das seguintes ferramentas: ProDom (SERVANT *et al.*, 2002), SMART (SCHULTZ *et al.*, 1998) e InterProScan (JONES *et al.*, 2014). A presença de peptídeo sinal foi analisada pelo software SignalP (PETERSEN *et al.*, 2011), seguida da análise de anotação funcional, realizada através da ferramenta Blast2Go (CONESA *et al.*, 2005). Também foi realizada a predição da estrutura tridimensional da proteína por meio de três ferramentas: I-TASSER (ZHANG, 2008), HHPred (SÖDING *et al.*, 2005) e SWISS-MODEL (BIASINI *et al.*, 2014), sendo que os modelos tridimensionais gerados tiveram suas estruturas alinhadas pelo software PyMOL (<http://www.pymol.org/>) e analisadas, posteriormente, através do gráfico de Ramachandran, disponível no servidor RAMPAGE (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>).

A busca por sequências homólogas dentro da espécie do gênero *Leptospira* foi realizada no Uniprot com a ferramenta BLAST (ALTSCHUL *et al.*, 1990), sendo as sequências resultantes alinhadas com a ferramenta MUSCLE (EDGAR, 2004). A ferramenta PSORT (<http://www.psort.org/>) foi utilizada para verificar se a proteína em questão é de membrana ou citosólica. A análise da região do domínio C-type foi realizada com o auxílio da ferramenta UGENE (<http://ugene.unipro.ru/>), sendo para esta região gerada uma *sequence logo*, com o auxílio da ferramenta WebLogo (CROOKS *et al.*, 2004) para visualização da conservação.

3. RESULTADOS E DISCUSSÃO

Através da análise dos domínios da sequência pelo banco de dados Interpro foram encontrados três domínios, com destaque para o Domínio C-type lectin, que é evidenciado por ser capaz de reconhecer diversas estruturas de carboidratos presentes em micro-organismos. Nessa sequência, este domínio localiza-se entre os aminoácidos 311 e 352. Na análise pelo banco de dados SMART, o domínio C-type lectin também foi identificado e, neste caso, a sequência referente a este domínio está compreendida entre os aminoácidos da posição 225 a 351. E, durante a análise pelo ProDom, também foi encontrado o domínio referente à lectina do tipo C, entre os aminoácidos 303 e 423.

Durante a análise do gráfico gerado pelo software SignalP, não foi detectada a presença de peptídeo sinal. O software Blast2go foi utilizado para realizar a anotação funcional da sequência e demonstrou que o domínio C-type lectin se encontra, nesta sequência, associado ao metabolismo de lipídeos, apesar de lectinas estarem mais comumente associadas ao metabolismo de carboidratos.

Os softwares utilizados para predição das estruturas tridimensionais dessa proteína foram: I-TASSER, HHPred e SWISS-MODEL, sendo suas

respectivas estruturas visualizadas pelo software PyMol (Figura 1), com o objetivo de identificar regiões mais conservadas, sendo estas apresentadas a seguir:

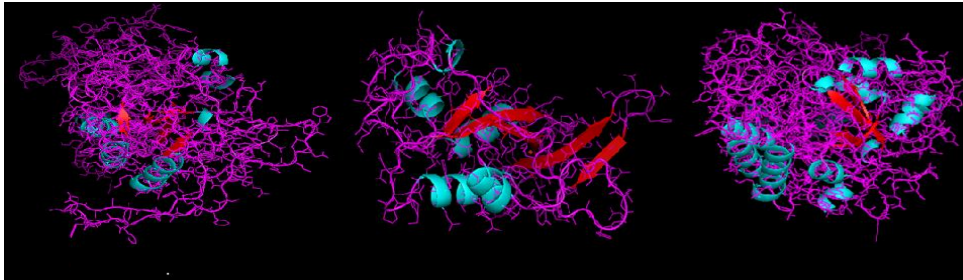


Figura 1: Representação dos Modelos gerados pelos preditores I-TASSER, HHpred e SWISS-MODEL, respectivamente.

Após a predição de estruturas, foi realizada a análise pelo gráfico de *Ramachandran*, com o objetivo de selecionar qual das estruturas possui mais aminoácidos na posição correta, sendo assim a mais confiável. Nos modelos gerados foi comprovado que o Modelo HHpred foi o que apresentou mais aminoácidos em posição favorável.

