

UNIVERSIDADE FEDERAL DO PARÁ INSTITUTO DE CIÊNCIAS BIOLÓGICAS PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

LUAN DANIEL SILVA FERREIRA

DIVERSIDADE ESTRUTURAL E FUNCIONAL DE PROTEÍNAS CAROTENOIDES LARANJA DE CIANOBACTÉRIAS ISOLADAS DE AMBIENTES AMAZÔNICOS

BELÉM - PARÁ 2022

LUAN DANIEL SILVA FERREIRA

DIVERSIDADE ESTRUTURAL E FUNCIONAL DE PROTEÍNAS CAROTENOIDES LARANJA DE CIANOBACTÉRIAS ISOLADAS DE AMBIENTES AMAZÔNICOS

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Genética e Biologia Molecular da Universidade Federal do Pará como requisito parcial para a obtenção do título de Mestre.

Orientador: Dr. Evonnildo Costa Gonçalves.

BELÉM - PARÁ 2022

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD Sistema de Bibliotecas da Universidade Federal do Pará Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)

F383d FERREIRA, LUAN DANIEL SILVA. DIVERSIDADE ESTRUTURAL E FUNCIONAL DE PROTEÍNAS CAROTENOIDES LARANJA DE CIANOBACTÉRIAS ISOLADAS DE AMBIENTES AMAZÔNICOS / LUAN DANIEL SILVA FERREIRA. — 2022. 76 f. : il. color.

> Orientador(a): Prof. Dr. Evonnildo Costa Gonçalves Dissertação (Mestrado) - Universidade Federal do Pará, Instituto de Ciências Biológicas, Programa de Pós-Graduação em Genética e Biologia Molecular, Belém, 2022.

1. Diversidade. 2. Análise filogenética. 3. Modelagem. 4. Simulação. I. Título.

CDD 579.135



UNIVERSIDADE FEDERAL DO PARÁ INSTITUTO DE CIÊNCIAS BIOLÓGICAS PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR



ATESTADO

ATESTAMOS, para os devidos fins, que LUAN DANIEL SILVA FERREIRA foi aprovado na DEFESA DE DISSERTAÇÃO em Mestrado no Programa de Pós-Graduação em Genética e Biologia Molecular, da Universidade Federal do Pará, sob a orientação do Prof. Dr. Evonnildo Costa Gonçalves (UFPA), tendo defendido no dia 19/07/2022, perante Banca Examinadora, a Dissertação intitulada: "DIVERSIDADE ESTRUTURAL E FUNCIONAL DE PROTEÍNAS CAROTENÓIDES LARANJA DE CIANOBACTÉRIAS ISOLADAS DE AMBIENTES AMAZÔNICOS". A Banca Examinadora, constituída pelos professores doutores: Dr. Luís Adriano Santos do Nascimento (UFPA), Dr. Diego Assis das Graças (UFPA) e Dr. Rafael Azevedo Baraúna (UFPA), aprovou a dissertação apresentada pelo mestrando, atribuindo-lhe média geral: 9,42 (EXCELENTE), a ser homologada na reunião do Colegiado.

Belém, 19/07/2022.

Esta declaração não exclui o aluno de efetuar as mudanças sugeridas pela banca nem vale como outorga de grau de MESTRADO, de acordo com o definido na Resolução 072/2004-CONSEPE.

> Profa. Dra. Andrea Kely Campos Ribeiro dos Santos Coordenadora do Programa de Pos-Graduação em Genética e Biologia Molecular

UFPA - Cidade Universitária Prof. José da Silveira Neto - Instituto de Ciências Biológicas PPPGBM – Programa de Pós-Graduação em Genética e Biologia Molecular Av. Augusto Corrêa, 01 – Guamá – 66.075-900 – Fone/Fax 3201-8411 – (<u>www.ppgbm.propesp.ufpa.br</u>) e-mail: poggbm@ufpa.br

INSTITUIÇÕES E FONTES FINANCIADORAS

Laboratório de Tecnologia Biomolecular da Universidade Federal do Pará Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES

Dedico este trabalho primeiramente à minha família e a todos aqueles que contribuíram para sua realização.

AGRADECIMENTOS

A minha família, pela ajuda e apoio, em todos os sentidos, dado a mim sempre que precisei ao longo dessa jornada.

Agradeço de forma especial à minha mãe Deise Sueli Silva Ferreira e ao meu pai Perkens Roberto de Araújo Ferreira, pela força e esperança que foram depositadas na minha pessoa, incentivando-me a cada dificuldade.

A minha querida vó, Maria de Jesus Monteiro Lopes, por estar sempre comigo e me dando forças para continuar dia após dia, cuidando de mim da melhor forma possível.

A minha noiva Dayana Caroliny da Silva Alves por todas as formas de apoio e incentivo que foi me dado ao longo desses dois anos, não deixando que eu viesse a desistir do meu sonho.

Ao meu orientador, professor Dr. Evonnildo Costa Gonçalves pela paciência e compreensão, pela assistência na elaboração desta dissertação, e por sua confiança colocada em minha pessoa, sendo uma pessoa incrível e especial para mim.

Ao professor Dr. Andrei Siqueira, pela sua ajuda, auxiliando-me sempre que precisei, tirando dúvidas e me dando sugestões essências para a elaboração desse trabalho, se tonando uma pessoa muito admirada por mim.

Aos colegas do Laboratório de Tecnologia Biomolecular (LTB) Andrei Gama, Lucas Barros e Alenna Lima, pela companhia, troca de conhecimento, pelas resenhas e risadas, momentos que levarei comigo para o resto da vida.

Aos docentes dessa instituição, não deixando de lado à direção e administração que fizeram um excelente papel e atenderam todas as minhas necessidades.

A CAPES pela concessão da bolsa, onde sem ela eu não poderia estar aqui. Aos professores e colegas do PPGBM.

"O homem não teria alcançado o possível se, repetidas vezes, não tivesse tentado o impossível."

(Max Weber)

RESUMO

A Proteína Carotenoide Laranja tem a função de permitir que as células evitem possíveis danos causados pela luz, por meio do mecanismo de fotoproteção. O objetivo do trabalho foi analisar a diversidade e caracterizar as propriedades estruturais e funcionais da Proteína Carotenoide Laranja (OCP) isoladas de Cianobactérias Amazônicas, a fim de validá-las para uso em biotecnologia. As sequências de OCPs foram obtidas através de anotação dos dados genômicos feito após extração de DNA, sequenciamento, montagem e separação dos genomas de cianobactérias de dois ambientes amazônicos distintos. Para a análise filogenética foi feita uma buscar no BLASTp por outra OCPs e montada a filogenia utilizando o programa MEGA X. A análise estrutural foi feita através da modelagem comparativa das OCPs no programa MODELLER, utilizando como molde a proteína de Tolypothrix sp. PCC 7601 e os modelos gerados foram validados conforme parâmetros de Ramachandran, QMEAN, RMSD e Verify3D. Todos os modelos de homologia gerados foram submetidos a simulações de Dinâmica Molecular (DM), com refinamento de 100 ns. Observou-se, que a filogenia corroborou as relações filogenéticas das OCPs estudadas, e evidenciou a funcionalidade destas através da formação um agrupamento único com OCPs do tipo OCP1. Quando avaliados os modelos construídos a partir do molde selecionado, todos foram validados com sucesso, com ótimos valores de RMSD, Ramachandran, QMEAN, DOPE score e Verify3D, evidenciando que a proteína em si é bastante conservada. Dessa forma, ao analisar a interação da proteína com o carotenoide hECN dos modelos de OCPs estudadas neste trabalho, mostrou que todos os complexos permaneceram estáveis durante a simulação de DM. Por fim, houve a formação de um agrupamento dessas OCPs com outras do tipo OCP1 em um clado específico, podendo-se inferir que suas funcionalidades são semelhantes por conta da estrutura conservadas das OCPs. Também foi possível observar que existe conservação estrutural de um grupo de resíduos, tais como Trp110, Arg155 e Leu107. Com isso, a busca por novas formas de OCPs obtidas de outras cianobactérias podem revelar novas aplicações para esta proteína.

Palavras-chave: Diversidade, Análise filogenética, Modelagem, Simulação.

ABSTRACT

The Orange Carotenoid Protein has the function of allowing cells to avoid possible damage caused by light, through the photoprotection mechanism. The objective of this work was to analyze the diversity and characterize the structural and functional properties of Orange Carotenoid Protein (OCP) isolated from Amazonian Cyanobacteria, in order to validate them for use in biotechnology. The sequences of OCPs were obtained through annotation of genomic data made after DNA extraction, sequencing, assembly and separation of cyanobacterial genomes from two different Amazonian environments. For the phylogenetic analysis, a search was made in BLASTp for other OCPs and the phylogeny was assembled using the MEGA X program. The structural analysis was performed through the comparative modeling of the OCPs in the MODELLER program, using the protein of Tolypothrix sp. PCC 7601 as a template and the generated models were validated according to parameters of Ramachandran, QMEAN, RMSD and Verify3D. All generated homology models were submitted to Molecular Dynamics (DM) simulations, with a refinement of 100 ns. It was observed that the phylogeny corroborated the phylogenetic relationships of the OCPs studied, and evidenced their functionality through the formation of a unique grouping with OCPs of type OCP1. When the models built from the selected template were evaluated, all were successfully validated, with excellent values of RMSD, Ramachandran, QMEAN, DOPE score and Verify3D, evidencing that the protein itself is highly conserved. Thus, when analyzing the interaction of the protein with the carotenoid hECN of the OCP models studied in this work, it was shown that all the complexes remained stable during the MD simulation. Finally, there was the formation of a grouping of these OCPs with others of the OCP1 type in a specific clade, and it can be inferred that their functionalities are similar due to the conserved structure of the OCPs. It was also possible to observe that there is structural conservation of a group of residues, such as Trp110, Arg155 and Leu107. Thus, the search for new forms of OCPs obtained from other cyanobacteria may reveal new applications for this protein.

Keywords: Diversity, Phylogenetic analysis, Comparative modeling, Simulation.

SUMÁRIO

1.	INTRODUÇÃO	. 14
1.1	CLASSIFICAÇÃO DAS CIANOBACTÉRIAS	. 16
1.2	FOTOSSÍNTESE CIANOBACTÉRIANA	. 17
1.3	FICOBILIPROTEÍNAS	. 20
1.4	PROTEÍNA CAROTENOIDE LARANJA	. 22
2.	OBJETIVOS	. 33
2.1	OBJETIVO GERAL	. 33
2.2	OBJETIVOS ESPECÍFICOS	. 33
3.	MATERIAIS E MÉTODOS	. 34
3.1	AMOSTRAGEM	. 34
3.2	SEQUÊNCIAS DE OCP	. 35
3.3	ANÁLISE FILOGENÉTICA	. 35
3.4	MODELAGEM POR HOMOLOGIA	. 36
3.4.	1 Seleção do molde	. 36
3.4.	2 Construção e validação dos modelos	. 37
3.5	DINÂMICA MOLECULAR	. 38
4.	RESULTADOS E DISCUSSÃO	. 42
4.1	OBTENÇÃO DAS SEQUÊNCIAS DE OCP E ANÁLISE FILOGENÉTICA	. 42
4.2	MODELAGEM POR HOMOLOGIA	. 44
4.2.	1 Seleção do molde	. 44
4.2.	2 Construção e validação dos modelos	. 46
4.3	DINÂMICA MOLECULAR	. 54
5.	CONCLUSÃO	. 62
6.	REFERÊNCIAS BIBLIOGRÁFICAS	. 63

LISTA DE FIGURAS

Figura 1 – Diversidade morfológica em cianobactérias14
Figura 2 – Ilustração do Fotossistema Cianobacteriano
Figura 3 - Representação do complexo antena, que pode ser constituído por diversos tipos
de pigmentos com diferentes espectros de absorção da energia da radiação luminosa, e a
transferência dessa energia absorvida entre os pigmentos, através da ressonância, até
alcançar o centro de reação 19
Figura 4 – Composição estrutural do ficobilissomo
Figura 5 – Estrutura cristalina de Synechocystis sp. PCC 6803
Figura 6 – Ligação entre domínio N-terminal/C-terminal e hECN
Figura 7 – Mecanismo de fotoproteção da OCP com a participação da FRP 25
Figura 8 - Diversidade de homólogos de domínio do OCP: HCPs e CTDHs 28
Figura 9 - Valores médios de RMSD (em Å) para o NTD calculado usando como
referência o NTD das estruturas cristalinas de OCPO (linha laranja) e RCP (linha
vermelha) e para o CTD calculado usando como referência o CTD da estrutura cristalina
de OCP ^O (linha azul)
Figura 10 - Representação gráfica do complexo OCP:Ligante a partir de sua estrutura
cristalina (4XB5)
Figura 11 - Mapa da localização do lago da usina e do lago Bolonha
Figura 13 – Visualização das sequências no ESPript 45
Figura 14 – Alinhamento dos modelos de OCP com o molde no programa PyMol 47
Figura 15 – Pontuação QMEAN das OCPs modelos, convertido para Z-Score 50
Figura 16 – Gráficos de Ramachandran gerados no MolProbity para cada modelo 52
Figura 17 – Comparação entre gráficos RMSD em Angstrom dos modelos com o molde
em simulações de DM 55
Figura 18 – Contribuição individual de energia por resíduos de acordo com o método
MM-GBSA na complexação OCP-ligante 58

LISTA DE TABELAS

Tabela 1 – Lista das sequências das OCP analisadas no presente estudo
tabela 2 – Seleção dos modelos das OCPs com base na pontuação Molpdf e DOPE score.
Tabela 3 – Parâmetro RMSD de validação dos modelos de OCP das linhagens analisadas
neste estudo
Tabela 4 – Avaliação de validação dos modelos com valores acima de 90% dos resíduos
em regiões favoráveis por Ramachandran 53
Tabela 5 – Avaliação da qualidade do enovelamento no servidor Verify3D. 54

NOMENCLATURA

- ALO Aloficocianina
- BLAST Basic Local Alignment Search Tool
- CACIAM Coleção Amazônica de Cianobactérias e Microalgas
- CTD C-Terminal Domain
- DM Dinâmica Molecular
- FE Ficoeritrina
- FICO Ficocianina
- FRP Fluorescence Recovery Protein
- hECN 3'- hydroxyechinenone
- ICT Intramolecular Charge Transfer
- IMG Integrated Microbial Genomes & Microbiome Samples
- MEGA Molecular Evolutionary Genetics Analysis
- NCBI National Center for Biotechnology Information
- NGS Next Generation Sequencing
- NPQ Non-Photochemical Quenching
- NTD N-Terminal Domain
- **OCP** Orange Carotenoid Protein
- PBS Ficobiliproteínas
- PDB Protein Data Bank
- PLP piecewise linear potential
- QMEAN Qualitative Model Energy Analysis
- RMSD Root Mean Square Deviation

1. INTRODUÇÃO

Estima-se que as cianobactérias habitem a superfície terrestre há cerca de 3,5 bilhões de anos (SCHOPF, 2000), tempo suficiente para que esses organismos se adaptassem, tornando-se resistentes às mais adversas condições ambientais. Isso é notável pela grande variabilidade de habitats ocupados pelas cianobactérias e também pela grande diversidade fisiológica e morfológica (GRAHAM & WILCOX, 2000), a diversidade morfológica e de arranjo celulares é mostrada por alguns representantes do grupo na figura 1.



Figura 1 – Diversidade morfológica em cianobactérias. a - Chroococcus subnudus, b -Ch. limneticus, c - Cyanothece aeruginosa, d – Snowella litoralis, e - Microcystis aeruginosa, f - Pleurocapsa minor, g - Planktothrix agardhii, h - Limnothrix redekei, i -Arthrospira jenneri, j - Johanseninema constricum, k - Phormidium sp., l e m -Oscillatoria sp., n - Schizothrix sp., o - Tolypothrix sp., p - Katagnymene accurata, q -Dolichospermum planctonicum, r - Dolichospermum sp, s - Nostoc sp, t - Nodularia

moravica, u e v - *Stigonema* sp. Barra de escala: $a - u = 10 \mu m$, $v = 20 \mu m$. Fonte: Dvořák *et al.* (2015).

