POTENCIAL BIOTECNOLÓGICO DAS PROTEÍNAS ANTICONGELANTES DE PRASIOLA CRISPA

Darlene Lopes Rangel [1]

Lucas Ferreira Maciel [2]

Evelise Leis Carvalho [3]

Thalita Fonseca de Araújo [4]

Maria Eduarda Tabarez de Abreu [5]

Paulo Marcos Pinto [6]

Resumo:

Prasiola crispa é uma macroalga verde taloide, pertencente a classe Treboxiophyceae, de distribuição biogeográfica cosmopolita, estando presente do Ártico ao continente Antártico. Na Antártica, P. crispa atua como importante produtor primário, resistindo a condições ambientais extremas como baixas temperaturas, estresse osmótico, radiação ultravioleta, entre outros. Devido a sua capacidade de colonizar um ambiente tão inóspito, esta alga deve possuir mecanismos adaptativos, cujos genes envolvidos são de grande interesse na área da Biotecnologia. Entre as principais biomoléculas de interesse, as proteínas de ligação ao gelo (IBPs) destacam-se. IBPs são polipeptídeos que permitem a sobrevivência de células à baixas temperaturas. Este grupo de proteínas possui grande aplicabilidade na agricultura, biomedicina e indústria alimentícia. Assim, este trabalho teve como objetivo identificar possíveis IBPS no transcriptoma já montado e anotado de P. crispa. A amostra foi coletada na Ilha do Rei George, Antártica. O RNA total foi extraído, a biblioteca montada e sequenciada por Illumina HiSeq 2000, com estratégia paired-end reads. Reads de baixa qualidade e adaptadores foram removidos e as reads processadas submetidas a montagem pelo montador Trinity. Os transcritos montados foram anotados pelo servidor web-based MG-RAST em pipeline automatizado e indicaram a contaminação com a microbiota. O transcriptoma de P. crispa foi então descontaminado e posteriormente anotado funcionalmente através da ferramenta Blast2GO. IBPs de P. crispa e microbiota foram buscadas por algoritmo BLASTX, alinhando os transcritos contra as sequências de IBPs em um banco de dados local. Transcritos identificados através da busca tiveram a sequência aminoacídica predita utilizando o Trandescoder e modeladas computacionalmente através do servidor Phyre2. Da busca com o BLASTX, identificou-se 127 transcritos como possíveis IBPs, com homologia a proteínas de plantas, bactérias e fungos. A modelagem computacional da estrutura tridimensional indicou que 17 destas possíveis IBPs possuem estrutura semelhante a outras proteínas do grupo, sendo que 8 delas possuem maior potencial. Em conclusão, a montagem e anotação funcional foram realizadas de maneira satisfatória e permitiu a identificação de potenciais IBPs. Como perspectiva, simulações de Dinâmica Molecular serão realizadas para verificar o comportamento destas proteínas em diferentes temperaturas para a escolha de quais serão expressas heterologamente.

