

Potencial biotecnológico das lectinas da família Musaceae: um estudo de bioinformática estrutural

ANA CLAUDIA OLIVEIRA DE FREITAS¹;
ALESSANDRA NEIS²; LUCIANO DA SILVA PINTO³

¹Universidade Federal de Pelotas – anaclaudia.olvf@outlook.com;

²Universidade Federal de Pelotas – alennneis@gmail.com - alennneis@gmail.com;

³Universidade Federal de Pelotas – Luciano da Silva Pinto - ls_pinto@hotmail.com;

1. INTRODUÇÃO

Lectinas são um grupo heterogêneo de proteínas que se ligam reversivelmente a diversos tipos de carboidratos (DAMME *et al.*, 1998). A importância do estudo dessas proteínas reside em seu amplo potencial biotecnológico. Lectinas vegetais, como a extraída da banana, apresentam diversas atividades biológicas já elucidadas, como: capacidade imunomoduladora, antiproliferativa e antiviral (COVÉS-DATSON *et al.*, 2019).

As bananas pertencem a Família Musaceae, que agrupa diferente gêneros e dentre eles se encontram o *Musa* e *Ensete*. Atualmente, existe uma escassez de informações sobre lectinas dos diferentes gêneros dessa família, e os estudos disponíveis de concentram no gênero *Musa*, com grande destaque para espécie *Musa acuminata*. A lectina com maior destaque é a Banlec (*M. acuminata*), que também possui uma versão recombinante nomeada H84T, com elevada atividade antiviral e sua mitogenicidade anulada (COVÉS-DATSON *et al.*, 2019, 2020).

Para explorar as características das lectinas negligenciadas da família Musaceae, a bioinformática é uma ferramenta que pode auxiliar na tomada de decisão, pois promove a redução de custos e tempo de pesquisa e, ainda, disponibiliza uma grande quantidade de dados permitindo avaliar características estruturais e funcionais das proteínas. Algumas das ferramentas mais utilizadas são a modelagem estrutural para obter a estrutura tridimensional da lectina e o *docking molecular* para analisar as possíveis interações entre a proteína e carboidratos (SINGH; DEVI; NG, 2014a).

Isto posto, o objetivo principal do presente trabalho é identificar lectinas da família Musaceae que são negligenciadas e, a partir de ferramentas de bioinformática, obter a estrutura tridimensional e analisar interação fina com glicanos preferenciais, com o intuito de subsidiar trabalhos futuros que explorem a expressão heteróloga e potencial biotecnológico dessas proteínas.

2. METODOLOGIA

2.1 Análise e obtenção das sequências:

A ferramenta de BLASTp (*Basic Local Alignment SearchTool*) foi utilizada para encontrar as lectinas mais semelhantes da família Musaceae utilizando como input a lectina de *Musa acuminata* (PDB ID: 1X1V). Para isso os seguintes parâmetros foram utilizados: banco de dados de proteínas não redundantes (nr) e adotando como limitante de busca a família Musaceae (ALTSCHUL *et al.*, 1990).

Foram selecionadas 15 sequências mais similares e que apresentaram a maior cobertura. Estas sequências foram alinhadas utilizando a ferramenta CLUSTAL a fim de avaliar as modificações específicas nos sítios de ligação ao carboidrato (HIGGINS, 1994). Para localizar os sítios de ligação a carboidrato nas sequências foi utilizado banco de dados UNIPROT (APWEILER, 2004). Por fim, foram analisadas pelo software SignalP para detectar a presença de peptídeo sinal.

2.2 Modelagem estrutural e refinamento das sequências:

As sequências foram submetidas a uma modelagem estrutural no software I-TASSER (ZHANG, 2008), que utiliza uma metodologia de *meta-threading* que consiste na procura em bancos de dados por proteínas com modelos de dobras semelhantes ao *input*. Logo após, as estruturas foram analisadas visualmente pelo software PyMol (PYMOL MOLECULAR GRAPHICS SYSTEM, SCHOREDINGER, versão 4.20) e foram sobrepostas a *Musa acuminata* (PDB ID: 3MIT) que está ligada a manose e possui identidade de 100% com a *Musa acuminata* (PDB ID: 1X1V).

Com a finalidade de checagem de qualidade e refinamento das modelagens foram utilizados os seguintes softwares: Galaxyrefine e PDBsum. Três lectinas foram escolhidas (THU46912.1; RWV92385.1; XP_018673706.1), a escolha levou em conta a falta de estudo com as espécies *M. balbisiana* e *E. ventricosum* em conjunto com a diferença nos sítios ativos e a similaridade dos mesmos na espécie de *M. acuminata* subsp. *malaccensis* pois, essa lectina pode apresentar atividades semelhantes as já descritas para a espécie (GOLDSTEIN *et al.*, 2001;).