O filo Cyanobacteria é constituído por bactérias gram-negativas capazes de realizar fotossíntese e, em alguns casos, fixadoras de nitrogênio. São organismos cosmopolitas, podendo ser encontrados nos mais variados ambientes, aquáticos ou terrestres, como florestas tropicais, desertos, manguezais e lagos na Antártica (COHEN & GUREVITZ, 2006). Além disso, é um dos mais antigos grupos de organismos na Terra, com um registro fóssil de organismos semelhantes à cianobactérias que remonta ao surgimento da vida (PATTANAIK & LINDBERG, 2015).

Essa longa história evolutiva e sua adaptação aos mais diversos ambientes proporcionou a este filo uma grande diversidade morfológica, fisiológica e ecológica. Quanto à morfologia, as cianobactérias podem se apresentar como unicelulares ou em forma filamentosa ou colonial (SINGH *et al.*, 2011).

Segundo Sá *et al.* (2010), as cianobactérias são organismos procarióticos capazes de fixar carbono através da fotossíntese, fazendo parte da comunidade fitoplanctônica e contribuindo, deste modo, com grande parte da produtividade primária e do fluxo de energia em ecossistemas aquáticos. Estes microrganismos habitam uma grande variedade de ambientes (dulcícolas, salobros, marinhos e terrestres) e estão presentes em todos os biótopos aquáticos (interface água/ar, coluna d'água e sedimento).

Estes organismos habitam uma grande variedade de ambientes sejam estes, dulcícolas, salobros, marinhos e terrestres, porém os habitats com maior ocorrência de cianobactérias se encontram nos ecossistemas de água doce (SANT'ANNA *et al.*, 2006) por isso são excelentes bioindicadores em estudos de caracterização ambiental e poluição de diferentes ecossistemas aquáticos, pois, apesar de viverem em ambientes extremos, elas apresentam caráter dinâmico e sensível frente às mudanças físico-químicas do ambiente (COSTA *et al.*, 2017).

Apesar dos estudos de manipulação genética em linhagens de cianobactérias crescerem lentamente, os últimos avanços em biologia sintética e na compreensão da fisiologia desses microrganismos fizeram deles uma rica fonte em produtos naturais a serem explorados pela biotecnologia (LINDBERG *et al.*, 2010; SINGH *et al.*, 2011; DITTMANN *et al.*, 2015). Podem ser utilizadas na obtenção de diversos produtos, como fármacos com diferentes atividades (antiviral, anti-inflamatória, anticâncer, antimalárica, antifúngica, entre outras), marcadores fluorescentes, biopigmentos, enzimas,

antioxidantes, exopolissacarídeos usados como gelificantes, emulsificantes, floculantes e hidratantes, proteases, inibidores de proteases (AVER *et al.*, 2015).

1.1 CLASSIFICAÇÃO DAS CIANOBACTÉRIAS

Inicialmente, as cianobactérias foram classificadas de acordo com os padrões de classificação botânica. Com a descoberta das características procarióticas desses organismos, eles também passaram a ser regidos por normas bacteriológicas (STANIER *et al.*, 1978). Caracterizam-se por serem organismos autotróficos fotossintéticos e procarióticos pertencentes ao domínio Bacteria. Alguns são unicelulares, outros formam filamentos ramificados e espécies raras formam placas ou colônias irregulares.

A classificação atual das cianobactérias inclui a abordagem polifásica que se baseia em análises moleculares, morfológicas, fisiológicas, bioquímicas e ecológicas. De acordo com a abordagem polifásica, Anagnostidis e Komárek (1985) tentaram conciliar o sistema botânico e bacteriológico e propuseram quatro ordens (Chroococcales, Nostocales, Oscillatoriales e Stigonematales), subdivididos em famílias, subfamílias, gêneros e espécies (KOMÁREK, 2006).

O trabalho taxonômico baseado apenas em caracteres moleculares ou morfológicos tem capacidade limitada para reconhecer a importância ecológica de diferentes genótipos, variabilidade morfológica *in situ*, processos de adaptação em andamento e a origem contínua de novos ecótipos e morfótipos de cianobactérias. Portanto, a combinação de abordagens moleculares e morfológicas para a taxonomia moderna de cianobactérias é essencial (KOMÁREK, 2006).

Rippka e colaboradores (1979) realizaram estudos comparando a morfologia de 178 linhagens de cianobactérias e classificaram os gêneros em 5 grupos (Chroococcales, Pleurocapsales e Oscillatoriales não têm o suporte de estudos filogenéticos, Nostocales e Stigonematales, são monofiléticos, compondo as cianobactérias heterocísticas): as espécies unicelulares são classificadas nas subseções I e II, que são diferenciadas pela sua capacidade de reproduzir através de fissão binária ou múltipla. As espécies multicelulares estão nas subseções III-V. As cianobactérias classificadas na subseção III possuem células vegetativas, enquanto que as subseções IV e V são diferenciadas pela sua capacidade de se reproduzir em filamentos falsos ou verdadeiramente ramificados, respectivamente. Atualmente o grupo é composto por 414 gêneros e 1619 espécies, dos quais 394 gêneros e 1553 espécies possuem taxonomia validada por pelo menos um dos dois códigos, isto é, bacteriano e botânico (HAUER & KOMÁREK, 2021).

A filogenia molecular das cianobactérias, baseada em genomas e genes marcadores, tem passado por várias revisões na última década (DEXTRO *et al.*, 2021). O avanço das técnicas de sequenciamento de nova geração (NGS) e redução dos custos por base sequenciada levaram a um aumento exponencial da disponibilidade de genomas em bancos de dados públicos (ZHAO *et al.*, 2020). Porém, a montagem de genomas de cianobactérias exigiu avanço significativo das ferramentas de bioinformática para superar a dificuldade do número de linhagens não-cultiváveis e das culturas não-axênicas. Esse fato contribui para o número reduzido de genomas de cianobactérias acessíveis em relação à disponibilidade total de genomas de procariotos (DEXTRO *et al.*, 2021).

As cianobactérias somam 653 montagens disponíveis no *National Center for Biotechnology Information Search database* (GenBank), em relação ao total de 31.889 genomas de bactérias (acesso em 10/10/2021). Esse número é aproximadamente seis vezes maior do que há cinco anos, porém ainda há que se considerar a qualidade desses dados (WALTER, *et al.*, 2017). Em 2018 uma análise realizada em 440 genomas revelou que aproximadamente 5% deles apresentavam contaminação com sequências de microrganismos de outros grupos (CORNET *et al.*, 2018).

1.2 FOTOSSÍNTESE CIANOBACTERIANA

As cianobactérias destacam-se por serem produtores primários bem como os primeiros organismos capazes de realizar a fotossíntese oxigênica. Elas possuem clorofila A e os fotossistemas I (PSI) e II (PSII) realizando, assim, a fotossíntese em presença de oxigênio. São também as únicas a possuírem ficobilissomos, com os pigmentos acessórios ficoeritrina, ficocianina e aloficocianina (REVIERS, 2006).

A fotossíntese representa o principal modo de metabolismo energético destes organismos. Os pigmentos fotossintéticos (clorofila A, ficocianinas, aloficocianinas e ficoeritrinas) estão localizados nas lamelas fotossintéticas dos tilacóides, com movimento livre no citoplasma (na periferia celular), ao contrário dos organismos eucariotos fotossintetizantes onde a fotossíntese se processa no interior de organelas especializadas, os cloroplastos (MUR *et al.*, 1999).

Com efeito, a fotossíntese requer primeiramente a hidrólise da água, ou seja, a quebra da molécula para extração de elétrons, além de uma ação coordenada dos dois fotossistemas (PSI e PSII) gerando, por fim, um gradiente de prótons para produção de energia para células na forma de ATP. No caso específico das cianobactérias, a cadeia transportadora de elétrons caracteriza-se por ser mista (CTE MISTA), abrigando em um mesmo trecho de membrana componentes das cadeias fotossintética e respiratória (TRIPP *et al.*, 2010).

As cianobactérias, podem realizar fotossíntese oxigênica ou anoxigênica, a fim de produzir ATP através da fotofosforilação. Neste processo, moléculas orgânicas são sintetizadas (anabolismo) a partir de reações envolvendo a luz, CO2 água (ou H₂S) e compostos de baixa energia. A fotossíntese oxigênica realizada por estes microrganismos é semelhante àquela realizada em plantas, processo pelo qual há síntese de compostos orgânicos a partir da presença de luz, água e gás carbônico (PACIFICO *et al.*, 2016), conforme ilustrado na figura 2.



Figura 2 – Ilustração do Fotossistema Cianobacteriano. Processo semelhante da mitocôndria. Fonte: Adaptado Voet & Voetóva (1995).

A absorção da energia proveniente da radiação eletromagnética se dá através de uma estrutura denominada de complexo antena, composta por pigmentos como clorofilas e carotenoides, e também por proteínas (MARTINS, 2011). Uma vez a energia absorvida pelos pigmentos, ocorre a excitação dessas moléculas, tornando-as carregadas energeticamente. Como esses pigmentos não possuem a capacidade de armazenar essa energia, ela precisa ser transferida para outras moléculas, através de um processo conhecido como ressonância indutiva, ou será perdida na forma de fluorescência e/ou calor. A ressonância é um processo físico pelo qual um pigmento consegue transferir energia para outro através da vibração das moléculas em uma determinada frequência. Essa transferência se dará entre os pigmentos até alcançar um complexo pigmento-proteico conhecido como centro de reação, formando uma espécie de funil, concentrando a energia em um determinado ponto (MCCORMICK *et al.*, 2015).

As proteínas que fazem parte do complexo antena são essências nesse processo de captação da radiação eletromagnética por terem o papel de organizar os pigmentos de forma a facilitar a transferência de energia entre as moléculas mantendo a disposição destas na forma de funil (Figura 3).



Figura 3 - Representação do complexo antena, que pode ser constituído por diversos tipos de pigmentos com diferentes espectros de absorção da energia da radiação luminosa, e a transferência dessa energia absorvida entre os pigmentos, através da ressonância, até alcançar o centro de reação. No centro de reação, a molécula de clorofila a reduzida (Clo a-) absorve essa energia, passando a um estado excitado (Clo-*), sendo logo em seguida oxidada (Clo a+), doando elétrons para a cadeia de transporte de elétrons. Os elétrons

poderão ser repostos pela oxidação da água, levando a formação de H+ e O². Fonte: adaptado de McCormick *et al.* (2015).

1.3 FICOBILIPROTEÍNAS

As ficobiliproteínas (FBP) são proteínas providas de cor e solúveis em água (MORAES *et al.*, 2010). Essas proteínas podem ser isoladas como complexos pigmentares por serem fluorescentes (ABALDE *et al.*, 1998) e representam aproximadamente 20% do peso seco total da biomassa cianobacteriana (PRASANNA *et al.*, 2010).

As ficobiliproteínas são um grupo de proteínas que possuem um cromóforo associado, responsáveis pela captação de luz para fotossíntese em cianobactérias e em algumas algas. Eles são divididos em quatros tipos principais: ficoeritrina, ficocianina, ficoeritrocianina e aloficocianina, e são caracterizados de acordo com sua estrutura e absorção da qualidade da luz (PAGELS *et al.*, 2019). Nas cianobactérias, existem principalmente três ficobiliproteínas: ficocianina e aloficocianina, que são azuis, e a ficoeritrina de cor vermelha (GANTT, 1981).

Segundo Grossman (1993), as ficobiliproteínas formam a família de macromoléculas acessórias coletoras de luz, organizando-se em complexos supramoleculares que são denominados ficobilissomas (Figura 4), que atuam como componentes do aparelho fotossintético em cianobactérias e em parte de algumas algas eucarióticas. As ficobiliproteínas de cianobactérias têm sido descritas como potenciais compostos bioativos e reconhecidas como produtos naturais de alto valor para aplicações biotecnológicas (PAGELS *et al.*, 2019).



Figura 4 – Composição estrutural do ficobilissomo. Fonte: Adaptado Castro et al. (2018).

Alguns estudos têm identificado vários organismos marinhos como potencial fonte de componentes bioativos com aplicações biotecnológicas e, num período de cerca de 30 anos, a indústria biotecnológica das cianobactérias tem crescido e expandido para as cianobactérias de água doce como por exemplo os géneros *Anabaena, Aphanizomenon* e *Microcystis* (KUMAR *et al.*, 2019).

Os compostos biológicos presentes nas cianobactérias, constituem fontes nutritivas alternativas aos alimentos convencionais, podendo ter benefícios para a saúde humana e animal (CHRISTAKI *et al.*, 2011; RAPOSO *et al.*, 2013). As cianobactérias produzem também diversas enzimas, como o beta-lactase e a protease, que podem ser comercialmente explorados, com a vantagem de que os custos de produção da biomassa de cianobactérias são inferiores aos de bactérias (THAJUDDIN & SUBRAMANIAN, 2005).

Esses compostos supracitados são utilizados como marcadores fluorescentes de células em pesquisas biomédicas, terapêutica e em técnicas que são sensíveis à fluorescência (GLAZER, 1994). Silva (2008) relata que na faixa de pH de 5,0 e 7,5 estas proteínas são estáveis e em temperaturas mais baixas esta estabilidade permanece por um tempo mais prolongado.

A importância industrial das FBP provém das propriedades relacionadas aos altos coeficientes de absorbância, fluorescência, e estabilidade e características como cor única, natureza proteica, não-tóxica e capacidade antioxidante, que as tornam importantes tanto ecologicamente quanto economicamente (RASTOGI *et al.*, 2015). Uma aplicação interessante das FBP é como corantes naturais, que desempenham papel importante na indústria alimentícia (RIZZO *et al.*, 2015), substituindo parcial ou completamente os corantes artificiais, agregando valor nutricional e reduzindo a toxicidade dos alimentos. Destaca-se para esta finalidade a ficocianina, após purificação (PRADO & GODOY, 2003).

No âmbito clínico e biotecnológico, as FBP são aplicadas para o diagnóstico de doenças como marcadores fluorescentes. Tal aplicação se deve a suas propriedades espectroscópicas, tendo inúmeras aplicações em imunoensaios, histoquímica, citometria de fluxo, detecção de espécies reativas de oxigênio. Além disso, atuam como agentes terapêuticos, sendo que sua aplicação como antiviral, antifúngica, antibacteriana e antitumoral tem sido explorada. Para este fim, necessitam de alto grau de pureza (YADAV *et al.*, 2011).

A ficocianina (FC) e a aloficocianina (AFC) são as ficobiliproteínas encontradas em maior quantidade nas cianobactérias, representando cerca de 20% da proteína total das mesmas. A FC é o principal constituinte dos ficobilissomos, sendo localizada ao seu redor, e a AFC está contida no núcleo dos ficobilissomos e é o pigmento de ligação entre os ficobilissomos e a lamela fotossintética. As outras ficobiliproteínas, ficoeritrina (FE) e ficoeritrocianina (FEC) são encontradas nas extremidades da cianobactéria. A FE é flexível e responsável por facilitar a adaptação a mudanças ambientais, captando a luz do ambiente (JOHNSON *et al.*, 2014).

1.4 PROTEÍNA CAROTENOIDE LARANJA

A Proteína Carotenoide Laranja (OCP) é uma proteína de origem cianobacteriana, consideravelmente solúvel, e está vinculada ao fotossistema II, pois se localiza na antena do ficobilissomo (PBS). Possui peso molecular de 35 kDa, e se liga de forma nãocovalente a um carotenoide denominado de 3'-hidroxiequinenona (hECN), que é derivado do β -caroteno. A função da OCP é por meio do mecanismo de fotoproteção, permitindo que as células se evitem possíveis danos causados pela luz, além de inibir o crescimento causado por luz intensa ou estresses de nutrientes (KIRILOVSKY & KERFELD, 2012; BAO *et al.*, 2017).

A OCP foi descoberta por David Krogmann e colaboradores em extratos brutos de cianobactérias insolúveis; foi purificada de três gêneros diferentes de cianobactérias, a saber: *Arthrospira*, *Aphanizomenon* e *Microcystis* (HOLT & KROGMANN, 1981), fornecendo a primeira descrição, em cianobactérias, de uma carotenoproteína solúvel em água.

