UNIVERSIDADE DE SÃO PAULO

INSTITUTO DE FÍSICA DE SÃO CARLOS

FLÁVIO RODOLFO ROSSETO

Estudos estruturais e funcionais de STM3615 de Salmonella enterica: uma proteína contendo ambos os domínios GGDEF-EAL envolvidos na biossíntese de c-di-GMP

São Carlos
2016

FLÁVIO RODOLFO ROSSETO

Estudos estruturais e funcionais de STM3615 de Salmonella enterica: uma proteína contendo ambos os domínios GGDEF-EAL envolvidos na biossíntese de c-di-GMP

Tese apresentada ao Programa de Pós-Graduação em Física do Instituto de Física de São Carlos da Universidade de São Paulo, para obtenção do título de Doutor em Ciências.

Área de concentração: Física Aplicada
Opção: Física Biomolecular
Orientador: Prof. Dr. Marcos Vicente de Albuquerque Salles Navarro

Versão Original

São Carlos
2016

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica revisada pelo Serviço de Biblioteca e Informação do IFSC, com os dados fornecidos pelo(a) autor(a)

Rosseto, Flávio Rodolfo
Estudos estruturais e funcionais de STM3615 de Salmonella enterica: uma proteína contendo ambos os domínios GGDEF-EAL envolvidos na biossíntese de c-di-GMP / Flávio Rodolfo Rosseto; orientador Marcos Vicente de Albuquerque Salles Navarro -- São Carlos, 2016.
95 p.

Tese (Doutorado - Programa de Pós-Graduação em Física Biomolecular) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2016.

1. Biofilme bacteriano. 2. GGDEF-EAL. 3. c-di-GMP. I. Navarro, Marcos Vicente de Albuquerque Salles, orient. II. Título.

FOLHA DE APROVAÇÃO

Flávio Rodolfo Rosseto

Tese apresentada ao Instituto de Física de São Carlos da Universidade de São Paulo para obtenção do título de Doutor em Ciências. Área de Concentração: Física Aplicada - Opção: Física Biomolecular.

Aprovado(a) em: 12/12/2016

Comissão Julgadora

Dr(a). Marcos Vicente de Albuquerque Salles Navarro

Instituição: (IFSC/USP)

Dr(a). Glaucius Oliva

Instituição: (IFSC/USP)

Dr(a). Antônio José da Costa Filho

Instituição: (FFCLRP/USP)

Dr(a). Frederico José Gueiros Filho

Instituição: (IQ/USP)

Dr(a). Artur Torres Cordeiro

Instituição: (CNPEM/Campinas)

À Deus e a meus pais por me proporcionarem
a oportunidade de me dedicar aos estudos
independente de qualquer dificuldade.

AGRADECIMENTOS

À Deus acima de tudo;

Aos meus pais Antônio e Dirce e meu irmão Fábio por me apoiar e aconselhar durante as diversas fases da minha vida até o presente momento;

Ao Professor Marcos Navarro, pela orientação, ensinamentos e oportunidades;

A todos os professores do IFSC nos quais me proporcionam um ótimo ambiente de aprendizado durante esses 10 anos em que estive na USP São Carlos;

Aos amigos de laboratório por toda a jornada que tivemos juntos durante o doutorado, todo aprendizado compartilhado, “gordices”, momentos felizes e conselhos nos momentos não tão felizes assim, especialmente para Naiara, Bruno, Nathalya, Deia, Helton, Juliana e Everton;

Aos amigos, que mais próximos ou a distância, fizeram parte direta ou indiretamente de toda minha vida acadêmica e compartilharam de seus conhecimentos no laboratório e sobre a vida e perseverança na busca de um sonho: Guilherme, Lis, Renatinha, Malu, Vivi, Fran, Marisa, Caio e Ana.