Palavras-chave: Bioinformática, algas verdes, IBPs

Modalidade de Participação: Pesquisador

POTENCIAL BIOTECNOLÓGICO DAS PROTEÍNAS ANTICONGELANTES DE PRASIOLA CRISPA

[1] Aluno de pós-graduação. darlenelrangel@gmail.com. Autor principal

[2] Aluno de graduação. lucasmacielbiotec@gmail.com. Co-autor

[3] Aluno de pós-graduação. eveliseleis@gmail.com. Co-autor

[4] Aluno de graduação. thalita.fonseca23@gmail.com. Co-autor

[5] Aluno de graduação. mariaeduardatabarez@gmail.com. Co-autor

[6] Docente. paulomarcospinto@gmail.com. Co-autor

POTENCIAL BIOTECNOLÓGICO DAS PROTEÍNAS ANTICONGELANTES DE *Prasiola crispa*

1. INTRODUÇÃO

Proteínas de ligação ao gelo (IBPs, do inglês *Ice Binding Proteins)* são polipeptídeos expressos em uma ampla gama de organismos que permitem a sobrevivência das células à baixas temperaturas (USTUN, TURHAN, 2015). A primeira IBP foi identificada em 1969, quando Arthur DeVries demonstrou que proteínas presentes no sangue de peixes teleósteos, habitantes da região Antártica, diminuem o seu ponto de congelamento (DEVRIES, WOHLSCHLAG, 1969). Estas proteínas foram inicialmente denominadas proteínas anticongelantes (AFPs, do inglês *antifreeze proteins*), um termo utilizado de maneira genérica por muitos anos, mas que atualmente compreende apenas um subconjunto das IBPs. Além dos peixes, este grupo de proteínas já foram identificadas em grande número de bactérias (CHATTOPADHYAY, 2007), fungos (XIAO et al., 2010), insetos (DUMAN, 2015) e plantas (GUPTA, DESWAL, 2014). Ademais, IBPs foram encontradas em algumas algas da Antártica (BAYER-GIRALDI et al., 2010). A distribuição de IBPs em tantas espécies distintas ocorreu provavelmente através de eventos de convergência evolutiva e transferência horizontal de genes (RANDY CHI FAI, TZI, JACK, 2017).

A produção em larga escala de IBPs ainda precisa de melhorias para se tornar viável e amplamente difundida, mas as possibilidades são variadas. Entre as aplicabilidades podem ser citadas a utilização para a preservação e melhoria da qualidade de alimentos congelados (HASSAS-ROUDSARI, GOFF, 2012), produção de culturas tolerantes ao frio (DUMAN, WISNIEWSKI, 2014), crioproteção de tecidos e órgãos (BROCKBANK, 2011) e até mesmo a aplicação na indústria petroleira (PERFELDT, 2014).

A predição da estrutura tridimensional de proteínas através de abordagens computacionais é uma importante área da Bioinformática, que está em franco desenvolvimento e possui grandes impactos sobre a Biotecnologia

A modelagem computacional de proteínas pode seguir por dois métodos: os independentes de estrutura molde e os baseados em estrutura molde. Os métodos independentes de proteína molde são divididos em *ab initio* e *de novo*, enquanto o baseado em estrutura molde dividem-se em *threading* e modelagem comparativa. Estas abordagens são assim classificadas de acordo com as informações que utilizam a partir dos bancos de dados, sendo o *Protein Data Bank* (PDB) o mais utilizado (VERLI, 2014)

Neste trabalho objetivamos identificar IBPs presentes no transcriptoma de *Prasiola crispa*, sequenciado por nosso grupo de pesquisa. *P. crispa*, que habita o continente Antártico, é um dos organismos mais presentes e importante produtor primário (CONVEY, 2006). Esta alga está presente nas regiões supralitorais, formando grandes tapetes verdes, sendo encontrada principalmente em locais próximos a colônias de aves, substrato rico em guano, onde há altas concentrações de nitrogênio e ácido úrico. Por habitar um ambiente inóspito, é de grande interesse o estudo e a identificação de IBPs, uma vez que essas proteínas possuem grande potencial biológico e biotecnológico.

2. METODOLOGIA

Prasiola crispa foi coletada na Ilha do Rei George, Antártica. A coleta foi realizada no verão Antártico, em janeiro de 2014. O RNA total foi extraído e o transcriptoma foi sequenciado pelo serviço da Macrogen, utilizando a plataforma Solexa-Illumina HiSeq 2000. *Reads* brutas foram filtradas para a remoção dos adaptadores e das *reads* de baixa qualidade com *Fastx-toolkit*. Após isso, a ferramenta Trinity versão r2014-07-17 foi utilizada como montador DBG para a geração dos contigs, utilizando k-mers de 25 pb.

Os *contigs* montados e recuperados foram alinhados contra o banco de proteínas não-redundantes do NCBI, sendo o *cut-off* escolhido 1e-10. Identificou-se os transcritos de acordo com o melhor hit contra sequências de função conhecida. Blast2GO foi utilizado para mapear e anotar as sequências, associando termos do GO e predizendo suas funções.