O acúmulo de evidências sugere que a OCP evoluiu a partir da fusão de duas proteínas de ligação a carotenoides primitivas (LECHNO-YOSSEF *et al.*, 2017; MOLDEN-HAUER *et al.*, 2017). Usando as primeiras sequências disponíveis de um genoma cianobacteriano, Krogmann e colaboradores (1997) foram capazes de identificar a OCP como o produto do gene OCP1 do *locus* slr1963 em *Synechocystis sp.* PCC 6803, com aproximadamente 34 kDa (Figura 5).



Figura 5 – Estrutura cristalina de *Synechocystis* sp. PCC 6803. Fonte: Adaptado de Lechno-Yossef *et al.* (2017).

A OCP possui um domínio α -hélice N-terminal (NTD, de 15-165 resíduos), efetor, e um domínio α/β C-terminal (CTD, de 190-317 resíduos), regulatório, unidos por um *linker* flexível e que interagem através de duas regiões distintas. Os 19 primeiros aminoácidos N-terminais (que formam uma pequena α -hélice) se estendem ao longo do domínio N-terminal e interagem com a face hidrofóbica exposta folha- β do domínio Cterminal. O grupo ceto é essencial para a função; é ligado por hidrogênio ao CTD por meio dos resíduos conservados Y201 e W288 (KERFELD *et al.*, 2003; WILSON *et al.*, 2008).

Ambos os domínios interagem com o carotenoide hECN. O grupo carbonila do hECN se liga aos resíduos Tyr201 e Trp288 do domínio C-terminal através de ligações de hidrogênio, enquanto o anel hidroxila está aninhado dentro de um grupo de resíduos aromáticos conservados (Trp41, Tyr44, Trp110) no domínio N-terminal (POLIVKA *et al.*, 2012), como mostra a figura 6.



Figura 6 – Ligação entre domínio N-terminal/C-terminal e hECN. Fonte: Carbon *et al.* (2015).

Esta interação entre a proteína e o carotenoide altera a conformação do ligante e suas propriedades espectrométricas. O anel ceto-terminal do hECN possui uma configuração s-*cis* quando em solução, porém assume uma configuração s-*trans* quando está no sítio ativo da OCP. Isso causa a redução do tempo no estado S1 (um estado eletrônico excitado, que antecede o estado eletrônico excitado permanente – S2), a alteração da energia S1 e também para a estabilização de um estado de transferência intramolecular de carga (*Intramolecular Charge Transfer* – ICT) de hECN (POLIVKA *et al*, 2012).

A presença de ligações de hidrogênio entre o grupo carbonila do carotenoide e a proteína desempenha um papel crucial na modulação e estabilização do estado ITC. Para hECN, este estado é observado exclusivamente na OCP (KIRILOVSKY & KERFELD, 2012).

A presença do grupo carbonila em hECN e sua interação com a OCP são essenciais para a fotoatividade. Na fotoativação com forte luz azul-verde, a OCP sofre uma mudança conformacional e converte-se de uma forma laranja, estável (OCP^O) para uma forma vermelha e metaestável (OCP^R) que facilita a dissipação térmica da energia de excitação da luz resultante da antena através da interação com o ficobilissomo (PBS) (KIRILOVSKY & KERFELD, 2012; BAO *et al.*, 2017). Esta dissipação de energia como calor nas antenas é um processo que ajuda a diminuir a quantidade de energia que atinge os centros de reação e depois reduz a probabilidade de fotoinativação do fotossistema (BOULAY *et al.*, 2008). Na ausência de luz, a OCP^R sozinha consegue retornar espontaneamente à forma OCP^O. Esta reversão é bastante afetada pela presença de outra proteína, *fluorescence recovery protein* (FRP). A FRP é uma proteína de 13-14 kDa fortemente ligada à membrana, que interage com a forma vermelha ativa da OCP. *In vitro*, esta interação acelera grandemente a conversão da OCP^R para a OCP^O. *In vivo*, a FRP é essencial para recuperar a capacidade total da antena, presumivelmente desempenhando um papel na separação do OCP^R do ficobilissomo (BOULAY *et al.*, 2010), como mostra a figura 7.



Figura 7 – Mecanismo de fotoproteção da OCP com a participação da FRP. Na escuridão / luz ambiente, OCP^O é desvinculado de seu local de ligação (X) no núcleo PBS. (2) Em condições de alta luminosidade, a absorção de um fóton por 3'-hECN desencadeia a fotoquímica iniciada pelo domínio C-terminal. (3) O domínio N-terminal e o carotenoide ligam-se ao PB e formam um OCP^R apagado. (4) Para dissociar o OCP^R, complexo de PB, FRP interage seletivamente com o domínio C-terminal de OCP^R. (5) O FRP catalisa a reversão escura do OCP para OCP^O, restaurando a estrutura terciária relativamente compacta desta forma de OCP e dissociando-a do sítio de ligação de PB. Fonte: adaptado de Kirilovsky *et al.* (2014).

Análises de genomas de cianobactérias mostraram que nas linhagens que possuíam o gene que codifica a OCP (*ocp1*), a maioria também possuía o gene para a FRP, que tipicamente é adjacente ao *ocp1*. Acredita-se que a FRP se tornou evolutivamente associada à OCP como uma forma de melhorar seu mecanismo de fotoproteção, sendo uma recente inovação na regulação da OCP (KIRILOVSKY & KERFELD, 2013; BAO *et al.*, 2017).

A proteína de ligação de carotenoides de laranja solúvel em água (OCP) de cianobactérias é um sistema ideal para o estudo dos efeitos do ambiente proteico nas propriedades fotofísicas dos carotenoides. Contém um único pigmento, o carotenoide 3'-hidroxiequinenona (hECN) (POLIVKA *et al.*, 2005).

Primeiro, as ligações de hidrogênio entre o grupo carbonil conjugado da molécula hECN e os resíduos Tyr203 e Trp290 foram observados na estrutura OCPs, sugerindo que são funcionalmente importantes (KERFELD, 2004). Além disso, embora a molécula hECN adote uma configuração trans na OCP, ela é arqueada, exibindo um desvio médio de 16° da conformação trans (KERFELD, 2003). Outra característica interessante da estrutura OCP é que a bolsa de ligação para o hECN é cercada por um número incomum de resíduos de metionina; isso resulta no posicionamento de seis átomos de enxofre dentro de 6,5 Å da molécula hECN.

A OCP também pode ser convertida em outra forma, a proteína carotenoide vermelha (RCP), que é caracterizada por um espectro de absorção desviado para o vermelho. A conversão de OCP para RCP é conseguida por proteólise ou por acidificação da OCP. Pode ser que a molécula de hECN no RCP esteja mais exposta ao solvente, causando mais desvio para o vermelho do espectro de absorção de hECN (KERFELD, 2004).

Polivka e colaboradores (2005) desenvolveram vários modelos para explicar as observações à luz de interpretações recentes de outros sistemas modelos. Dessa forma, foi possível evidenciar que a incorporação de hECN na OCP altera significativamente tanto a energia, quanto a dinâmica dos estados excitados de hECN, e que essas mudanças podem ser correlacionadas com as interações de carotenóides observadas na estrutura do OCP.

A fenda de ligação de hECN na OCP é revestida por vários aminoácidos apolares com o grupo carbonila acoplado em uma bolsa formada por resíduos polares. O ambiente da proteína, a ligação de hidrogênio através do grupo carbonila e a mudança conformacional da hECN induzida pela interação com a proteína sublinham as propriedades espectroscópicas incomuns da OCP. Portanto, uma das questões-chave é se a proteína fornece um ambiente que estabiliza o estado ICT em hECN (KERFELD, 2004).

A OCP pode interconverter entre duas formas: a forma laranja inativa (OCP^O) e a forma vermelha ativa (OCP^R). No escuro, OCP está "adormecido" em sua forma inativa de OCP^O com um espectro de absorção característico de 3'- hECN com pico em 475 nm

e 495 nm. No estado OCP, o pigmento carotenoide interage com NTD e CTD através de múltiplas forças moleculares fracas, como ligações de hidrogênio, interações hidrofóbicas e contatos de van der Waals (LOU *et al*, 2020).

A luz branca ou azul intensa leva o 3'- hECN a adotar uma conformação completamente diferente que quebra as interações com o CTD, resultando em uma separação completa de NTD e CTD. quase totalmente envolto em NTD com apenas os β - ceto anéis expostos (LEVERENZ *et al.*, 2015). OCP^R é metaestável e reverte espontaneamente para OCP^O no escuro. Acredita-se que OCP^R é a única forma de interagir com o PBS e induzir a extinção da fluorescência do PBS (BERERA *et al.*, 2012).

1.5 EVOLUÇÃO E DIVERSIDADE DAS OCPs

Desde a primeira descrição estrutural da OCP e o primeiro sequenciamento de genomas de cianobactérias, os domínios individuais de OCP foram reconhecidos como também codificados como polipeptídeos independentes (KERFELD, 2004). Pelo menos nove clados diferentes de homólogos ao NTD podem agora ser identificados filogenomicamente em diversas cianobactérias, com membros de diferentes clados muitas vezes co-ocorrendo no mesmo genoma (MELNICKI *et al.*, 2016).

Esses parálogos foram denominados "Proteínas Carotenoides Helicoidais" (HCPs), pois todos eles são previstos para conservar tanto a estrutura helicoidal quanto a bolsa de ligação de carotenoides do OCP-NTD (LOPEZ-IGUAL *et al.*, 2016). Os homólogos do domínio C-terminal (CTDHs) que são filogeneticamente basais às sequências OCP-CTD também são encontrados em quase todos os genomas que codificam um HCP, embora não pareçam ter cópias duplicadas e sua função permanece enigmática (MELNICKI *et al.*, 2016).

A expansão da família HCP em diferentes subtipos filogenéticos implica que os parálogos podem ter desenvolvido funções divergentes. A expressão e caracterização de quatro subtipos diferentes de *Nostoc* PCC 7120 mostrou que cada um é de cor vermelha como a OCP-NTD e RCP, confirmando sua capacidade de ligar carotenoides (LOPEZ-IGUAL *et al.*, 2016), conforme a figura 8.



Figura 8 - Diversidade de homólogos de domínio do OCP: HCPs e CTDHs. A) Filogenia Máxima Verossimilhança (*Maximum Likelihood*) de HCPs e OCP NTD. B) Filogenia Junção de Vizinhos (*Neighbor Joining*) da família OCP-CTD e NTF2. Fonte: adaptado de Bao *et al.* (2017).

Uma função específica ainda não foi demonstrada para HCP1, embora estruturas cristalinas com diferentes carotenoides sugiram, e sua capacidade de se ligar a uma variedade de carotenoides diferentes, incluindo mixoxantofila que contém um grupo glicosil volumoso terminal. Portanto, parece que uma gama mais ampla de carotenoides ligados a HCPs pode ser capaz de mediar a supressão de ${}_{1}O^{2}$, em comparação com a ligação mais específica de cetocarotenoides que é necessária para a supressão de antena induzível pela OCP (SEDOUD, *et al.*, 2014). Além disso, a ligação de HCP1 com mixoxantofila levou a propriedades ópticas alteradas em comparação com HCP1 ligado com cantaxantina. Essa capacidade da família HCP/NTD de se ligar a diferentes tipos de carotenoides implica o cromóforo como outro módulo que pode ser variado para ajustar as propriedades funcionais das proteínas que contêm esse domínio.

À luz da interação entre o NTD e o CTD da OCP, parece plausível que os CTDHs também possam interagir com alguns HCPs, formando um heterodímero que se assemelharia a OCP sem seu ligante. O *locus* do gene CTDH é comumente adjacente a um *locus* HCP4 ou HCP5 (MELNICKI *et al.*, 2016), que são os clados que se ramificam mais próximos das sequências primitivas OCP-NTD e incluem o parálogo *Nostoc* 7120 (All4941, um HCP4) que foi apenas HCP para demonstrar extinção no PBS (LOPEZ-IGUAL *et al.*, 2016). Essa evidência sugere que um antigo evento de fusão gênica entre um HCP e um CTDH pode ter ocorrido, dando origem a uma OCP ancestral.

Pelo menos duas famílias adicionais de parálogos da OCP de comprimento total também foram recentemente identificadas, ocorrendo em diversos genomas de cianobactérias na maioria das espécies filogenéticas.

Uma dessas novas famílias, OCP2, forma um clado distinto que mapeia entre a família primitiva "OCPX" e a família OCP1 canônica e bem caracterizada, sugerindo uma história evolutiva intermediária. Usando *Fremyella diplosiphon*, um modelo de cianobactéria de adaptação cromática que contém OCP1 e OCP2, as propriedades desses parálogos foram comparadas. A OCP2 foi capaz de extinguir o PBS, porém com propriedades substancialmente diferentes em relação ao OCP1; OCP2 mostrou fotodinâmica mais rápida, incapacidade de se acumular de forma estável na forma ativa de OCP^R à temperatura ambiente, ausência de dimerização e ausência de interação FRP (BAO *et al.*, 2018).

Em relação ao OCP2, o OCP1 tem vários níveis de controle regulatório fino, permitindo acesso flexível a uma ampla gama de atividades de extinção e, portanto, é mais refinado evolutivamente do que o OCP2 e OCPX recém-identificados. Isso é consistente com a hipótese de fusão de domínios e justifica a prevalência de OCP1 entre a maioria das cianobactérias.

Kuznetsova e colaboradores (2020), usando espectroscopia mostraram que apesar das diferenças nas características de fotoativação de OCP1 e OCP2, as propriedades espectroscópicas desses dois tipos de OCP são comparáveis, mostrando que a diferença não está diretamente relacionada às propriedades espectroscópicas do carotenoide ligado. Estruturalmente, foi descoberto que a OCP2 é menos flexível que OCP1, sugerindo que as diferenças no mecanismo de fotoativação se devem à rigidez da proteína. A primeira caracterização da nova família OCP, OCP2, apontou diferenças funcionalmente relevantes em relação ao OCP1 canônico, uma delas é a sua fotoconversão e retroreversão mais rápidas (BAO *et al.*, 2017).

Quando comparadas as propriedades de OCP1 e OCP2 de *Tolypothrix* por combinação de métodos espectroscópicos estruturais e ultrarrápidos, foi observado que, embora OCP1 e OCP2 exibam dinâmicas muito diferentes do fotociclo OCP^O-OCP^R e eficiência de supressão de ficobilissomas, sua fotofísica inicial é quase idêntica (KUZNETSOVA *et al.*, 2020).

1.6 DINÂMICA MOLECULAR DE OCPs

Técnicas experimentais sondaram a evolução da OCP após a fotoativação com alta resolução no tempo ou no espaço (MEZZETTI *et al.*, 2019). No entanto, uma imagem clara e completa caracterização dos mecanismos moleculares através dos quais o carotenoide transloca entre os dois domínios e a proteína dissociada não está disponível (GUPTA *et al.*, 2019).

Usando este paradigma, pode-se reproduzir toda a evolução do complexo, desde sua forma em repouso (OCP^O) até sua forma final aberta. As simulações feitas por Bondanza e colaboradores (2020) demonstram que o carotenoide atua como uma "trava" que une os dois domínios em OCP^O, estabelecendo interações de força iguais com os dois domínios (NTD-CTD). Esta trava, no entanto, é levantada assim que o movimento do carotenoide é permitido na direção do NTD graças à transição eletrônica, e a dissociação pode prosseguir (KONOLD *et al.*, 2019).

Quando analisado o RMSD, durante o processo, o NTD sofre pequenas, mas significativas modificações. Estes podem ser quantificados observando o RMSD do NTD ao longo da evolução da Variação Coletável (CV) quando calculado em relação ao OCP^O e RCP (Figura 9).