Às técnicas Andressa, Bel e Patrícia pelos conselhos, auxílios e companheirismo;

Ao Instituto de Física de São Carlos por todas as oportunidades;

Ao grupo de cristalografia por disponibilizar suas instalações e equipamentos para o desenvolvimento de alguns experimentos;

Ao Laboratório Nacional de Luz Síncrotron pelo disponibilização das linhas de luz MX2 e SAXS2;

Ao LNBio por disponibilizar suas instalações e equipamentos;

Ao Diamond Light Source, pela oportunidade cedida para coleta de dados de Cristalografia;

À Fundação de Amparo à Pesquisa do Estado de São Paulo, pela concessão da bolsa de doutorado e pelo apoio financeiro para a realização desta pesquisa (2012/01711-5);

À todos direta ou indiretamente envolvidos para a realização deste trabalho;

Ao professores e funcionários do IFSP - campus Itapetininga e colaboradores da Libbs-Biotec pela oportunidade, apoio e compreensão nos momentos finais do doutorado.

“A menos que modifiquemos a nossa maneira de pensar, não
seremos capazes de resolver os problemas causados
pela forma como nos acostumamos a ver o mundo”

Albert Einstein

RESUMO

ROSSETO, F. R. **Estudos estruturais e funcionais de STM3615 de Salmonella enterica: uma proteína contendo ambos os domínios GGDEF-EAL envolvidos na biossíntese de c-di-GMP.** 2016. 95 p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2016.

A formação de biofilmes bacterianos é um fenômeno bem conhecido, caracterizado pela formação de uma comunidade bacteriana estática, embebida em uma matriz exopolimérica, regulada pela molécula sinalizadora c-di-GMP. Os domínios proteicos que catalisam a síntese (GGDEF) e degradação (EAL e HD-GYP) de c-di-GMP estão presentes em grande quantidade em quase todos os genomas bacterianos sequenciados até hoje. Dentre as diversas proteínas envolvidas nas vias de sinalização mediadas por esse nucleotídeo, uma grande parcela são proteínas transmembranares que possuem ambos domínios GGDEF e EAL. Funcionalmente, esses domínios se apresentam em todas combinações: ambos degenerados ou conservados e combinações GGDEF-degenerado/EAL-conservado ou vice-versa. Enquanto que domínios conservados potencialmente apresentam atividade catalítica, os degenerados geralmente convertem-se em domínios estruturais ou receptores de c-di-GMP. Embora recentes estudos estruturais revelaram detalhes de proteínas com ambos domínios degenerados (LapD) ou ativos (MorA), pouco se sabe sobre uma das combinações mais representativas: GGDEF-degenerado/EAL-conservado. Nesse trabalho, realizamos um estudo estrutural e funcional da proteína STM3615 de *Salmonella enterica*, que apresenta um domínio periplasmático de função desconhecida, seguido pelos domínios citoplasmáticos HAMP, GGDEF-degenerado e EAL-conservado. Através de diferentes construções citoplasmáticas solúveis de STM3615, confirmamos que essa proteína apresenta atividade fosfodiesterase, mesmo quando o domínio EAL encontra-se isolado. Corroborando com sua atividade catalítica, estudos em solução, tais como SAXS e cromatografia de exclusão molecular, mostraram que o EAL isolado de STM3615 é dimérico, um pré-requisito para ser ativo. Utilizando uma construção com os domínios GGDEF-EAL determinamos sua estrutura cristalográfica a uma resolução de 2,5 Å. Comparada com proteínas de arquitetura próxima, como o receptor de c-di-GMP LapD de *Pseudomonas fluorescens*, ou a enzima bifuncional MorA de *Pseudomonas aeruginosa,* sua estrutura se assemelha muito mais a essa última. Em particular, a hélice que conecta os domínios GGDEF e EAL possui a mesma extensão que a de MorA, maiores que a encontrada em LapD. Como a hélice pequena de LapD está relacionada com sua plasticidade conformacional interdomínios, a estrutura apresentada nesse trabalho sugere as proteínas *dual domain* cataliticamente ativas (EAL-mono ou bifuncionais) sejam estruturalmente rígidas. Combinando esses resultados com uma análise computacional feita em outras 150 sequências representativas de proteínas *dual domain*, propomos mecanismos catalíticos distintos para as enzimas bifuncionais e as EAL-monofuncionais. Enquanto que

essas últimas formam dímeros estáveis através do domínio EAL, numa conformação apta para interagir e degradar c-di-GMP, as enzimas bifuncionais apresentam transições oligoméricas mediadas por interação de c-di-GMP com EAL, impondo atividades ciclase (GGDEF) e fosfodiesterase (EAL) excludentes. Por fim, baseados nesses mecanismos e na arquitetura de STM3615, ainda especulamos mecanismos funcionais *in vivo* compatíveis com o tema emergente de interações proteicas e localização do sinal nas vias de sinalização mediadas por c-di-GMP.