2.3 Docking molecular

Para analisar a interação entre proteína-carboidrato, foram escolhidos os softwares Autodock Vina e Autodock Tools (MORRIS *et al.*, 2006; TROTT; OLSON, 2009). O *docking* foi realizado apenas com as três lectinas selecionadas com suas estruturas obtidas pelo I-TASSER. O ligante utilizado foi a manose e a sua estrutura 3D foi obtida RCSB PDB (<http://ligand-expo.rcsb.org/>). Também foi utilizado o software Plip (*Protein Ligand Interaction Profiler*) que forneceu as informações sobre as ligações os aminoácidos envolvidos na ligação com o carboidrato (Salentin *et al.*, 2015).

3. RESULTADOS E DISCUSSÃO

Na ferramenta de BLASTp, foram encontradas 22 sequências com 100% de cobertura e similares ao input *M. acuminata* (PDB ID: 1X1V) que eram pertencentes as espécies *M. acuminata* (7 isoformas), *M. balbisiana* (4 isoformas), e *Ensete ventricosum* (4 isoformas). Todas apresentaram alta similaridade (100% - 89%) em relação a de referência. Para as análises seguintes apenas 3 lectinas foram selecionadas pois as outras eram isoformas da mesma lectina. No alinhamento global, realizado pelo CLUSTAL foi possível localizar os sítios de interação fornecidos pelo Uniprot e pôde-se observar a presença de alterações pontuais nos aminoácidos, que podem ou não alterar o perfil de interação a carboidrato, porém para essa análise são necessárias mais estas de estudo *in vitro* e *in vivo*.

Tabela 1 Comparação entre os sítios de interação a carboidratos fornecidos pelo Uniprot para a *M. acuminata* (PDB ID 1X1V) e as três lectinas selecionadas para as análises no presente trabalho. As modificações dos resíduos estão assinaladas em negrito e sublinhadas.

Organismo de origem	ID	Locais de ligação ao carboidrato			
		30-38	60-63	105-106	130-133
<i>Musa acuminata</i>	1X1V	SGDVVD	GTPH	GN	DFID
<i>Musa balbisiana</i>	THU46912.1	SGDVVD	GTPH	GN	<u>QFLD</u>
<i>Musa acuminata</i> subsp. malaccensis	<u>XP_018673706.1</u>	SGDVVD	GTPH	GN	DF <u>LD</u>
<i>Ensete ventricosum</i>	<u>RWV92385.1</u>	SG G VVD	G <u>N</u> PH	GN	<u>QFLD</u>

Ao realizar a modelagem no I-TASSER é obtido os valores de *c-score* que são uma pontuação de confiança, que varia de -5 a 2, sendo que os valores mais altos apresentam modelos de melhor confiança. Logo após, no Pymol, foi realizada a sobreposição das estruturas geradas pelo I-TASSER utilizando a molécula de *M. acuminata* resolvida (PDB ID: 3MIT). O intuito foi analisar a pontuação RMSD, para observar as diferenças estruturais entre os modelos gerados e a referência. Os valores de *c-score* dos modelos selecionados foram próximos a 1, o que pode ser considerado um resultado com qualidade e o valor de RMSD foi inferior a 1, o que demonstra que os modelos eram semelhantes a referência, o que é esperado, já que possuem elevada similaridade (Tabela 2).

Tabela 2 Valores de *c-score* e RMSD das lectinas selecionadas no estudo

ID	Nome	<i>c-score</i>	RMSD
<u>THU46912.1</u>	<i>Musa balbisiana</i>	-0,97	0,914
<u>XP_018673706.1</u>	<i>Musa acuminata</i> subsp. malaccensis	0,86	0,581
<u>RWV92385.1</u>	<i>Ensete ventricosum</i>	0,86	0,527

Após o refinamento e checagem de qualidade, foi realizado o docking molecular e todas as lectinas apresentaram as energias de ligação a carboidrato muito semelhantes ao modelo *M. acuminata* (PDB ID 3MIT). Além disso, na disposição espacial realizada no Pymol, as poses geradas pelo *docking* utilizando a manose como ligante, as três lectinas selecionadas (*M. balbisiana* THU46912.1; *E. ventricosum* RWV92385.1; *M. acuminata* subsp. *malaccensis* XP_018673706.1) apresentaram proximidade a disposição espacial das moléculas do ligante com a referência (Figura 1).