Existe um consenso geral sobre o fato que a dissociação é de longe o processo mais lento no fotoativação de OCP. Espera-se que esta etapa tenha grande contribuições entrópicas (devido à maior conformação liberdade dos dois domínios dissociados em OCP^R) que compensa a perda de interações de ligação entre as principais interfaces. Todos os resíduos de Trp em OCP são altamente conservados e (exceto W101) contribuem para o chamado "túnel de carotenoides", que fornece um ambiente hidrofóbico específico para o carotenoide (BONDANZA *et al.*, 2020), conforme mostra a figura 10.

A compreensão das características estruturais que influenciam propriedades como a força da associação de domínio, ajuste de propriedades ópticas, dinâmica de fotoconversão e interações com FRP ou PBS permitirá que estratégias baseadas em OCP se tornem parte do kit de ferramentas de biologia sintética, que conta com componentes modulares que podem ser trocados alternadamente. Componentes fotorreceptores modulares já revolucionaram o campo da optogenética (SHCHERBAKOVA *et al.*, 2015).

A engenharia de fotoproteção em cianobactérias pode ser um meio de melhorar seu desempenho como fábricas microbianas. A regulação do OCP foi identificada como um alvo para aumentar o desempenho e/ou eficiência das linhagens de produção (SANTOS *et al.*, 2014). Ao mesmo tempo, espera-se que uma compreensão abrangente da distribuição natural de OCPs/HCPs/CTDHs em relação aos metadados ambientais forneça informações fundamentais sobre como as cianobactérias se adaptam para ocupar uma ampla gama de nichos ecofisiológicos (KIRILOVSKY & KERFELD, 2013).



Figura 9 - Valores médios de RMSD (em Å) para o NTD calculado usando como referência o NTD das estruturas cristalinas de OCP^O (linha laranja) e RCP (linha vermelha) e para o CTD calculado usando como referência o CTD da estrutura cristalina de OCP^O (linha azul). Fonte: Adaptado de Bondanza *et al.* (2020).



Figura 10 - Representação gráfica do complexo OCP:Ligante a partir de sua estrutura cristalina (4XB5). Cores são usadas para mostrar diferentes domínios (magenta para NTE, amarelo para NTD, cinza para o *linker*, azul para CTD e verde para CTT). Ligante e alguns resíduos relevantes (Y44, W110, Y201, W277 e W288) são mostrados em amarelo e azul. Fonte: Bondanza *et al.* (2020).

2. OBJETIVOS

2.1 OBJETIVO GERAL

Avaliar a diversidade estrutural e funcional da Proteína Carotenoide Laranja (OCP) em Cianobactérias isoladas de ambientes Amazônicos.

2.2 OBJETIVOS ESPECÍFICOS

- Executar análise filogenética para analisar estruturalmente e funcionalmente as OCPs coletadas;
- Construir e validar modelos tridimensionais das OCPs nas linhagens estudadas;
- Caracterizar a estabilidade e afinidade do ligante no sítio ativo das proteínas encontradas.

3. MATERIAIS E MÉTODOS

3.1 AMOSTRAGEM

As linhagens utilizadas foram *Microcystis* sp. CACIAM 03, *Cyanobium* sp. CACIAM 14 e *Alkalinema* sp. CACIAM 70d, isoladas a partir de amostras de água do reservatório da Usina Hidrelétrica de Tucuruí – UHE Tucuruí (3°49'55''S e 49°38'50''W), localizado no Estado do Pará (Figura 11), e *Synechocystis* sp. CACIAM 05, *Nostoc* sp. CACIAM 19 e *Tolypothrix* sp. CACIAM 22, isoladas a partir de amostras de água do lago Bolonha (01°25'15''S e 48°26'02''W), localizado no município de Belém, Pará (Figura 11). Estas linhagens fazem parte da Coleção Amazônica de Cianobactérias e Microalgas (CACIAM) do Laboratório de Tecnologia Biomolecular da Universidade Federal do Pará.



Figura 11 - Mapa da localização do lago da usina e do lago Bolonha.

Os DNAs genômicos foram extraídos das culturas não axênicas das linhagens isoladas e sequenciados em duas plataformas de sequenciamento de nova geração – NGS (*Next Generation Sequencing*): Illumina MiSeq (Illumina, EUA) e Genome Sequencer FLX 454 (Roche, Suíça). As leituras (*reads*) brutas, provenientes dos sequenciamentos, foram filtrados qualitativamente com valor mínimo de phred20. Uma montagem conjunta de todas as leituras foi realizada pelo *software* Newbler v2.9 (Roche, Suíça), com os seguintes parâmetros: sobreposição mínima de 20 pb, identidade mínima da sobreposição de 80%, modo heterozigoto e opção de extensão de sobreposições de baixa cobertura ativados.

Desde que se tratou de cultura não axênica, o software MaxBin 2.0 (WU *et al.*, 2016), amplamente utilizado em análises metagenômicas, foi utilizado para separar as

sequências montadas de diferentes genomas (processo conhecido como *binning*). Para classificar taxonomicamente as sequências separadas (*bins*) obtidas, foi utilizado a ferramenta *Basic Local Alignment Search Tool* - BLASTp (protein-protein BLAST) para cada *bin*, nas sequências contendo os modelos ocultos de Markov para os genes essenciais identificados pelo MaxBin 2.0 (WU *et al.*, 2016), contra o banco de dados não redundante do NCBI.

3.2 SEQUÊNCIAS DE OCP

Foram obtidas seis sequências após anotação dos dados genômicos gerados na etapa anterior, e foram listadas conforme espécie, tamanho, ponto isoelétrico provável, massa (kDa) e domínio, além da atribuição de códigos para melhor identificação (Tabela 1).

Cádigo	Egnésis	Tamanho	PI	Massa	Domínio
Courgo	Espècie	(AA)		(kDA)	
OCP C03	Microcystis sp. CACIAM 03	320	4,96	35,51	Carot_N NTF2
OCP C14	Cyanobium sp. CACIAM 14	318	4,71	35,06	Carot_N NTF2
OCP C70d	Alkalinema sp. CACIAM 70d	319	5,04	35,31	Carot_N NTF2
OCP C05	Synechocystis sp. CACIAM 05	317	4,94	34,67	Carot_N NTF2
OCP C19	Nostoc sp. CACIAM 19	319	5,01	35,28	Carot_N NTF2
OCP C22	Tolypothrix sp. CACIAM 22	319	4,89	35,43	Carot_N NTF2

Tabela 1 – Lista das sequências das OCP analisadas no presente estudo.

AA = Aminoácidos; PI = Ponto Isoelétrico; kDA = kiloDalton.

3.3 ANÁLISE FILOGENÉTICA

As seis sequências de OCP obtidas na etapa anterior foram submetidas a ferramenta BLASTp, onde foram selecionadas outras sequências de OCPs que possuíam até 95% de percentual de identidade, além de duas buscas separadas utilizando como referência o gene *ocpX* e *ocp2*, respectivamente, em seguida todas as sequências foram submetidas em formato FASTA ao programa MEGA - *Molecular Evolutionary Genetics Analysis*, versão X (TAMURA *et al.*, 2018) para alinhamento múltiplo utilizando o método ClustalW (CHENNA *et al.*, 2003).
Os parâmetros utilizados no ClustalW foram no alinhamento em pares na penalidade de aberta de lacuna de 10,00 e na penalidade de extensão de lacuna de 0,10. Os parâmetros para alinhamento múltiplo foram 10,00 para penalidade de aberta de lacuna e 0,20 penalidade de extensão de lacuna. Foi utilizada uma matriz de distância com o método de estimativa de variância *bootstrap*, com 1000 replicatas. O *bootstrap* é um método de simulação baseado em dados para estimar o tamanho da amostra e analisar dados: incluindo teste de hipótese (valores p), erro padrão (SE) e estimativa de intervalo de confiança (CI); que envolve a retirada repetida de amostras aleatórias dos dados originais, com substituição (WALTER, 2017).

Após, foi construída uma árvore filogenética pelo mesmo programa utilizando o método de Máxima Verossimilhança, buscando gerar a árvore com a maior probabilidade de produzir e explicar os dados obtidos pelo alinhamento (BRAUN *et al.*, 2014). Dessa forma, para gerar uma inferência filogenética observável, e com isso mostrar o quão diverso são as espécies estudadas, foi utilizado o método de Máxima Verossimilhança (*Maximum Likelihood*). Então, são computadas todas as árvores possíveis, sendo escolhida aquela com maior probabilidade de que os resultados tenham origem de acordo com o modelo evolutivo selecionado, baseado em um determinado alinhamento (PEREIRA, 2012). O modelo de substituição aplicado para gerar a árvore foi o WAG (WHELAN & GOLDMAN, 2001).

3.4 MODELAGEM POR HOMOLOGIA

3.4.1 Seleção do molde

As sequências obtidas nos genomas analisados foram visualizadas utilizando a ferramenta ESPript, com base na estrutura. ESPript é um utilitário, cuja saída é um arquivo PostScript de sequências alinhadas com aprimoramentos gráficos. O programa calcula uma pontuação de similaridade para cada resíduo das sequências alinhadas. As sequências dentro de cada grupo são alinhadas de forma rápida.

A maior identidade entre as sequências, possibilitou a seleção de um único molde a ser utilizado para gerar o modelo tridimensional das seis sequências encontradas. Assim, as sequências foram submetidas ao *Protein Data Bank* (PDB) para seleção do molde. O molde escolhido foi *Tolypothrix* sp. PCC 7601 (ID 6PQ1), que apresentou maior identidade variando de 80 a 90% com as sequências de OCP obtidas anteriormente. A estrutura da proteína-molde apresenta duas cadeias, porém apenas uma delas foi utilizada na modelagem. A estrutura do molde estar resolvida pelo método de difração de raios X e apresentou os melhores valores de identidade e resolução, além de possuir a estrutura cristalizada do ligante hECN (45D).

3.4.2 Construção e validação dos modelos

Para gerar os modelos tridimensionais foi utilizado o programa MODELLER versão 10.2 (FISER & SALI, 2003). Para cada sequência de OCP (OCP C3, OCP C14, OCP C70d, OCP C05, OCP C19 E OCP C22), foram gerados 100 modelos baseados no alinhamento da sequência alvo com a proteína molde, considerando diferentes conformações e o melhor modelo foi classificado por função de densidade de probabilidade molecular (Molpdf) e DOPE *score*, posteriormente as estruturas 3D foram visualizadas no PyMol 2.5.2 (SCHRÖDINGER, 2000.).

A função Molpdf avalia a qualidade dos átomos selecionados no modelo usando o método DOPE (*Discrete Optimized Protein Energy*) (SHEN & ŠALI, 2006). DOPE utiliza a função de energia padrão do *software*. A pontuação do modelo DOPE é projetada para selecionar a melhor estrutura de uma coleção de modelos construídos pelo Modeller. A pontuação não é normalizada em relação ao tamanho da proteína e tem uma escala arbitrária, portanto, pontuações de diferentes proteínas não podem ser comparadas diretamente.

Para validar os modelos gerados foram analisados alguns parâmetros como sua qualidade estereoquímica, a qualidade do enovelamento, o valor de *Root Mean Square Deviation* (RMSD) entre a cadeia principal do molde e de cada modelo construído, e também, o valor de pontuação QMEAN (*Qualitative Model Energy Analysis*), que usa vários parâmetros estatísticos expressos como potenciais de força média: características geométricas do modelo (distâncias de pares atômicas, ângulos de torção, acessibilidade ao solvente) são comparadas com distribuições estatísticas obtidas de estruturas já resolvidas experimentalmente e será avaliado no servidor *online* SWISS-MODEL (BIASINI *et al.*, 2014).

A ideia de QMEAN é colocada conforme Benkert *et al.* (2011, p.3), QMEAN é uma função de pontuação composta que é capaz de derivar estimativas de qualidade absoluta globais (ou seja, para toda a estrutura) e locais (ou seja, por resíduo) com base em um único modelo.

O servidor MolProbity foi utilizado para gerar o gráfico de Ramachandran, e assim, analisar a qualidade estereoquímica do modelo. O gráfico de Ramachandran testa a qualidade de estruturas tridimensionais de proteínas, pois mostra os valores permitidos dos ângulos de torção ψ e ϕ para cada resíduo de aminoácido da proteína (CHEN *et al.*, 2010). A conformação da estrutura secundária de uma proteína pode ser expressa em função de seus diedros de espinha dorsal expressos em pares (ϕ , Ψ) que podem ser representados em um gráfico do tipo Ramachandran para uma interpretação mais fácil. Esses lotes são normalmente divididos em regiões permitidas (RAMAKRISHNAN & RAMACHANDRAN, 1965). Cerca de 40% de todos os aminoácidos na estrutura estão contidos em apenas 2% do gráfico de Ramachandran – as chamadas "áreas permitidas" (LOVELL, 2003).

Já a qualidade do enovelamento foi analisada pelo servidor Verify3D, que analisa a compatibilidade do modelo tridimensional com sua sequência de aminoácidos (sequência primária) e gera um perfil 3D. Esse perfil é uma tabela calculada a partir das coordenadas atômicas da estrutura, onde é mostrada a compatibilidade do modelo tridimensional da estrutura com qualquer sequência de aminoácidos. É necessário que no mínimo 80% dos resíduos tenham um *score* maior ou igual a 0,2 no perfil 3D/1D para que o modelo seja considerado bom (EISENBERG *et al.*, 1997).

Os melhores modelos de cada gene foram submetidos a simulações de Dinâmica Molecular (DM), para produção de uma trajetória de 100 ns.

3.5 DINÂMICA MOLECULAR

O ligante (45D – beta, beta-caroteno-4,4'-diona), considerado isômero da Cantaxantina, cuja formúla molecular é C₄₀H₅₂O₂, e sua incorporação na OCP altera de forma significativa a energia, e também a dinâmica dos estados excitados da enzima. As coordenadas do ligante foram obtidas da estrutura cristalizada do molde depositado no PDB, e foram complexadas utilizando o AMBER 18.

O campo de força aplicado foi o *ff14sb* (MAIER *et al.*, 2015) para a proteína e GAFF (BONDANZA *et al.*, 2020) para o ligante, respectivamente. Contra-íons Na+ ou Cl– foram adicionados para neutralizar as cargas e moléculas de água *tip3p* em uma caixa octogonal com 10 Å em cada direção da estrutura 3D da proteína.

A minimização de energia foi realizada em cinco etapas, quatro delas usando 3.000 ciclos de descida mais íngreme e 5.000 ciclos de gradientes conjugado para cada

modelo; os átomos pesados serão restringidos por um potencial harmônico de 1000 Kcal / mol * Å² (SIQUEIRA *et al*, 2018). Na última etapa, foi aplicado 5.000 ciclos de descida mais íngreme e 30.000 ciclos de gradientes conjugados e sem restrições. O estágio de aquecimento e equilíbrio foi dividido em 14 etapas. A temperatura foi aumentada gradativamente, até atingiu 300 K. Langevin Dynamics (termostato) foram empregados com uma frequência de colisão de 3,0 ps⁻¹.

Um potencial harmônico de 25 Kcal / mol*Å² foi empregado nas etapas iniciais e desligado durante a etapa 13. O procedimento de aquecimento durou 650 ps até a etapa 13 e foi realizada usando um conjunto NVT. Posteriormente, uma fase de equilíbrio de 2 ns foi empregada em um conjunto NPT. O algoritmo SHAKE foi empregado para restringir a vibração das ligações de todo os átomos de hidrogênio. O método *Particle Mesh Ewald* foi usado para calcular as interações eletrostáticas usando um valor de corte de 10,0 Å (SIQUEIRA *et al.*, 2018).

Para executar essas simulações de DM, um servidor PDB2PQR (http: //nbcr-222.ucsd. edu / pdb2pqr_2.0.0 /) foi usado para determinar o estado de protonação da proteína considerando um nível de pH de 7,0. Todas as etapas de preparação e produção de DM foram executadas usando os pacotes do programa AMBER 18. Por fim, foram realizados cálculos de energia livre de ligação MMGBSA e MMPBSA no próprio AMBER (CASE *et al*, 2021).