Para a criação de um banco de dados local e busca por IBPs os termos *Ice-binding protein*, *Anti-freezing protein*, *Thermal hysteresis* e *Ice-structuring protein* foram buscados no banco de dados de proteínas do NCBI. Todas as sequências associadas a estes termos foram baixadas e um banco de dados local foi criado. Após a criação do banco, todos os transcritos, inclusive os contaminantes, foram submetidos a busca contra o banco de dados local utilizando o algoritmo BLASTX (*cut-off*: 1e-10), a partir do alinhamento de sequências proteicas contra as seis fases de leitura dos transcritos.

Para a modelagem computacional, todas as sequências que tiveram alinhamento por BLASTX contra as IBPs, foram submetidas a análise pela ferramenta *Transdecoder* para a identificação de fases de leitura aberta (ORFs) e a sequência proteica correspondente. As sequências de aminoácidos dos transcritos com ORFs identificadas foram submetidas a modelagem no web-server Phyre 2.

3. RESULTADOS e DISCUSSÃO

A busca identificou 127 sequências com ao menos um *BLAST hit*, com E-value variando entre 1e-10 e 1e-111. A origem das proteínas apresentou grande variação, contendo sequências oriundas de plantas, algas, bactérias e fungos, sendo o maior número de hits com proteínas bacterianas. Os 127 contigs foram submetidos para a análise de presença de ORFs e obtenção da sequência de aminoácidos. Apenas 11 contigs apresentaram ORF completa e 45 não continham ORFs. As sequências de aminoácidos de *contigs* com ORFs, completas ou parciais, foram submetidas a modelagem pelo web-server Phyre 2. Ao todo, 81 modelagens foram realizadas, sendo que em 27 destas a abordagem *ab initio* foi aplicada. É possível que as mesmas representem IBPs, contudo, devido à baixa quantidade de estruturas tridimensionais conhecidas e a não confiabilidade dos algoritmos *ab initio* é equivocada qualquer inferência que possa ser feita atualmente e por isso estas sequências não foram consideradas para análises posteriores.

Entre as modelagens comparativas, 38 tiveram como molde principal proteínas de diversas funções como imunoglobulinas, fibronectinas e moléculas de sinalização celular. Apesar da conhecida dualidade funcional de IBPs, ainda não há estudos que indiquem qualquer associação direta das mesmas com as proteínas que foram utilizadas como molde.

Por fim, 17 contigs foram modelados utilizando como molde IBPs ou proteínas com clara associação de dualidade demonstrada na literatura, como as quitinases em plantas. A confiabilidade foi descrita como a porcentagem de resíduos que foram modelados com precisão > 90%, sendo que alta confiabilidade significava mais de

90% dos resíduos, confiabilidade média entre 50 e 90% e baixa confiabilidade < 50%.

A IBP de *Flavobacterium frigoris*, uma bactéria Antártica, foi o principal molde utilizado em 7 modelagens de alta e média confiabilidade, com uma das sequências chegando a identidade de 58%. O enovelamento β-hélice, identificado através de estudos de cristalografia de raio-x, mostrou-se essencial para a interação os cristais de gelo. Já a IBP da bactéria marinha Antártica *Colwellia sp*. foi molde principal de outras 4 proteínas, de estrutura muito similar à de *F. frigoris*, sendo a grande diferença as regiões *motifs* (Figura 2).

Proteínas dos fungos *Typhula ishikariensis* e *Leucosporidium sp.* também foram utilizadas como moldes, porém as precisões da estrutura de ambas foram baixas (Figura 3).

E por fim, três isoformas foram identificadas com identidade entre 45 e 49% a estrutura tridimensional semelhante à quitinase de *Oryza sativa*. O alinhamento de sequências com outras quitinases confirmou a presença do domínio da família 19 das Glicohidrolases.

Sendo assim, 8 transcritos apresentaram maior potencial, sendo três deles oriundos diretamente de *P. crispa* e o restante a partir da microbiota. Estes três transcritos são os únicos pertencentes ao transcriptoma de *P. crispa* que foram identificados com o potencial de atividade IBP.