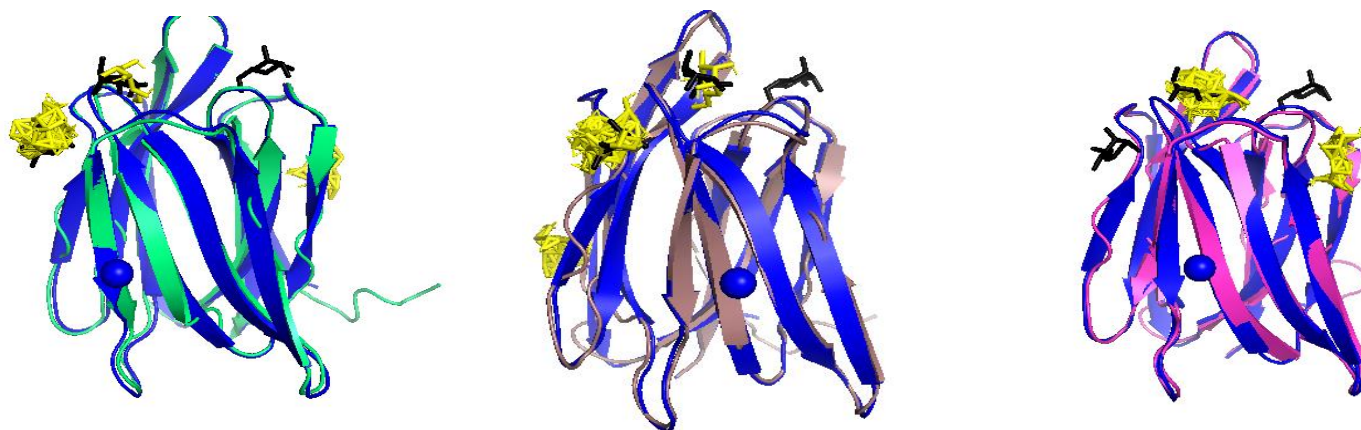


Figura 1 Sobreposição do *docking* molecular no software PLIP, utilizando a manose como ligante e a *M. acuminata* (PDB ID 3MIT) como referência, para a análise da disposição espacial dos ligantes. *E. ventricosum* RWV92385.1 (verde claro) (A) *M. balbisiana* THU46912.1 (rosa escuro) (B) *M. acuminata* subsp. *malaccensis* XP_018673. As nove poses da manose fornecida pelo *docking* molecular estão em amarelo e as poses referência estão em preto.

4. CONCLUSÕES

O presente trabalho, através das ferramentas de bioinformática conseguiu demonstrar uma elevada similaridade de diferentes lectinas da família Musaceae e, portanto, possível aplicação biotecnológica, já que estas possuem semelhança de interação a carboidrato. Esses resultados impulsionam futuras pesquisas biotecnológicas com essas lectinas negligenciadas.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- COVÉS-DATSON, Evelyn M. *et al.* Inhibition of Ebola Virus by a Molecularly Engineered Banana Lectin. **PLoS Neglected Tropical Diseases**, [S. l.], v. 13, n. 7, p. 1–20, 2019. Disponível em: <https://doi.org/10.1371/journal.pntd.0007595>
- DAMME, Els J. M. Van *et al.* Plant Lectins: A Composite of Several Distinct Families of Structurally and Evolutionary Related Proteins with Diverse Biological Roles. **Critical Reviews in Plant Sciences**, [S. l.], v. 17, n. 6, p. 575–692, 1998. Disponível em: <https://doi.org/10.1080/07352689891304276>. Acesso em: 30 abr. 2020.
- SINGH, Senjam Sunil; DEVI, Sanjenbam Kunjeshwori; NG, Tzi Bun. Banana lectin: A brief review. **Molecules**, [S. l.], v. 19, n. 11, p. 18817–18827, 2014. Disponível em: <https://doi.org/10.3390/molecules191118817>
- ALTSCHUL, Stephen F. *et al.* Basic local alignment search tool. **Journal of Molecular Biology**, [S. l.], v. 215, n. 3, p. 403–410, 1990. Disponível em: [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2). Acesso em: 15 ago. 2020.
- ZHANG, Yang. I-TASSER server for protein 3D structure prediction. **BMC Bioinformatics**, [S. l.], v. 9, n. 1, p. 1–8, 2008. Disponível em: <https://doi.org/10.1186/1471-2105-9-40>. Acesso em: 15 mar. 2021.
- GOLDSTEIN, Irwin J. *et al.* Carbohydrate binding properties of banana (*Musa acuminata*) lectin: II. Binding of laminaribiose oligosaccharides and β -glucans containing β 1,6-glucosyl end groups. **European Journal of Biochemistry**, [S. l.], v. 268, n. 9, p. 2616–2619, 2001. Disponível em: <https://doi.org/10.1046/j.1432-1327.2001.02149.x>. Acesso em: 20 jul. 2020.