3.6 ANÁLISE DE ESTABILIDADE

As análises de estabilidade foram feitas observando o RMSD (*Root Mean Square Deviation*). Basicamente, quando é analisado o RMSD pode-se medir a distância média entre cada átomo nos modelos comparados, e é usado frequentemente como forma de comparar a similaridade entre ambas (VERLI, 2014). Outro ponto importante, é que os valores gerados por RMSD evidenciam informações acerca do equilíbrio do sistema, mostrando o quando a estrutura começa a assumir sua a conformação mais estável dependo da média (JOSHI *et al.*, 2017).

O RMSD pode ser calculado para qualquer tipo e subconjunto de átomos; por exemplo, átomos C α de toda a proteína, átomos C α de todos os resíduos em um subconjunto específico (por exemplo, as hélices transmembranares, bolsa de ligação ou um loop), todos os átomos pesados de um subconjunto específico de resíduos ou todos os

átomos pesados em um pequeno -ligantes moleculares (ABAGYAN & KUFAREVA, 2012).

A principal desvantagem do RMSD reside no fato de ser dominado pelas amplitudes dos erros. Duas estruturas que são idênticas com exceção de uma posição de um único loop ou um terminal flexível normalmente têm um grande *backbone* global RMSD e não podem ser sobrepostos de forma eficaz por qualquer algoritmo que otimize o RMSD global. Portanto, o RMSD usa como uma de suas avaliações de quão bem uma estrutura submetida corresponde à estrutura alvo conhecida.

O RMSD é útil para a análise de movimentos dependentes do tempo da estrutura. É frequentemente usado para discernir se uma estrutura é estável na escala de tempo das simulações ou se está divergindo das coordenadas iniciais. Na maioria das vezes, a divergência das coordenadas iniciais é interpretada como um sinal de que a simulação não está equilibrada. Quando uma simulação é equilibrada, ou seja, quando a estrutura de interesse flutua em torno de uma conformação média estável, faz sentido calcular as flutuações de cada subconjunto da estrutura (MARTINEZ, 2015).

3.7 ENERGIA LIVRE

Para a análise de energia livre foram usados os métodos de MM/GBSA e MM/PBSA. O tamanho de simulação pode influenciar os resultados de MM/GBSA dependendo do sistema utilizado e o tempo de convergência da simulação. Nos estudos entre a interação proteína e ligante por meio de DM, simulações entre 2-6 ns são amplamente utilizadas (XU *et al.*, 2013; KUMARI *et al.*, 2015), e por este motivo foi dada maior importância para esta faixa de tempo.

A energia livre é muito importante para o entendimento de processos de solvatação, descrevendo a tendência à associação e a espontaneidade do processo. A capacidade de prever o valor dessa propriedade usando modelos de química teórica é de grande valia, embora a obtenção de estimativa mais acurada envolva um custo computacional muito elevado (GENHEDEN *et al.*, 2010).

O tempo de 10 ns foi o escolhido para o restante da metodologia, concordando com o tempo estipulado anteriormente para o GBSA e o tempo preconizado pela literatura na área (HOU *et al.*, 2011; XU *et al.*, 2013; KUMARI *et al.*, 2014).

Em MMPBSA e MMGBSA, a energia livre é avaliada como uma soma de um termo de energia conformacional (a parte MM - mecânica molecular) e um termo de

solvatação sem energia que é calculado usando eletrostática contínua (tornando estes métodos abordagens implícitas de solvatação). MM refere-se ao tipo de função de energia usada para calcular a energia potencial de uma estrutura molecular (FERREIRA, 2017).

4. **RESULTADOS E DISCUSSÃO**

4.1 OBTENÇÃO DAS SEQUÊNCIAS DE OCP E ANÁLISE FILOGENÉTICA

As OCPs (OCP C03, OCP C14, OCP C70d, OCP C05, OCP C19 e OCP C22) foram alinhadas com outras OCPs selecionadas através do BLASTp, totalizando 20 sequências (OCP1, OCP2 e OCPX). Essas sequências coletadas foram inseridas no programa MEGA X, e todas foram usadas no alinhamento utilizando ClustalW. O alinhamento realizado foi utilizado para montar a filogenia, com intuito de corroborar as relações filogenética (Figura 12).



Figura 12 – Árvore filogenética das sequências de OCPs. Os círculos em azuis são as sequências analisadas neste estudo.

Com base na árvore filogenética produzida, observou-se que as sequências de OCPs (OCP C03, OCP C14, OCP C70d, OCP C05, OCP C19 e OCP C22) se agruparam em um clado com OCPs do tipo OCP1 com valor de suporte do nó de 100, mostrando que o clado apareceu em 100% das replicatas produzidas. OCP2 e OCPX agruparam mais próximas entre si, contudo, em clados distintos, com valor de suporte do nó de 100 e 90, onde 100% e 90% das replicatas suportaram os clados OCP2 e OCPX, respectivamente.

Esses achados corroborando Bao *et al.* (2017), que identificaram estas duas famílias adicionais de parálogos de OCP em diversos genomas cianobacterianos.

Uma dessas novas famílias, OCP2, forma um clado distinto que mapeia entre a família primitiva "OCPX" e a família OCP1, sugerindo uma história evolutiva intermediária. Bao e colaboradores (2019), analisaram algumas OCP1 e OCP2, e observaram que a OCP2 foi capaz de extinguir o PBS, porém com propriedades substancialmente diferentes em relação a OCP1. Com isso, a OCP2 mostrou fotodinâmica mais rápida, incapacidade de se acumular de forma estável no ativo OCP^R forma-se à temperatura ambiente, ausência de dimerização e ausência de interação FRP (BAO *et al.* 2017).

Em relação a OCP2, a OCP1 possui vários níveis de controle regulatório fino, permitindo flexibilidade acesso a uma gama mais ampla de atividade de extinção e, portanto, é mais refinada evolutivamente do que as recém-identificadas OCP2 e OCPX. Isso é consistente com a hipótese de fusão de domínios e justifica a prevalência de OCP1 entre a maioria das cianobactérias (BAO *et al.* 2017).

Já no trabalho de Muzzopappa *et al.* (2019), a análise da árvore sugeriu que houve uma diversificação inicial que deu origem a um ancestral OCPX e ao ancestral comum de OCP1 e OCP2. A diversificação deste ancestral levou as atuais proteínas OCP1 e OCP2. Contudo, foi possível mostra que o clado OCP2 não é evolutivamente mais antigo que OCP1, como proposto anteriormente (BAO *et al.*, 2019).

Além disso, para estudar as diferentes propriedades dos clados da OCP e entender sua evolução, em relação a funcionalidade, a OCP de *Synechocystis* sp. (PCC 6803) é muito estudada por Wilson *et al* (2008), onde observaram que nesta linhagem, durante períodos de iluminação, a forma ativa OCP^R acaba se concentrando em quantidades consideráveis, como consequência do aumento da intensidade da luz e/ou da redução de temperatura, sendo a fotoativação significativamente resistente a mudanças de temperatura. Na ausência de OCP, a extinção da fluorescência induzida pela forte luz branca ou azul-esverdeada é totalmente inibida e as células de *Synechocystis* são mais sensíveis à alta luz (WILSON *et al.*, 2006).

Estudos *in vitro* mostram que OCP1 de *Tolypothrix* é funcionalmente equivalente a OCP1 *Synechocystis*, incluindo sua regulação por FRP, que mostramos ser estruturalmente semelhante à forma dimérica da FRP de *Synechocystis*. *Tolypothrix* é um sistema modelo ideal para estudar a diversificação funcional de carotenoproteínas relacionadas ao OCP: seu genoma codifica membros dos clados OCP2 e OCP1, bem como um FRP. *Tolypothrix* OCP1 é 76% idêntico a *Synechocystis* OCP1 (BAO *et al.*, 2017).

Como visto no trabalho de Bao e colaboradores (2017), onde as sequências de aminoácidos de comprimento total da OCP foram recuperadas de genomas de cianobactérias disponíveis no banco de dados *The Integrated Microbial Genomes* (IMG), em fevereiro de 2016, e foram globalmente alinhadas para análise filogenética, a árvore baseada em máxima verossimilhança não enraizada resultante mostrou grande maioria das sequências agrupadas em um grande clado que inclui a OCP1 funcionalmente bem caracterizado (canônico) de *Synechocystis* (Slr1963) e os dois ortólogos estruturalmente caracterizados de *Arthrospira* e *Nostoc* (OCP_ARTMA e All3149, respectivamente). Portanto, esses dados corroboram a filogenia gerada nesse trabalho, pois há formação de três grandes clados, onde a OCP1, OCP2 E OCPX estão bem definidos, evidenciando assim, as semelhanças estrutural e funcional dessas OCPs.

4.2 MODELAGEM POR HOMOLOGIA

4.2.1 Seleção do molde

As seis sequências analisadas de OCP (OCP C3, OCP C14, OCP C70d, OCP C05, OCP C19 E OCP C22) foram submetidas por uma busca no PDB. O molde escolhido foi o *Tolypothrix* sp. PCC 7601 (ID 6PQ1), que apresentou maior identidade e similaridade com as sequências de OCP, ambas foram visualizadas no servidor ESPript 3.0 (Figura 13). O molde possui um peso total de 37 kDa, contagem de átomos de 2810, resolvido pelo método de difração de raio X, com resolução de 1,61 Å e uma estrutura cristalizada de ligante (ID 45d).



Figura 12 – Visualização das sequências no ESPript. Resíduos marcados na cor vermelha são conservados. Resíduos marcados com triângulo azuis são conservados e descritos em outros trabalhos.

Os grupos são alinhados usando consistência de perfil com estruturas secundárias previstas. Dessa maneira, foi possível observar que resíduos tais como Trp41, Tyr44, Trp110, Trp101, Tyr201, Tyr203 e Trp276 e Trp290, descritos em trabalhos anteriores (POLIVKA *et al.*, 2012; KIRILOVSKY & KERFELD, 2012), e outros próximos ao sitio ativo são bastante conservados.

4.2.2 Construção e validação dos modelos

Foi utilizado um script em linguagem Python, disponível no próprio Modeller versão 10.2 (FISER & SALI, 2003), e executado com objetivo de construir modelos que representasse cada OCP. A escolha do molde usado para construir os modelos, foi crucial para iniciar as análises, pois quanto maior a identidade das sequências com o molde melhor e mais significativa é a criação dos modelos pelo programa Modeller.

Dessa forma, os 100 modelos para cada uma das seis OCPs da coleção em estudo, foram analisados e com base no alinhamento da sequência alvo com a proteína molde, foi considerado o melhor modelo através da pontuação por função de densidade de probabilidade molecular (Molpdf) e DOPE *score* (Tabela 2).

Modelos das OCPs	Molpdf	DOPE score
OCP C70d	122.142.175	-3.869.178.125
OCP C14	120.196.265	-3.972.132.813
OCP C03	120.437.366	-3.855.217.188
OCP C19	129.998.108	-3.906.528.906
OCP C05	128.957.837	-3.813.422.266
OCP C22	120.262.646	-4.002.052.734

Tabela 2 – Seleção dos modelos das OCPs com base na pontuação Molpdf e DOPE score.

A estrutura 3D prevista das proteínas OCPs foram geradas pelo MODELLER e a estrutura com as pontuações de DOPE mais baixas foi selecionada. O melhor modelo foi selecionado escolhendo o modelo com menor valor de energia da função objetivo do Modeller (DOPE score) e maior valor de Molpdf score de uma coleção de modelos gerados pelo MODELLER.

Como mostrado na tabela acima, a pontuação Molpdf para os modelos, sofreu variação, indo de 120.142.175 a 128.998.108 e a DOPE, de -4.002.052.734 a - 3.813.422.266, menores e maiores valores respectivamente. Então, o Modeller em si, não calcula função de avaliação, e sim, de pontuação. A pontuação mostra como o modelo se comporta, diferente da avaliação, que mostra como o modelo é, no caso, a proteína normal e real.

Posteriormente, as estruturas 3D dos modelos das OCP foram alinhadas com o molde e visualizadas no programa PyMol 2.5.2 (SCHRÖDINGER, 2000), conforme a figura 14.



Figura 13 – Alinhamento dos modelos de OCP com o molde no programa PyMol. O molde estar representado pela cor vermelha e os modelos das OCPs estão explicitados conforme a ordem alfabética. A) OCP C70d em branco; B) OCP C14 em verde; C) OCP C03 em azul; D) OCP C19 em roxo; E) OCP C05 em amarelo; F) OCP C22 em rosa.

Todos os modelos gerados apresentaram um baixo RMSD, indicando que a conformação estrutural da proteína foi mantida, conforme Tabela 3.

Modelos OCP	RMSD
OCP C70d	0,135 Å
OCP C14	0,126 Å
OCP C03	0,109 Å
OCP C19	0,105 Å
OCP C05	0,127 Å
OCP C22	0,102 Å

Tabela 3 – Parâmetro RMSD de validação dos modelos de OCP das linhagens analisadas neste estudo.

Assim, quanto menor o RMSD, melhor é o modelo em comparação com a estrutura alvo. Os resultados apresentados na tabela 3 indicam que as estruturas se mantiveram estáveis, não apresentando valores de RMSD superior a 0,135 Å, sendo assim, os modelos assumiram a conformação mais próxima quando comparados ao molde, caso que foi possível observar na validação dos modelos de OCPs, com valores de mínima de 0,102 Å para o modelo de OCP C22 e máxima de 0,135 Å para o modelo de OCP C70d.

A OCP tem um domínio N-terminal a-helicoidal (resíduos 15-165) e um domínio C-terminal a/b (resíduos 190-317). No alinhamento feito entre as sequências mostrou que a sobreposição das proteínas pode haver vários resíduos conservados, resíduos esses, importantes para a atividade biológica da própria proteína, como Trp290, Tyr203 e Arg155 (CARBON *et al*, 2015; MELNICKI *et al*, 2016). Outro ponto importante é analisar o sitio-ativo da proteína, que após a modelagem e o alinhamento, se manteve visualmente semelhante a proteína-molde.

No servidor *online* SWISS-MODEL foi possível gerar os valores de QMEAN para cada OCP. Por padrão, eles são transformados em *Z-Score*¹ para relacioná-los com o que esperaríamos de estruturas de raios X de alta resolução (Figura 14).

¹*Z-Score* mede a distância entre um ponto de dados e a média usando desvios padrão. Z-scores podem ser positivos ou negativos. O sinal informa se a observação está acima ou abaixo da média. Por exemplo, um z-score de +2 indica que o ponto de dados está dois desvios padrão acima da média, enquanto um -2 significa que está dois desvios padrão abaixo da média. Um z-score de zero é igual à média.

Q	MEAN4 Value: -0.34 🚱	A)			
Sequence colored by local		/	Local quality score		
au	ality:			-	
-1	, -	0.0)	0.5	1.0
A:	MSFTIDSARNIFPGTLSADAVP.	AI	5 A R	FGQLSAEDQLALIWFAYL	45
A:	EMGKTITIAAPGAASMQFAEST	LA	Q V K	QMSYREQTQFMCDLANRA	90
A:	DTPLCRTYATWSPNIKLGFWYQ	LGI	RWM	EEGLVAPIP <mark>EGYRLSANA</mark>	135
A:	AAVLQSIRNIEPGQQITVLRNS	VVI	DМG	FDPAKLGSYNRVAEPVVA	180
A:	PKADDQRRKVTIEGVDNPTVLS	ΥMΙ	DNL	NANDFDALVQLFTPDGAI	225
A:	Q P P F Q R P I V G K E A V L R F F R E E C	QN	LNL	I P E R G V S E P A A D G Y T Q I K	270
A:	ITGKVQTPWFGGGVGMNIAWRF	LLI	N P E	NQIFFVAIDLLASP <mark>KELL</mark>	315
A:	NLIR				319



QMEAN4 Value: -0.35 @

C)

Sequence colored by local quality:



49

Continuação - Pontuação QMEAN das OCPs modelos, convertido para Z-Score.



Figura 14 – Pontuação QMEAN das OCPs modelos, convertido para Z-Score. A) OCP C70d em branco; B) OCP C14 em verde; C) OCP C03 em azul; D) OCP C19 em roxo; E) OCP C05 em amarelo; F) OCP C22 em rosa.