Além disso, foi realizado um BLAST das sequências de *Leptospira spp* e de *Leptospira interrogans* obtidas do Uniprot e um subsequente alinhamento através da ferramenta MUSCLE, com o objetivo de buscar similaridades e conservações entre as sequências de *Leptospira* com a sequência proteica de interesse. Para visualizar o alinhamento da região do domínio C-type, utilizou-se o software UGene, sendo observada uma grande conservação quando o gene foi alinhado com as sequências da espécie *interrogans* e uma menor conservação, quando alinhado com a sequência de todas as espécies de *Leptospiras*. Para avaliar com mais precisão as regiões de conservação foi gerada uma sequence logo, através da ferramenta WebLogo, obtida da comparação entre a sequência de interesse e a sequência de *Leptospira interrogans* e outra sequence logo obtida da comparação entre a sequência de interesse e a sequência de todas as espécies de *Leptospira*. Com essas sequências geradas houve a confirmação dos resultados gerados pelo alinhamento.

4. CONCLUSÕES

Através das análises in silico não foi possível determinar com exatidão a localização e a função da lectina. Por esse motivo é necessário que sejam realizadas análises mais detalhadas, em laboratório, para a obtenção de respostas que irão auxiliar na compreensão a nível molecular da lectina e até mesmo sua possível associação com a patogênese ou outros processos biológicos.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- ALTSCHUL, S. F.; GISH, W.; MILLER, W.; MYERS, E. W.; LIPMAN, D. J. Basic local alignment search tool. **Journal of Molecular Biology**, v.215, n.3, p. 403-410, 1990.
- ARAÚJO, V. E. M.; MOREIRA, E. C.; NAVEDA, L. A. B. Frequência de aglutininas anti *Leptospira interrogans* em soros sanguíneos de bovinos, em

Minas Gerais, de 1980 a 2002, **Arq. Bras. Med. Vet. Zootecnia**, v.57, n.4, p.430-435, 2005.

BIASINI, M.; MARIANI, V.; HAAS, J.; SCHEUBER, S.; SCHENK, A. D.; SCHWEDE, T.; PHILIPPSSEN, A. OpenStructure: a flexible software framework for computational structural biology. **Bioinformatics**, v.26, n.20, p. 2626-2628, 2010.

CONESA, A.; GÖTZ, S.; GOMEZ, J. M. G.; TEROL, J.; TALÓN, M.; ROBLES, M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. **Bioinformatics**, v.21, n.18, p. 3674–3676, 2005.

CULLEN, P. A.; XU, X.; MATSUNAGA, J.; SANCHEZ, Y.; KO, A.I.; HAAKE, D.A.; ADLER, B. Surfaceome of *Leptospira* spp. **Infection and Immunity**, New York, v.73, n.8, p.4853-4863, 2004.

CROOKS, G. E.; HON, G.; CHANDONIA, J. M.; BRENNER, S. E. WebLogo: A Sequence Logo Generator. **Genome Res**, California, v.14, n.6, p.1188-1190, 2004.

EDGAR, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, v.32, n.5, p. 1792-1797, 2004.

FREIRE, M.G. **Isolamento, caracterização físico-química e estudo das atividades inseticida, microbicida e inflamatória da lectina de sementes de *Talísia Esculenta***. 2003. 189p. Tese (Doutorado em Bioquímica) - Instituto de Biologia, Universidade Estadual de Campinas.

JONES, P.; BINNS, D.; CHANG, H.; FRASER, M.; LI, W.; MCANULLA, C.; MCWILLIAM, H.; MASLEN, J.; MITCHELL, A.; NUKA, G.; PESSEAT, S.; QUINN, A. F.; VEGAS, A. S.; SCHEREMETJEW, M.; YONG, S.; LOPEZ, R.; HUNTER, S. InterProScan 5: genome-scale protein function classification. **Bioinformatics**, Hinxton, vol. 30, n.6, p. 1236-1240, 2014.

PETERSEN, T. N.; BRUNAL, S.; VON HEIKNE, G.; NIELSEN, H. Signalp 4.0: discrimination signal peptides for transmembrane regions. **Nature Methods**, v. 8, p. 785-786, 2011.

SCHULTZ, J.; MILPETZ, F.; BORK, P.; PONTING, C.P. SMART, a simple modular architecture research tool: Identification of signaling domains. **PNAS**, California, vol. 95, n.11, p. 5857–5864, 1998.

SHARON, N; LIS, H. History of lectins: from hemagglutinins to biological recognition molecules. **Glycobiology**, Israel, vol. 14, n. 1, p. 53R–62R, 2004.

SÖDING, J.; BIEGERT, A.; LUPAS, A.N. The HHpred interactive server for protein homology detection and structure prediction. **Nucleic Acids Research**, Tübingen, vol 33, n.2, p. w244-w248, 2005.

ZHANG, Y. I-TASSER server for protein 3D structure prediction. **BMC Bioinformatics**, v.9, n.40, p. 1471-2105, 2008.