Palavras-chave: Biofilme bacteriano. GGDEF-EAL. c-di-GMP.

ABSTRACT

ROSSETO, F. R. **Structural and functional studies of STM3615 from *Salmonella enterica*: a GGDEF-EAL-containing protein involved in c-di-GMP biosynthesis.** subtitle. 2016. 95 p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2016.

The formation of bacterial biofilms is a well-established phenomenon regulated by the signaling molecule c-di-GMP, characterized by the establishment of a static bacterial community embedded in a exopolymeric matrix. The domains responsible for the synthesis (GGDEF) or degradation (EAL and HD-GYP) of c-di-GMP are present in multiple proteins in nearly all bacterial genomes sequenced to date. Among the multiple and structurally diverse proteins involved in c-di-GMP signaling and biosynthesis, a large class are transmembrane proteins bearing both EAL and GGDEF domains. Functionally, these domains are presented in all combinations: both degenerate or conserved and combinations GGDEF-degenerated/EAL-conserved or vice versa. While the predicted conserved domains exhibit catalytic activity, the degenerate usually converted into structural domains or c-di-GMP receptors. While structural studies have revealed details of proteins with both domains degenerated (LapD) or conserved (MorA), little is known about one of the most representative combinations: GGDEF-degenerated/EAL-conserved. In this work, we conducted a structural and functional study of *Salmonella enterica* STM3615 protein, which has a periplasmic domain of unknown function, followed by cytoplasmic domains HAMP, GGDEF-degenerated and EAL-conserved. Through different soluble cytoplasmic constructs of STM3615, we confirmed that this protein has phosphodiesterase activity, even with the isolated EAL domain. In agreement with its catalytic activity, solution studies, such as SAXS and size exclusion chromatography, showed that STM3615 isolated EAL is dimeric, a prerequisite for phosphodiesterase activity. Using a construct with the isolated EAL-GGDEF domains, we determine its crystal structure to a resolution of 2.5 Å. Compared to the architectural closed c-di-GMP receptor LapD from *Pseudomonas fluorescens* and the bifunctional enzyme MorA from *Pseudomonas aeruginosa*, STM3615 structure is more similar to the latter. In particular, the α-helix connecting the domains GGDEF and EAL has similar extension, longer than the helix found in LapD. Given that this helix in LapD is essential for its inter-domain conformational plasticity, the structure presented in this study suggests the dual domain catalytically active proteins are structurally rigid. Combining these results with a computational analysis with 150 representative sequences containing the tandem GGDEF-EAL domains, we propose distinct catalytic mechanisms for bifunctional and monofunctional EAL enzymes. While the latter form stable dimers through the EAL domain, a conformation prompted to interact and degrade c-di-GMP, the bifunctional enzymes present oligomeric transitions mediated by interaction of c-di-GMP with EAL domain, imposing excluding cyclase (GGDEF) or phosphodiesterase (EAL) activities. Finally,

based on these mechanisms and STM3615 architecture, we also speculated about functional mechanisms *in vivo* consistent with the emerging theme of protein interactions and localized signal involved in signaling pathways mediated by c-di-GMP.

Keywords: Biofilm bacterial. GGDEF-EAL. c-di-GMP.