A pontuação QMEAN atribui um potencial de interação para o conjunto dos átomos e dependente da distância desses átomos, além de um potencial de ângulo de torção que alcança três resíduos posteriores ao resíduo em análise. Os termos individuais do QMEAN estão disponíveis na saída e sua análise pode revelar possíveis explicações para a baixa pontuação de um modelo, em contraste com as duas abordagens de aprendizado de máquina mencionadas acima que retornam uma única pontuação.

Quando comparado a outros *softwares* de avaliação de qualidade de modelo bastante utilizados, o QMEAN mostra uma melhoria significativa na estatística empregada em quase todas as medidas de qualidade que descrevem a capacidade da função de pontuação de identificar a estrutura nativa e discriminar modelos bons de maus (BERKENT *et al.*, 2008).

Em seguida, o servidor MolProbity foi utilizado com intuito de gerar os gráficos de Ramachandran para cada modelo em estudo e analisar suas qualidades estereoquímica, conforme a figura 16.



Figura 15 – Gráficos de Ramachandran gerados no MolProbity para cada modelo. A) OCP C70d; B) OCP C14; C) OCP C03; D) OCP C19; E) OCP C05; F) OCP C22.

A análise de Ramachandran atribui resíduos a uma das três categorias: Favorecido, Permitido e Atípico (RAMACHANDRAN *et al.*, 1963). Todos os resíduos, com exceção de um em OCP C70d (Asp129), um em OCP C14 (Asp168) e um em OCP C19 estão em áreas permitidas ou favorecidas. O mais importante, todavia, é a observação de que todas os resíduos do sítio ativo estão localizados em zonas favorecidas, o que confirma a estrutura tridimensional do modelo como confiável, o que aumenta consideravelmente a confiabilidade do modelo, e apenas um Ácido Aspártico em três dos modelos, de todos os aminoácidos, apareceu em uma área externa.

Todos os modelos apresentaram bons resultados para o gráfico de Ramachandran, com no mínimo 97,8% e o máximo de 99,4% de resíduos nas regiões favoráveis, e todos os modelos foram considerados aceitáveis (Tabela 4), onde é necessário que 90% dos resíduos, no mínimo, estejam nas regiões favoráveis para que a estrutura seja aceitável (LASKOWSKI, 1993).

Modelos OCP	Ramachandran
OCPC 70d	98,1%
OCP C14	98,4%
OCP C03	99,4%
OCP C19	97,8%
OCP C05	98,4%
OCP C22	98,4%
	1

Tabela 4 – Avaliação de validação dos modelos com valores acima de 90% dos resíduos em regiões favoráveis por Ramachandran.

Para avaliar a qualidade do enovelamento foi utilizado o servidor Verify3D. O Verify3D é um servidor online de uso livre e gratuito, sendo usado para verificar a avaliação da qualidade de modelos de proteínas com perfis tridimensionais (https://www.doembi.ucla.edu/verify3d/). Todos os arquivos estavam no formato de arquivo PDB e foram fornecidos como entrada para gerar um gráfico de janela de perfil (Tabela 5).

Modelos OCP	Pontuação Média 3D-1D>= 0,2
OCP C70d	94,98%
OCP C14	93,71%
OCP C03	93,75%
OCP C19	94,98%
OCP C05	95,58%
OCP C22	97,81%

Tabela 5 – Avaliação da qualidade do enovelamento no servidor Verify3D.

O servidor Verify3D fornece avaliação de estruturas no nível de resíduos, o que permite ao usuário localizar partes da proteína que provavelmente terão a conformação correta ou procurar regiões mal dobradas (GROTTHUSS *et al.*, 2003).

O Verify3D examina a semelhança de um modelo nuclear (3D) com a sua própria sequência de aminoácidos particular que é unidimensional. Cada depósito distribuía uma classe básica de radiância de sua área e condição (alfa, beta, circular, polar, apolar e assim por diante). A pontuação varia de -1 (pontuação ruim) a +1 (pontuação boa). Dessa forma, o menor valor foi de 93,71% para ocpC14 e o maior valor foi de 97,81% para OCP de OCP C22, as demais sequências ficaram dentro desse intervalo para o valor médio da pontuação 3D-1D >= 0,2. Vale ressaltar que, para uma avaliação ser considerada aceitável precisa conter no mínimo 80% da média de pontuação 3D-1D, e os modelos em estudo satisfazem essa condição.

4.3 DINÂMICA MOLECULAR

Simulações de DM de 100ns para cada modelo de OCP e também para o molde foram realizadas objetivando avaliar a estabilidade e afinidade dos complexos. Desse modo, mudanças nas conformações que puderam ser observadas nessas simulações foram importantes para aprimorar os testes de validação desses modelos, que foram construídos com base em dados cristalográficos obtidos do molde.

É comum que no começo das simulações os valores tendem a aumentar de forma abrupta, pois toda a estrutura busca um equilíbrio, de forma ao decorre do tempo atinge um pico, sugerindo que elas atingiram o equilíbrio. O RMSD apresenta valores que são retornados em função do tempo, onde permite enxergar qual o momento no tempo que as estruturas levaram para se estabilizar.

Os modelos de OCP de OCP C03 e OCP C05 apresentaram os maiores valores de RMSD e a estabilidade estrutural dos modelos com o molde são apresentadas na figura 17.



Figura 16 – Comparação entre gráficos RMSD em Angstrom dos modelos com o molde em simulações de DM. A) OCP C70d; B) OCP C14; C) OCP C03; D) OCP C19; E) OCP C05; F) OCP C22.

No trabalho de Bondanza e colaboradores (2020), foi usado a OCP de *Synechocystis* sp. PCC 6803 para simulações de DM, e notaram que os dois domínios NTD e CTD são bastante estáveis; isso também é confirmado pelo fato de que o RMSD desses domínios permanece em torno de 2 Å para toda a simulação e pela observação direta das estruturas extraídas da simulação. Afirmando o fato de que todos os complexos neste estudo permaneceram estáveis durante a simulação e o ligante (45D), apresentando valores de RMSD menores que 1 Å nos setes casos.

A avaliação funcional das proteínas foi estimada pela estabilidade do complexo e cálculos de energia livre de ligação. A afinidade do sítio de ligação e a contribuição individual dos resíduos foram estimadas através de cálculos de energia livre de ligação.

Segundo Su e colaboradores (2015), os resultados obtidos por MM/GBSA são sensíveis as variações no tempo de simulação. Os valores obtidos reforçam os dados descritos na literatura (HOU *et al.*, 2011), onde os métodos de GBSA não são acurados para predizer os valores experimentais aproximados. Porém, estes métodos descrevem valores relativos, apresentando a tendência entre um grupo de compostos.

Os últimos 5000 quadros (10 ns) de DM foram utilizados para calcular a energia livre de ligação pelo método MM-GBSA e MM-PBSA, conforme a tabela 6. O pacote de programas AMBER (Construção Assistida de Modelos com Refinamento de Energia - do inglês *Assisted Model Building with Energy Refinement*) (CASE *et al.*, 2012) é um conjunto de programas que reúne diversos campos de força de mecânica molecular, bastante conhecido e utilizado para realizar simulações de dinâmica molecular. É um pacote de aproximadamente 50 programas, no qual inclui programas licenciados e livres. Ele fornece vários protocolos para decompor as energias livres calculadas em contribuições específicas dos resíduos, usando modelos de solventes implícitos GB ou PB (METZ *et al.*, 2011).

MODEL OS	MM-	Std dov	Std orror	MM DDCA	Std. dev.	Std.
MODELOS	GBSA	Stu. uev.	Stu. error	WINI-F DSA		error
OCP C70d	-87,29	2,97	0,09	-79,61	3,47	0,10
OCP C14	-88,45	3,02	0,09	-76,52	3,70	0,11
OCP C03	-86,43	3,47	0,11	-75,83	3,78	0,11
OCP C19	-86,77	3,03	0,09	-75,78	3,53	0,11
OCP C05	-81,88	3,12	0,09	-71,67	3,68	0,11
OCP C22	-83,53	2,90	0,09	-74,43	3,38	0,10
<i>Tolypothrix</i> sp. PCC 7601	-88,81	3,17	0,10	-79,48	3,66	0,11

Tabela 6 – Resultados de cálculos de energia livre de ligação com base nos últimos 10 ns da simulação MD. Todos os valores de energia estão em kcal Mol^{-1} .

Com base nos resultados encontrados, pressupõem-se que as OCPs analisadas possuem alta afinidade ao ligante, com valores de energia livre de ligação abaixo de -80 Kcal/Mol⁻¹, esse fato é corroborado pelo elevado nível de conservação dos resíduos presentes e ao redor do sítio ativo.

Quando esses resultados são avaliados individualmente, é notório que OCP C22 e OCP C70d apresentaram maior afinidade ao ligante, valores esperados de estabilidade e energia de ligação. Desvio padrão de 2,90 e erro padrão 0,09 no cálculo de energia livre GB e desvio padrão de 3,38 e erro padrão de 0,10 no cálculo de energia livre PB para OCP C22. A OCP C70d apresentou desvio padrão de 2,97 e erro padrão de 0,09 no cálculo de energia livre GB e desvio padrão de 3,47 e erro padrão de 0,10 no cálculo de energia livre PB, sendo considerado o modelo com maior afinidade ao ligante.

A contribuição individual dos resíduos foi avaliada pela decomposição do método MM-GBSA e os resultados são apresentados na figura 18.



Figura 17 – Contribuição individual de energia por resíduos de acordo com o método MM-GBSA na complexação OCP-ligante. A) OCP C70d; B) OCP C14; C) OCP C03; D) OCP C19; E) OCP C05; F) OCP C22; G) *Tolypothrix* sp. PCC 7601. Fonte: Elaborado pelo autor (2022).

As interações podem ser decompostas para cada resíduo, incluindo apenas as interações nas quais um dos átomos do resíduo está envolvido, é a decomposição por resíduo. Alternativamente, as interações podem ser decompostas por pares de resíduos específicos, incluindo apenas as interações nas quais um átomo de cada um dos resíduos analisados está participando, é a chamada decomposição em pares. Esses protocolos de decomposição podem fornecer informações importantes nos cálculos de energia livres (GOHLKE *et al.*, 2003).

Alguns resíduos que fazem interações com o ligante, descritos em outros trabalhos, como os Tyr201, Trp288, Trp41, Tyr44 e Trp110, também foram observados nos modelos analisados aqui.

Bondanza *et al.* (2020), menciona que à medida que a translocação prossegue e o carotenoide se aprofundou no NTD, suas interações Leonard-Jones com este domínio continuam aumentando. Em particular, observamos que um grupo de resíduos tem um grande efeito de estabilização (Glu34, Leu37, Trp41, Thr80, Asn104, Trp110, Tyr111, Ile125, Pro126 e Tyr129). O conjunto desses resíduos define uma bolsa de ligação principalmente apolar/aromática para o carotenoide.

De fato, destaca-se o papel crucial desempenhado pelo Trp277, que estabiliza o complexo proteína-ligante via ligação H ao carotenoide, e por muitos resíduos aromáticos na cavidade do NTD (MAKSIMOV *et al.*, 2020).

Quanto a outras OCPs estudas, esses resíduos em todos os nossos modelos também fizeram interações favoráveis com a hECN, apesar de certos resíduos mostrarem maior interação em relação aos descritos anteriormente. Apesar das diferenças estruturais entre as seis OCPs analisadas nesse trabalho, foi possível observar que existe conservação estrutural de um grupo de resíduos. O trio Trp110, Arg155 e Leu107 foi visto em cinco dos seis modelos.

A dupla de resíduos Leu250 e Tyr44, fizeram-se presentes em 3 dos 6 modelos. Outra dupla de resíduos visto em 2 dos modelos foram Ile151 e Trp279. Vale ressaltar que esses resíduos supracitados, quando não presentes na mesma posição sofriam uma variação de até 5 resíduos da sua posição original, onde também fizeram interações favoráveis.

O sistema da OCP C05 não conservou o resíduo Ile151 e nem estava em posição próxima, apresentou os maiores resultados nos cálculos de energia livre de ligação.

Resíduos de Trp também foram conservados no sítio de ligação e apresentaram valores de energia relevantes na análise de decomposição.

Dessa forma, resíduos polares como Tyr, Trp e Ile, e resíduo básico como Arg em posições próximas ao sítio ativo, são capazes de fazer interações eletrostáticas favoráveis com hECN, além de colaborarem com o trio Trp110, Arg155 e Leu107 na complexação OCP-ligante. Trp110, Arg155 e Ile305, foram os resíduos que mais contribuíram em OCP C70d. Em contrapartida, os resíduos Ile305, Ile151 e Trp290 contribuíram de forma significativa em OCP C22.

O carotenoide da OCP ocupa duas posições distintas na proteína dependendo do estado de fotoativação: repouso (OCP^O) e ativado (OCP^R). Os resíduos que definem essas duas posições distintas são referidos como configurações carotenoide-proteína cpcO e cpcR, respectivamente, e esses resíduos são altamente conservados dentro da família de OCP (KERFELD, 2017).

No trabalho de Miksimov (2015), confirma o supracitado, através da destruição da ponte salina Arg155-Glu246 e outras alterações fotoinduzidas que levam a movimentos específicos dos domínios N e C da OCP e a um aumento na flexibilidade do OCP. As mudanças cíclicas na hidrofobicidade e geometria geral da OCP parecem ser cruciais para fornecer um microambiente em torno do cofator hECN, que é necessário para a implementação do NPQ de fluorescência de ficobilissomas.

As distâncias entre as pontes Arg155-Glu246, foram medidas e analisadas, conforme a tabela 7.

Modelos	Distâncias Inicial	Distância Final
OCP C14	22.5 Å	11.3 Å
OCP C70d	5.5 Å	4.8 Å
OCP C03	4.5 Å	3.7 Å
OCP C19	7.1 Å	3.4 Å
OCP C05	11.9 Å	5.1 Å
OCP C22	9.5 Å	6.1 Å

Tabela 7 – Distâncias inicias e finais da ponte Arg155-Glu246 medidas no PyMol

Pode-se perceber claramente que a ponte salina Arg155-Glu246 contribui para a estabilidade do complexo proteína-ligante, fazendo com que haja um encurtamento da

distância, trazendo mais rigidez e mantendo o complexo fechado ao longo das simulações de DM.

Nossa compreensão da modularidade estrutural do OCP, sua capacidade de mudar de cor em resposta a gatilhos ambientais e sua dissociação de domínio ativado por luz estabelecem as bases para seu reaproveitamento em biotecnologia.

Aplicações mais próximas da função nativa foram projetadas; o OCP foi introduzido em um modelo de antena de captação de luz artificial para controlar o fluxo de energia em função da excitação da intensidade da luz. A capacidade de tirar proveito da dissociação/reassociação de domínio controlado por luz do OCP é baseada na capacidade de projetar as afinidades relativas do NTD e CTD um para o outro (ANDREONI *et al.*, 2017). A introdução de mutações na interface principal ou secundária pode ser usada para ajustar a afinidade relativa do NTD e CTD entre si sob diferentes condições.

Além disso, o OCP é conhecido por se converter na forma espectral vermelha em pH baixo (pH = 3,5) ou na presença de caotrópicos. Isso sugere que o OCP poderia ter usos potenciais como um indicador de mudança de pH, presença de caotrópicos ou mesmo como um sensor geral para o estado de dobramento de proteínas (KING *et al.*, 2014).

A ligação específica do carotenoide também influencia a fotoativação; por exemplo, sabe-se que a ligação de 3'- hidroxiequinenona aumenta o rendimento quântico de OCP da fotoativação. O OCP, ao adicionar carotenóides aos cromóforos atualmente disponíveis (FMN, FAD, retinal, biliverdina, ficocianobilina, fitocromobilina) usados em biologia sintética, amplia a faixa espectral de luz que pode ser usada em aplicações optogenéticas (ZIEGLER & MOGLICH, 2015).