Figura 1 – Predição de estrutura tridimensional de proteínas que utilizaram como molde a IBP de *Flavobacterium frigoris*. As cores e as setas indicam a orientação, sendo o azul a região amino-terminal e o gradiente de cores até o vermelho a região carboxi-terminal.

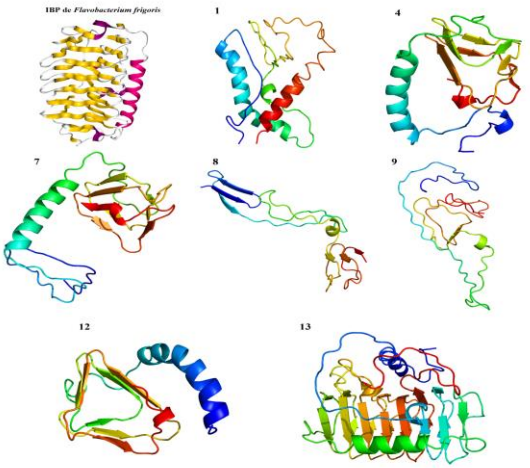


Figura 2 – Predição de estrutura tridimensional de proteínas que utilizaram como molde a IBP de Colwellia sp. As cores e as setas indicam a orientação, sendo o azul

a região amino-terminal e o gradiente de cores até o vermelho a região carboxi-terminal.

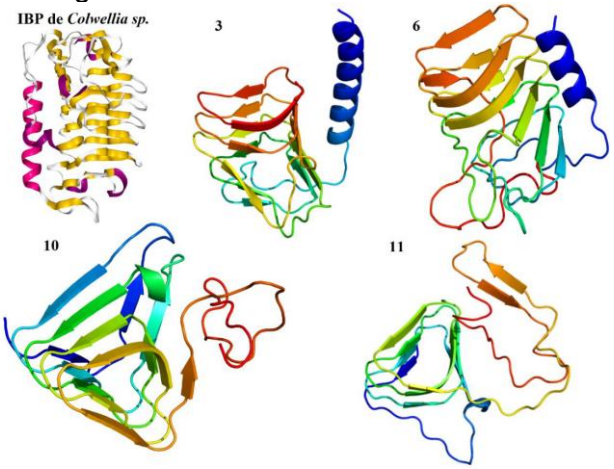
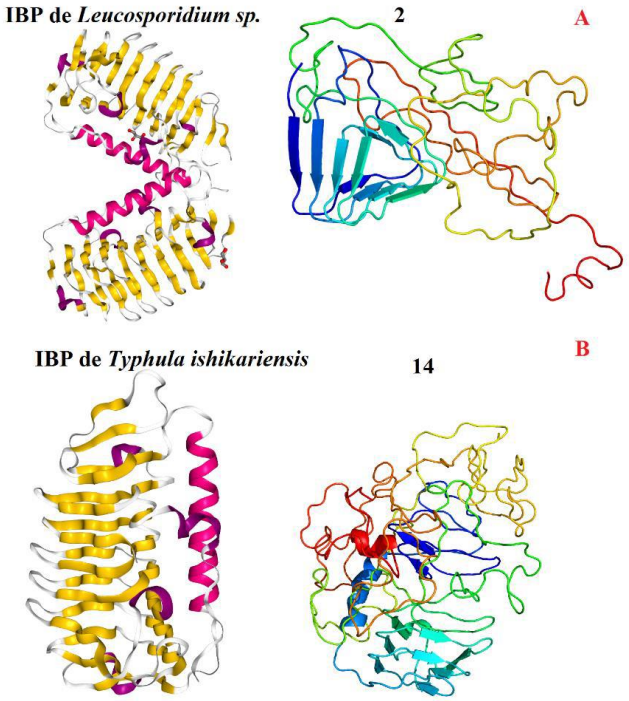


Figura 3 – Predição de estrutura tridimensional de proteínas que utilizaram como molde a IBP de (A) Leucosporidium sp. e (B) Typhula ishikariensis. As cores e as setas indicam a orientação, sendo o azul a região amino-terminal e o gradiente de cores até o vermelho a região carboxi-terminal.