LISTA DE FIGURAS

Figura 1 - Representação do ciclo de formação de biofilme. Bactérias podem reversivelmente aderir à superfícies. Expressão de moléculas de adesãofornece uma forte base para que as bactérias se tornem estáticas, levando à colonização. As células então secretam uma matriz extracelular consistente de moléculas poliméricas, induzindo a formação de macrocolônias. Cada estágio coincide com um padrão de expressão de distintos genes, sugerindo que a formação de biofilme é um processo altamente regulado. Por fim, as bactérias podem desprender-se da colônia, novamente através de processos ativos e regulados.[1] 27

Figura 2 - Estrutura e funções fisiológicas de c-diGMP. A nível celular, c-di-GMP é controlado por diguanilato ciclases que possuem domínios GGDEF (vermelho) e fosfodiesterases específicas que apresentam domínios EAL ou HD-GYP (azul). c-di-GMP pode, por exemplo, reduzir a mobilidade por baixa expressão flagelar em *Pseudomonas aeruginosa* ou interferir com a função motor de flagelos em *Escherichia coli* e *C. crescentus*. Baixos níveis de c-di-GMP são necessários para a expressão de genes de virulência aguda (por exemplo, em *Vibrio cholerae*).[11] Em todas as bactérias testadas, altas concentrações de c-di-GMP estimularam a variação das funções associadas a biofilmes, como a formação de fímbrias e várias outras matrizes de exopolissacarídeos. [14] 29

Figura 3 - Representação da predição e disposição das hélices transmembranares e dos domínios periplasmático, HAMP, GGDEF e EAL da LapD de *Pseudomonas fluorescens.* TM: Hélice transmembranar; DP: Domínio periplasmático. 33

Figura 4 - Modelo de como concentrações de Pi controlam a formação de biofilmes de *P. fluorescens*. Esquerda: Em altas concentrações de Pi, c-di-GMP se acumula nas células. LapD liga-se a c-di-GMP e sequestra a protease LapG na membrana interna, promovendo a manutenção da adesina LapA na superfície celular. Nessas condições, as bactérias se aderem irreversivelmente em um substrato permitindo a progressão do biofilme. Direita: Quando baixas concentrações extracelulares de Pi são sentidas pelo sistema dois-componentes PhoR/Pst, uma cascata de sinalização é disparada resultando na transcrição da fosfodiesterase da família EAL RapA. A atividade de RapA na conversão de c-di-GMP para pGpG, abaixa os níveis intracelulares dessa molécula levando à dissociação de c-di-GMP da LapD. Na forma apo, LapD não consegue interagir com LapG, liberando-a para clivar a parte N-terminal de seu substrato natural no periplasma, LapA, promovendo sua remoção da superfície celular. A liberação de LapA de bactérias no biofilme leva a sua desadesão da superfície,

sendo que as bactérias planctônicas sem LapA ficam impedidas de se aderirem novamente à uma superfície. 34

Figura 5 - A) Modelos estruturais para a proteína LapD completa na forma auto-inibida e ativa ligada a c-di-GMP baseados nas estruturas cristalinas e estudos *in vitro* e *in vivo*. Somente a forma ligada com c-di-GMP é capaz de ligar-se a LapG no periplasma. O domínio HAMP foi modelado por homologia com base em alinhamentos sequenciais e informações estruturais disponíveis.[3-4] B) Modelo para o controle da formação de biofilme mediado por LapD. O desenho representa o modelo atual baseado nas análises estruturais e outros resultados. [6-9] 35

Figura 6 - Representação da predição e disposição dos domínios REC, PAS, GGDEF e EAL da FimX de *Pseudomonas aeruginosa.* 36