Os organismos vivos precisam sentir constantemente as mudanças ambientais e se adaptar a elas para sobreviver. Para organismos fotossintéticos, é crucial detectar mudanças na intensidade da luz e responder rapidamente a elas. Esses organismos são capazes de modificar seu aparato fotossintético para coletar luz com eficiência sob condições de luz flutuantes, que mudam rapidamente de intensidades de luz subsaturadas para supersaturadas. Ao mesmo tempo, os organismos fotossintéticos devem evitar a foto dano, gerado pelo excesso de energia que atinge os centros de reações fotoquímicas (MUZZOPAPA, 2020).

5. CONCLUSÃO

✓ No presente estudo, foi possível observar que as OCPs (OCP C03, OCP C14, OCP C70d, OCP C05, OCP C19 e OCP C22) coletadas das cianobactérias isoladas de ambientes amazônicos se agruparam com outras sequências de OCPs obtidas em banco de dados públicos e globais. Contudo, houve a formação de um agrupamento dessas OCPs com outras do tipo OCP1 em um clado específico, podendo-se inferir que suas funcionalidades são semelhantes por conta da estrutura conservadas das OCPs. OCP2 e OCPX agruparam mais próximas entre si, em clados distintos, confirmando trabalhos anteriores.

✓ Quando avaliados os modelos construídos a partir do molde selecionado de *Tolypothrix* sp. PCC 7601, todos foram validados com sucesso, com ótimos valores de RMSD, Ramachandran, QMEAN, DOPE *score* e Verify3D, evidenciando que a proteína em si é bastante conservada, principalmente quando analisadas as estruturadas tridimensionais de cada modelo.

✓ Ao utilizar de simulações de DM para analisar a interação da proteína com o ligante, observou-se que houve pouca variação da distância média entre cada átomos nos modelos comparados (RMSD). O trio Trp110, Arg155 e Leu107 parecem estar intimamente ligados ao processo de interação com o ligante hECN no sítio ativo, sendo conservados na maioria dos modelos. Além disso, o estudo detalhado da conservação de resíduos e interações de ligação ajuda a selecionar os melhores candidatos para aplicações futuras.

✓ Por fim, a busca por novas formas de OCPs obtidas de outras cianobactérias podem revelar novas aplicações para esta proteína, incluindo a utilização de organismos geneticamente modificados (OGMs) com maior capacidade desenvolvimento sob condições de estresse, indo desde cianobactérias capazes de produzir concentrações maiores de carotenoides, até plantas com maior capacidade de evitar danos causados pela intensidade luminoso e atividade fotossintética.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- AVER, G. M.; KREUTZ, O. C.; SUYENAGA, E. S. METHODS OF OBTAINING DRUGS FROM THE PERSPECTIVE OF MEDICINAL CHEMISTRY. Revista Conhecimento Online, Novo Hamburgo, a. 7, v. 2, 2015.
- ABALDE, J.; BETANCOURT, L.; TORRES, E.; CID, A.; BARWELL, C. Purification and characterization of phycocyanin from the marine cyanobacterium *Synechococcus* sp. IO9201. Plant Science, v. 136, n. 1, p. 109-120, 1998.
- ABAGYAN, R.; KUFEREVA, R. Methods of protein structure comparison. Methods Mol Biol. v. 857, p. 231–257, 2012.
- ANDREONI, A.; LIN, S.; LIU, H.; BLANKENSHIP, R. YAN, H.; WOODBURY, N. Orange carotenoid protein as a control element in an antenna system based on a DNA nanostructure. Nano Lett, v. 17, p. 1174-1180, 2017.
- BAO, H. MELNICKI, M.; PAWLOWSKI, E.; SUTTER, M.; AGOSTONI, M; LECHNO-YOSSEF, S.; CAI, F.; MONTGOMERY, B. KERFELD, C. Additional families of orange carotenoid proteins in the photoprotective system of cyanobacteria. Nature Plants, v. 3, n. 8, p. 1-11, 2017.
- BENKERT, P.; BIASINI, M.; SCHWEDE, T. Toward the estimating the absolute quality of individual protein structure models. **Bioinformatic**, v. 27, p. 343-350, 2011.
- BERERA, R.; STOKKUM, I. H. V.; GWIZDALA, M.; WILSON, A.; KIRILOVSKY, D; GRONDELLE, R. V. The photophysics of the orange carotenoid protein, a lightpowered molecular switch. J. Phys. Chem. B, v. 116, n. 8, p. 2568–2574, 2012.
- BIASINI, M. BIENERT, S.; WATERHOUSE, A.; ARNOLD, K.; STUDER, G.; SCHMIDT, T.; KIEFER, F.; CASSARINO, T. G.; BERTONI, M.; BORDOLI, L.; SCHWEDE, T. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Research, v. 42, n. 1, p. 252-258, 2014.

- SANTOS, B. F.; DU, W.; HELLINGWERF, K. J. Synechocystis: not just a plug-bug for CO2, but a green E. coli. Frontiers in Bioengineering and Biotechnology, v. 2, n. 36, 2014.
- BONDANZA, M.; CUPELLINI, L.; FACCIOLI, P.; MENNUCCI, B. Molecular Mechanisms of Activation in the Orange Carotenoid Protein Revealed by Molecular Dynamics. J. Am. Chem. Soc. v. 142, p. 21829-21841, 2020.
- BOULAY, C.; WILSON, A.; D'HAENE, S.; KIRILOVSKI, D. Identification of a protein required for recovery of full antenna capacity in OCP-related photoprotective mechanism in cyanobacteria. Proceedings Of The National Academy Of Sciences, v. 107, n. 25, p. 11620-11625, 2010.
- BOULAY, C.; ABASOVA, L.; SIX, C.; VASS, I.; KIRILOVSKI, D. Occurrence and function of the orange carotenoid protein in photoprotective mechanisms in various cyanobacteria. Biochimica Et Biophysica Acta (bba) - Bioenergetics, v. 1777, n. 10, p. 1344-1354, 2008.
- BRAUN, R. L.; JUNQUEIRA, D. M.; VERLI, H. Filogenia Molecular. *In*: VERLI, H. (Org.). Bioinformática: da biologia à flexibilidade molecular. 1. ed. Porto Alegre. p. 81-114, 2014.
- CARBON, C. B.; THUROTTE, A.; WILSON, A.; PERREAU, F.; KIRILOVSKI, D. Biosynthesis of soluble carotenoid holoproteins in Escherichia coli. Scientific Reports, v. 5, n. 1, p.1-8. 2015.
- CASE, D. A.; AKTULGA, H. M.; BELFON, K.; BEN-SHALOM, I. Y.; BROZELL, S. R.; CERUTTI, D. S.; KOLLMAN, P. A. Amber 2021. University of California, San Francisco, 2021.

CASE, D. A.; SALOMON-FERRER, R.; WALKER; R. C. Amber 12. p. 1–826, 2012.

- CASTRO, F. C. EXTRAÇÃO SÓLIDO-LÍQUIDO DE FICOCIANINA DE Arthospira platensis USANDO LÍQUIDO IÔNICO PRÓTICO (2-HEAF): MODELAGEM DOS DADOS EXPERIMENTAIS. Monografia (Engenharia Química) - Centro de Tecnologia da Universidade Federal do Ceará, fortaleza, 2018.
- CHENNA, R.; SUGAWARA, H.; KOIKE, T.; LOPEZ, R.; GIBSON, T. J.; HIGGINS, D. G.; THOMPSON, J. D. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Research, v. 31, n. 13, p. 3497-3500, 2003.
- CHEN, V. B. ARENDALL, W. B.; HEADD, J. J.; KEEDY, D. A.; IMMORMINO, R. M.; KAPRAL, G. J.; MURRAY, L. W.; RICHARDSON, J. S.; RICHARDSON, D. C. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallographica Section D Biol Crystallogr, v. 66, n. 1, p. 12-21, 2010.
- CHRISTAKI, E.; FLOROU-PANERI, P.; BONOS, E. Microalgae: a novel ingredient in nutrition. International Journal of Food Sciences and Nutrition, v. 62, p. 794–799, 2011.
- COHEN, Y.; GUREVITZ, M. The Cyanobacteria-Ecology, Physiology and Molecular Genetics. **The Prokaryotes**, p. 1074-1098, 2006.
- CORNET, L.; BERTRAND, A. R.; HANIKENNE, M.; JAVAUX, E. J.; WILMOTTE, A.; BAURAIN, D. Metagenomic assembly of new (Sub)polar Cyanobacteria and their associated microbiome from non-axenic cultures. **Microbial Genomics**, v. 4, 2018.
- COSTA, V. B.; GOMES, A. L.; OLIVEIRA, G. J. PINHEIRO, S. C. C.; CUNHA C. J.
 S.; SOUSA, E. B. CIANOBACTÉRIAS DO PARQUE ESTADUAL DO CHARAPUCU (AFUÁ, PARÁ, BRASIL). *In*: ALFARO, A. T. S.; TROJAN, D. G. (Org.). Ciências Ambientais e o Desenvolvimento Sustentável na Amazônia. Editora Atena, 2017. 183 p.
- DEXTRO, R. B. DELBAJE, E.; COTTA, S. R.; ZEHR, J. P. FIORE, M. F. Trends in Free-access Genomic Data Accelerate Advances in Cyanobacteria Taxonomy. Journal of Phycology, v. 57, p. 1392–402, 2021.

- DITTMANN, E. GUGGER, M.; SIVONEN, K.; MENOS, D. P. Natural Product Biosynthetic Diversity and Comparative Genomics of the Cyanobacteria. Trends In Microbiology, v. 23, n. 10, p. 642-652, 2015.
- EISENBERG, D.; LUTHY, R.; BOWIE, J. U. VERIFY3D: assessment of protein models with three-dimensional profiles. **Methods in Enzimology**, v. 277, p. 396-404, 1997.
- FERREIRA, F. J. N. Investigação computacional de bromo-ariloxi-2-acetamida-etilbenzimidazólicos como inibidores não-peptídicos da proteinase cruzaína de Trypanosoma cruzi. 2017. Dissertação (Mestrado em Química Medicinal e Modelagem Molecular da Universidade Federal do Pará), 2017.
- FISER, A.; SALI, A. MODELLER: Generation and Refinement of Homology-Based Protein Structure Models. **Methods in Enzimology**, v. 374, p. 460-491, 2003.
- GANTT, E. Phycobilisomes. Annu Rev Physiol. v. 23, p. 327-347, 1981.
- GENHEDEN, S. KONGSTED, J.; SODERHJELM, P.; RYDE, U. Nonpolar solvation free energies of protein-ligand complexes. J. Chem. Theory Comput. v. 6, n. 11, p. 3558–3568, 2010.
- GLAZER, A. N. Phycobiliproteins a Family of Valuable, Widely Used Fluorophores. **Journal of Applied Phycology**, v. 6, n. 2, p. 105-112, 1994.
- GOHLKE, H. KIEL, C.; CASE, D. A. Insights into protein–protein binding by binding free energy calculation and free energy decomposition for the ras–raf and ras–ralgds complexes. Journal of molecular biology, v. 330, n. 4, p. 891–913, 2003.

GRAHAM, L. E.; WILCOX, L. W. Algae. Prentice Hall, New Jersey, p. 97-131, 2000.

- GROSSMAN, A. R.; SCHAEFFER, M. R.; CHIANG, G. G.; COLIER, J. L. The phycobilisone, a light-harvesting complex responsive to environmental conditions. Microbiol. Rev., v. 57, p. 725-749, 1993.
- GROTTHUSS, M. V. PAS, J. WYRWICZ, L.; GINALSKI, K.; RYCHLEWSKI, L.; Application of 3D-Jury, GRDB, and Verify3D in Fold Recognition. **PROTEINS: Structure, Function, and Genetics**, v. 53, p. 418 – 423, 2003.

- GUPTA, S. SUTTER, M.; REMESH, S. G.; DOMINGUEZ-MARTIN, M. A.; BAO, H.; FENG, X. A.; CHAN, L. G.; PETZOLD, C. J.; KERFELD, C. A.; RALSTON, C. Y. X-ray radiolytic labeling reveals the molecular basis of orange carotenoid protein photoprotection and its interactions with fluorescence recovery protein. J. Biol. Chem., v. 294, p. 8848–8860, 2019.
- HOLT, T.; KROGMANN, D. W. A carotenoid-protein from cyanobacteria. Biochim.Biophys. Acta, v. 637, p. 408–414, 1981.
- HAUER, T; KOMÁREK, J. CyanoDB 2.0 On-line database of Cyanobacterial genera.World-wide electronic publication, Univ. of South Bohemia & Inst. of Botany AS CR., 2021.
- HOU, T. WANG, J.; LI, Y.; WANG, W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. Journal of Chemical Information and Modeling, v. 51, p. 69–82, 2011.
- JOHNSON, E. M.; KUMAR, K. D. Physicochemical parameters optimization, and purification of phycobiliproteins from the isolated *Nostoc* sp. Bioresource Technology, v. 166, p. 541-547, 2014.
- JOSHI, H.; BHATIA, D.; KRISHNAN, Y.; MAITI, P. K. Probing the structure and in silico stability of cargo loaded. **Nanoscale**, v. 9, p. 4467-4477, 2017.
- KERFELD, C. A. Structure and function of the water-soluble carotenoid-binding proteins of cyanobacteria. **Photosynth. Res.** v. 81, p. 215-225, 2004.
- KERFELD, C. A.; SAWAYA M. R.; BRAHMANDAM, V.; CASCIO, D.; HO, K. K.; TREVITHICK-SUTTON C. C.; KROGMANN; D. W.; YEATES, T. O. The crystal structure of a cyanobacterial water-soluble carotenoid binding protein. Structure, v. 11, p. 55–65, 2003.
- KING J. D. LIU, H.; HE, G.; ORF, G. S.; BLANKENSHIP, R. E. Chemical activation of the cyanobacterial orange carotenoid protein. **FEBS Lett**, v. 588, p. 4561-4565, 2014.

- KIRILOVSKY, D.; KERFELD, C. A. Cyanobacterial photoprotection by the orange carotenoid protein. **Nature Plants**, v. 2, n. 12, p. 16180-16186, 2016.
- KIRILOVSKY, D.; KERFELD, C. A. The Orange Carotenoid Protein: a blue-green light photoactive protein. **Photochem Photobiol Sci**, v. 12, p. 1135-1143, 2013.
- KIRILOVSKY, D.; KERFELD, C. A. The Orange Carotenoid Protein: a blue-green light photoactive protein. Photochemical & Photobiological Sciences, v. 12, n. 7, p. 1135-1143, 2012.
- KOMÁREK, J. Cyanobacterial taxonomy: current problems and prospects for the integration of traditional and molecular appoaches. **Algae**, v. 21, p. 349-375, 2006.
- KOMÁREK, J.; ANAGNOSTIDIS, K. Cyanoprokaryota 1. Sussawasserflora von Mitteleuropa, p. 548, 1998.
- KONOLD, P. E.; STOKKUM, I. H. M. V.; MUZOPAPPA, F.; WILSON, A.; GROOT, M.; KIRILOVSKY, D.; KENNIS, J. T. M. Photoactivation Mechanism, Timing of Protein Secondary Structure Dynamics and Carotenoid Translocation in the Orange Carotenoid Protein. J. Am. Chem. Soc., v. 141, p. 520–530, 2019.
- KUMARI, R.; KUMAR, R.; LYNN, A. A. G_mmpbsa-A GROMACS Tool for High-Throughput MMPBSA Calculations. Journal of Chemical Information and Modeling. 2015.
- KUZNETSOVA, V.; DOMINGUEZ-MARTIN, M. A.; BAO, H.; GUPTA, S.; SUTTER,
 M.; KLOZ, M.; REBARZ, M.; PRECEK, M.; CHEN, Y.; PETZOLD, C. J.;
 RALSTON, C. Y.; KERFELD, C. A.; POLIVKA, T. Comparative ultrafast spectroscopy and structural analysis of OCP1 and OCP2 from Tolypothrix.
 Biochimica et Biophysica Acta (BBA) Bioenergetics, v. 1861, n. 2, 2020.
- LECHNO-YOSSEF, S.; MELNICKI, M. R.; BAO, H. MONTGOMERY, B. L.; KERFELD, C. A. Synthetic OCP heterodimers are photoactive and recapitulate the fusion of two primitive carotenoproteins in the evolution of cyanobacterial photoprotection. **Plant J**, v. 91, p. 646–656, 2017.