4. CONSIDERAÇÕES FINAIS

Neste trabalho identificamos inicialmente 127 possíveis IBPs, onde a modelagem da estrutura tridimensional permitiu a identificação de 17 possíveis IBPs, sendo que 8 delas apresentaram o maior potencial. Com os resultados obtidos neste trabalho e considerando o potencial inovador das IBPs, o próximo passo seria a simulação *in silico* da Dinâmica Molecular para avaliar o comportamento destas estruturas proteicas em diferentes temperaturas e sua interação com os cristais de gelo. Também se espera amplificar, utilizando *primers* degenerados, e sequenciar os genes de maior potencial, para identificação da sequência completa.

5. REFERÊNCIAS

BAYER-GIRALDI, M., UHLIG, C., JOHN, U., MOCK, T., VALENTIN, K. Antifreeze proteins in polar sea ice diatoms: diversity and gene expression in the genus Fragilariopsis. Environmental Microbiology, v. 12 (4), p. 1041-1052, 2010.

BROCKBANK, K.G.M., CAMPBELL, L.H., GREENE, E.D., BROCKBANK, M.C.G., DUMAN, J.G. Lessons from nature for preservation of mammalian cells, tissues, and organs. In Vitro Cellular & Developmental Biology - Animal, v. 47 (3), p. 210-217, 2011.

CHATTOPADHYAY, M.K. Antifreeze proteins of bacteria. Resonance, v. 12 (12), p. 25-30, 2007.

CONVEY, P. Encyclopedia of the Antarctic. 1ed, Nova Iorque: Routledge, 2006.

DEVRIES, A.L., WOHLSCHLAG, D.E. Freezing Resistance in Some Antarctic Fishes. Science, v. 163 (3871), p. 1073-1075, 1969.

DUMAN, J.G., WISNIEWSKI, M.J. The use of antifreeze proteins for frost protection in sensitive crop plants. Environmental and Experimental Botany, v. 106, p. 60-69, 2014.

DUMAN, J.G. Animal ice-binding (antifreeze) proteins and glycolipids: an overview with emphasis on physiological function. The Journal of Experimental Biology, v. 218 (12), p. 1846-1855, 2015.

GUPTA, R., DESWAL, R. Antifreeze proteins enable plants to survive in freezing conditions. Journal of Biosciences, v. 39 (5), p. 931-944, 2014.

HASSAS-ROUDSARI, M., GOFF, H.D. Ice structuring proteins from plants: Mechanism of action and food application. Food Research International, v. 46 (1), p. 425-436, 2012.

PERFELDT, C.M., CHUA, P.C., DARABOINA, N., FRIIS, D., KRISTIANSEN, E., RAMLØV, H., WOODLEY, J.M., KELLAND, M.A., VON SOLMS, N. Inhibition of Gas Hydrate Nucleation and Growth: Efficacy of an Antifreeze Protein from the Longhorn Beetle Rhagium mordax. Energy & Fuels, v. 28 (6), p. 3666-3672, 2014.

RANDY CHI FAI, C., TZI BUN, N., JACK HO, W. Antifreeze Proteins from Diverse Organisms and their Applications: An Overview. Current Protein & Peptide Science, v. 18 (3), p. 262-283, 2017.

USTUN, N.S., TURHAN, S. Antifreeze Proteins: Characteristics, Function, Mechanism of Action, Sources and Application to Foods. Journal of Food Processing and Preservation, v. 39 (6), p. 3189-3197, 2015.

VERLI, H., Bioinformática: da Biologia à flexibilidade. 1. ed., São Paulo: SBBq, 2014.

XIAO, N., INABA, S., TOJO, M., DEGAWA, Y., FUJIU, S., KUDOH, S., HOSHINO, T. Antifreeze activities of various fungi and Stramenopila isolated from Antarctica. North American Fungi, v. 5, p. 215-220, 2010.