- LEVERENZ, R. L. SUTTER, M.; WILSON, A. GRUPTA, S.; THUROTTE, A.; CARBON, C. B.; PETZOLD, C. J. PHOTOSYNTHESIS. A 12 A carotenoid translocation in a photoswitch associated with cyanobacterial photoprotection. Scienc, v. 348, n. 6242, p. 1463–1466, 2015.
- LINDBERG, P.; PARK, S.; MELIS, A. Engineering a platform for photosynthetic isoprene production in cyanobacteria, using *Synechocystis* as the model organism. Metabolic Engineering, v. 12, n. 1, p.70-79, 2010.
- LOPEZ-IGUAL, R.; WILSON, A.; LEVERENZ, R. L.; MELNICKI, M. R.; CARBON, C. B.; SUTTER, M.; TURMO, A.; PERREAU, F.; KERFELD, C. A.; KIRILOVSKI, D. Different Functions of the Paralogs to the N-Terminal Domain of the Orange Carotenoid Protein in the Cyanobacterium *Anabaena* sp. PCC 7120. Plant Physiol, v. 171, p. 1852-1866, 2016.
- LOVELL, S. C.; DAVIS, Y. W.; ARENDALL, W. B.; BAKKER, P. I. W.; WORD, J. M.; PRISANT, M. G.; RICHARDSON, J. S.; RICHARDSON, D. C. Structure validation by Calpha geometry: phi, psi and Cbeta Deviation. **Proteins**. v. 50, p. 437– 450, 2003.
- LOU, W.; NEIDZWIEDZKI, D. M.; JIANG, R. J.; BLANKENSHIP, R. E.; LIU, H. Binding of red form of Orange Carotenoid Protein (OCP) to phycobilisome is not sufficient for quenching. **BBA – Bioenergetics**, 2020.
- MAKSIMOV, E. G.; PROTASOVA, E. A.; TSORAEV, G. V.; YAROSHEVICH, G. V.;
 MAYDYKOVSKIY, A. I.; SHIRSHIN, E. A.; GOSTEV, T. S.; JELZOW, A.;
 MOLDENHAUER, M.; SLONIMSKIY, Y. B.; SLUCHANKO, N. N.; THOMAS
 FRIEDRICH. Probing of carotenoid-tryptophan hydrogen bonding dynamics in the single-tryptophan photoactive Orange Carotenoid Protein. Sci. Rep., v. 10, p. 11729, 2020.
- MARTINEZ, L. Automatic Identification of Mobile and Rigid Substructures in Molecular Dynamics Simulations and Fractional Structural Fluctuation Analysis. PLOS ONE, v. 10, 2015.

- MARTINS, F. N. Uma síntese sobre aspectos da fotossíntese. **Revista de Biologia e Ciências da Terra**, v. 11, p. 10-14, 2011.
- MAIER, J. A.; MARTINEZ, C.; KASAVAJHALA, K.; WICKSTROM, L.; HAUSER, K. E.; SIMMERLING, C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. Journal of chemical theory and computation. v. 11, p. 3696-3713, 2015.
- MEZZETTI, A.; ALEXANDRE, M.; THUROTTE, A.; WILSON, A.; GWIZDALA, M.; KIRILOVSKI, D. Two-Step Structural Changes in Orange Carotenoid Protein Photoactivation Revealed by Time-Resolved Fourier Transform Infrared Spectroscopy. J. Phys. Chem. B, v. 123, p. 3259–3266, 2019.
- MCCORMICK, A. J. BOMBELLI, P.; BRADLEY, R. W.; THORNE, R.; WENZEL, T.; HOWE, C. J. Biophotovoltaics: oxygenic photosynthetic organisms in the world of bioelectrochemical systems. Energy & Environmental Science, v. 8, p. 1092-1109, 2015.
- MELNICKI, M. R. LEVERENZ, R. L.; SUTTER, M.; LOPEZ-IGUAL, R.; WILSON, A.; PAWLOWVSKI, E. G.; PERREAU, F.; KIRILOVSKI, D.; KERFELD, C. A. Structure, Diversity, and Evolution of a New Family of Soluble Carotenoid-Binding Proteins in Cyanobacteria. Molecular Plant, v. 9, n. 10, p. 1379-1394, 2016.
- METZ, A.; PFLEGER, C.; KOPITZ, H.; PFEIFFER-MAREK, S.; BARINGHAUS, K.; GOLHKE, H. Hot spots and transient pockets: predicting the determinants of smallmolecule binding to a protein–protein interface. Journal of chemical information and modeling, v. 52, n. 1, p. 120–133, 2011.
- MOLDEN-HAUER, M.; SLUCHANKO. N. N.; BUHRKE, D.; ZLENKO, D. V.; TAVRAZ, N. N.; SCHMITT F. J.; HILDEBRANDT, P.; MAKSIMOV E. G.; FRIEDRICH, T. Assembly of photoactive orange carotenoid protein from its domains unravels a carotenoid shuttle mechanism. **Photosynth Res**, v. 133, p. 327–341, 2017.
- MORAES, C. C.; FIGUEIRA, F. S.; KALIL, S. J. C-phycocyanin extraction process for large-scale use. Journal of Food Biochemistry, v. 34, p. 133-148, 2010.

- MUR, L. R.; OLAV, M. S.; UTKILEN, H. Toxic Cyanobacteria in Water: A guide to their public health consequences, monitoring and management. WHO, 1999.
- MUZZOPPAPA, F.; WILSON, A.; KIRILOVSKY, D. Interdomain interactions reveal the molecular evolution of the orange carotenoid protein. **Nature Plants**, v. 5, p. 1076– 1086, 2019.
- NEI, M.; KUMAR, S. Molecular Evolution and Phylogenetics. **Oxford University Press**, Oxford, 2000.
- PAGELS, F.; BONOMI, B, J.; VEGA, J.; ABDALA-DÍAZ, R.; VASCONCELOS, V.; GUEDES, A. C.; FIGUEROA, F. L.; Light regulating metabolic responses of Cyanobium sp. (Cyanobacteria). Arch Hydrobiol, v. 193, p. 285–297, 2020.
- PAGELS, A. F.; GUEDES, A. C.; AMARO, H. M.; KIJJOA, A.; VASCONCELOS, V.
 Phycobiliproteins from cyanobacteria: Chemistry and biotechnological applications.
 Biotechnology Advances, v. 37, n. 3, 2019.
- PATTANAIK, B.; LINDBERG, P. Terpenoids and Their Biosynthesis in Cyanobacteria. Life, v. 5, n. 1, p.269-293, 2015.
- PEREIRA, S. L. Biologia Molecular e Evolução. Holos editora, p. 123-131, 2012.
- POLÍVKA, T.; CHÁBERA, P.; KERFELD, C. A. Carotenoid–protein interaction alters the S1 energy of hydroxyechinenone in the Orange Carotenoid Protein. Biochimica Et Biophysica Acta (bba) Bioenergetics, v. 1827, n. 3, p. 248-254, 2012.
- PRADO, M. A; GODOY, H. T. Corantes artificiais em alimentos. Alimentos e Nutrição, Araraquara, v.14, n. 2, p. 237-250, 2003.
- PRASANNA, R.; SOO, A.; JAISWAL, P. NAYAK, S.; GUPTA, V.; CHAUDHARY, V.; JOSHI, M.; NATARAJAN, C. Rediscovering cyanobacteria as valuable sources of bioactive compounds. Applied Biochemistry and Microbiology, v. 46, n. 2, p. 119-134, 2010.
- RAMAKRISHNAN, C. RAMACHANDRAN, G. N. Stereochemical criteria for polypeptide and protein chain conformations: II. Allowed conformations for a pair of peptide units, **Biophys. J**. v. 5, p. 909–933, 1965.
- RAMACHANDRAN, G. N. Estereoquímica das configurações da cadeia polipeptídica. Jornal de Biologia Molecular, v. 7. p. 95-99, 1963.
- RAPOSO, M. F. J.; MORAIS, R. M. S. C.; MORAIS, A. M. M. B. Health applications of bioactive compounds from marine microalgae. Life Sciences, v. 93, p. 479–486, 2013.
- RASTOGI, R. P.; SONANI, R. R. PATEL, A. B.; MADAMWAR, D. Occurrence of a functionally stable photoharvesting single peptide allophycocyanin α-subunit (16.4 kDa) in the cyanobacterium *Nostoc* sp. R76DM. **RSC Advances**, v. 5, n. 106, p. 87598-87608, 2015.
- RAVIERS, B. **Biologia e filogenia das algas**. 1. ed. Artmed; Porto Alegre: Brasil, p. 280, 2006.
- RIPPKA, R.; DERUELLES, J.; WATERBURY, J. B.; HERDMAN, M.; STANIER, R.,
 Y. Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. Microbiology, v. 111, n. 1, p.1-61, 1979.
- RIZZO, R. F.; SANTOS, B. N, C.; CASTRO, G. F. P. S.; PASSOS, T. S.; LIMA-ARAUJO, K. G. Production of phycobiliproteins by Arthrospira platensis under different lightconditions for application in food products. Food Science and Technology, v. 35, n. 2, p. 247-252, 2015.
- SÁ, L. L. C.; VIERA, J. M. S.; MENDES, R. A.; PINHEIRO, S. C. C.; VALE, E. R.; ALVES, F. A. S.; JESUS, I. M.; SANTOS, E. C. O.; COSTA, V. B. Ocorrência de uma floração de cianobactérias tóxicas na margem direita do rio Tapajós, no Município de Santarém (Pará, Brasil). **Rev. Pan-Amaz Saude**, v. 1, p. 159-166, 2010.
- SEDOUD A.; LOPEZ-IGUAL, R.; REHMAN, U.; WILSON, A. PERREAU, F.; VASS, I. The Cyanobacterial Photoactive Orange Carotenoid Protein Is an Excellent Singlet Oxygen Quencher. **Plant Cell**, v. 26, p. 1781-1791, 2014.
- SCHOPF, J. W. The Fossil Record: Tracing the toots of the Cyanobacterial Lineage. *In*:
 B. A. WHITTON; M. POTTS (Eds.). The Ecology of Cyanobacteria: Their Diversity in Time and Space. Kluwer Academic Publisher, Drodrecht, p. 13-35, 2000.

- SANT'ANNA, C. L.; CELIA, L.; AZEVEDO, M. T. P.; AGUJARO, L. F. CARVALHO, L. R. Manual Ilustrado para Identificação e Contagem de Cianobactérias Planctônicas de Águas Continentais Brasileiras. Rio de Janeiro: Interciência, Sociedade Brasileira de FicologiaSBFic, p. 58, 2006.
- SHCHERBAKOVA, D. M.; SHEMETOV, A. A.; KABERNIUK, A. A.; VERKHUSHA,
 V. V. Natural Photoreceptors as a Source of Fluorescent Proteins, Biosensors, and
 Optogenetic Tools. Annual Review of Biochemistry, v. 84, p. 519-550, 2015.

SHEN, M.-Y.; ŠALI, A. Statistical potential for assessment and prediction of **protein** structures. **Proteína Sci**. v. 15, p. 2507-2524, 2006.

- SILVA, L. A. Processo biotecnológico de produção, extração e recuperação do pigmento ficocianina da Spirula platensis. Dissertação (Mestrado em Processos Biotecnológicos) – Universidade Federal do Paraná – UFPR, Curitiba, 2008.
- SINGH, R. K.; TIWARI, S. P.; RAI, A. K.; MOHAPATRA, T. M. Cyanobacteria: an emerging source for drug discovery. The Journal of Antibiotics, v. 64, n. 6, p.401-412, 2011.
- SIQUEIRA, A.; LIMA, A. R. J.; AGUIAR, D. C. F.; SANTOS, A. S.; JÚNIOR, J. L. S. G. V.; GONÇALVES, E. C. Genomic screening of new putative antiviral lectins from Amazonian cyanobacteria based on a bioinformatics approach. **Proteins,** v. 86, p. 1047–1054, 2018.
- STANIER, R. Y.; SISTROM, W. R.; HANSEN, T. A.; WHITTON, B. A.; CASTENHOLZ, R. W.; PFENNIG, N.; GORLENKO, V. N.; KONDRATIEVA, E. N. EIMHJELLEN, K. E.; WHITTEMBURY, R.; GHERNA, R. L.; TRUPER, H. G. Proposal to place the nomenclature of the cyanobacteria (blue-green algae) under the rules of the International Code of Nomenclature of Bacteria. International Journal of Systematic Bacteriology, v. 28, p. 335–336, 1978.
- SU, P. C.; TSAI, C. C.; MEHBOOB, S. HEVENER, K. E.; JOHNSON, M. E.; Comparison of radii sets, entropy, QM methods, and sampling on MM-PBSA, MM-GBSA, and QM/MM-GBSA ligand binding energies of F. tularensis enoyl-ACP reductase (FabI). Journal of Computational Chemistry, 2015.

- TAMURA, K.; KUMAR, S.; STECHER, G.; LI, M.; KNYAZ, C. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular Biology and Evolution, v. 35, p. 1547-1549, 2018.
- THAJUDDIN, N.; SUBRAMANIAN, G. Cyanobacterial biodiversity and potential applications in biotechnology. **Current Science**, v. 89, p. 47–57, 2005.
- TRIPP, H. J.; BANCO, S. R.; TURK, K. A.; FOSTER, R. A.; DESANY, NIAZI, FAHEEM, JASON, P. A.; ZEHR, J. P. Metabolic streamlining in an open-ocean nitrogen-fixing cyanobacterium. Nature, v. 464, n. 7285, p. 90–94, 2010.
- VERLI, H. Dinâmica Molecular. In: VERLI, H. (Org). Bioinformática: da Biologia à Flexibilidade Molecular. 1. ed. São Paulo: SBBq, p. 173-187, 2014.
- VOET, D.; VOETOVÁ, J. Biochemie. Victoria Publishing, Praha, 1995.
- YADAV, S.; SINHA, R. P.; TYAGI, M. B.; KUMAR, A. Cyanobacterial secondary metabolites. International Journal of Pharma and Bio Sciences, v. 2, n. 2, p. 144-167, 2011.
- WALTER, J. M.; COUTINHO, F. H.; DUTILH, B. E.; SWINGS, J.; THOMPSON, F. L.; THOMPSON, C. C. Ecogenomics and Taxonomy of Cyanobacteria Phylum. Front. Microbiol, v. 8, 2017.
- WHELAN S.; GOLDMAN, N. A. General Empirical Model of Protein Evolution Derived from Multiple Protein Families Using a Maximum-Likelihood Approach. Molecular Biology and Evolution, v. 18, n. 5, p. 691–699, 2001.
- WILSON A.; PUNGINELLI, C.; GALL, A.; BONETTI, C.; ALEXANDRE, M.; ROUTABOUL, J.; KERFELD, C. A.; GRONDELLE, R. V.; ROBERT, B.; KENNIS, J. T. M.; KIRILOVSKI, D. A photoactive carotenoid protein acting as light intensity sensor. **Proc Natl Acad Sci USA**, v. 105, n. 33, p. 12075–12080, 2008.
- WILSON, A. AJLANI, G. VERBAVATZ, J.; VASS, M.; KERFELD, C. A.; KIRILOVSKI, D. A carotenoid protein has not been resolved in phycobilisome-related energy dissipation in cyanobacteria. **Plant Cell**, v. 18, p. 992-1007, 2006.

- WU, Y.; SIMMONS, B. A.; SINGER, S. W. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. **Bioinformatics**, v. 34, p. 605-607, 2016.
- XU, L.; SUN, H.; LI, Y.; WANG, J.; HOU, T. The impact of force fields and ligand charge models. **The Journal of Physical Chemistry B**., v. 117, p. 8408–21, 2013.
- ZIEGLER, T.; MOGLICH, A. Photoreceptor engineering. Front Mol Biosci, v. 2, p. 30, 2015.
- ZHAO Z.; CRISTIANE, A.; ROSEN, GAIL. Keeping up with the genomes: efficient learning of our increasing knowledge of the tree of life. **BMC Bioinformatics**, v. 21, p. 412, 2020.