Figura 7 - A) Alinhamento estrutural do complexo PilZ-FimXEAL e LapD GGDEF-EAL apo. As estruturas estão coloridas de forma a identificar as diferentes regiões, sendo a PilZ (amarelo), FimXEAL (verde e azul) e GGDEF (salmão), EAL (ciano) e hélice S (laranja) da LapDGGDEF-EAL. A ligação do c-di-GMP na FimX está disposta na cor vermelha. O barril TIM do domínio EAL da LapD e FimX foram usados como base do alinhamento estrutural. B) A figura ilustra a proposta de mecanismo para as duas possíveis conformações dos domínios GGDEF-EAL da FimX dadas através da interação com PilZ e c-di-GMP. Na linha superior, podemos observar os domínios GGDEF-EAL em conformação fechada na ausência de c-di-GMP e PilZ, de forma similar ao observado com a apo-LapDGGDEF-EAL.[44] Essa conformação é estabilizada pela interação do barril TIM do domínio EAL com a hélice S (superior esquerda), onde na LapDGGDEF-EAL, essa conformação impede a interação da proteína com c-di-GMP. Como os domínios GGDEF-EAL da XacFimX e P. fluorescens LapD apresentam, respectivamente, apenas 25% e 39% de identidade e similaridade de sequência, não foi possível elucidar se o c-di-GMP poderia se ligar à FimX em conformação fechada (como indicada na representação transparente – superior direita). Nas figuras na região central e abaixo, observamos que os domínios GGDEF e EAL podem se dissociar, passando então para a conformação aberta, que permite a ligação de c-di-GMP (estrela vermelha) e também da PilZ.[50] 37

Figura 8 - Representação Da predição e disposição das hélices transmembranares e domínios PAS, GGDEF e EAL da MorA de *Pseudomonas aeruginosa.* TM: Hélice transmembranar. 38

Figura 9 - Estrutura Cristalográfica dos domínios GGDEF-EAL da MorA. A) Na presença de c-di-GMP, a estrutura da MorA apresenta a formação de um dímero. B) Com a sobreposição das estruturas nos dois estados oligoméricos, monômero (azul claro) na ausência de c-di-GMP e dímero (azul escuro) na presença de c-di-GMP. A

hélice repressora, α5 destacada também em (A), muda a orientação e tamanho (C) de acordo com o estado oligomérico da enzima, afetando estruturalmente a posição dos aspárticos que coordenam a ligação do Mg^{2+}, onde no monômero não há atividade pois os aspárticos ficam afastados do sítio ativo, porém, a extensão do *loop* na forma dimérica permite que os resíduos se posicionem de forma a coordenar os metais e o substrato de forma adequada para promover a hidrólise.[10] 39

Figura 10 - Presença de mutantes de GG(D/E)EF/EAL de *Salmonella typhimurium* UMR1 no bolo fecal de ratos tratados com estreptomicina. Ratos tratados com estreptomicina foram infectados com quatro grupos de cepas (selvagem e três mutantes de GG(D/E)EF/EAL). As variações dos mutantes no intestino dos ratos foram avaliadas pela contagem de unidades de colônias formadas (CFU) por grama de fezes até 30 dias depois da infecção. Três dos vinte mutantes (STM2672, STM3615 e STM4551) sumiram de forma mais rápida do intestino dos camundongos.[5] 40

Figura 11 - Predição de hélices transmembranares da STM3615 realizada pelo servidor TMHMM. As regiões vermelhas demonstram a alta probabilidade de regiões transmembranares no N-terminal (peptídeo sinal) e logo antes do domíno HAMP. 45

Figura 12 - Análise da delimitação dos domínios HAMP (verde), GGDEF (vermelho) e EAL (azul) 45

Figura 13 - Análise computacional para predição de estruturas secundárias da STM3615 pelo servidor PSIPRED..... 46

Figura 14 - Representação da predição e disposição dos domínios periplasmático, HAMP, GGDEF, EAL e hélices transmembranares da STM3615 de Salmonella enterica. TM: Hélice Transmembranar, DP: domínio periplasmático..... 46

Figura 15 - Clonagem das construções da STM3615: Análise eletroforética do produto de PCR e de restrição em gel de agarose 0,8%. (A) e (C) MM) Padrão de massa molecular, 1) DP^{36-145} (327 Kb); 2) DP^{36-150} (342 Kb); 3) DP^{36-155} (357 Kb); 4) DP^{41-150} (312Kb); 5) DP^{41-155} (327Kb); 6) DP^{46-145} (342 Kb); 7) DP^{46-150} (297 Kb); 8) DP^{46-145} (312Kb). (B) e (D) MM) Padrão de massa molecular, 1) HAMP-GGDEF-$EAL^{168-657}$ (1467 Kb); 2) HAMP-GGDEF-$EAL^{173-657}$ (1452 Kb); 3) GGDEF-$EAL^{207-657}$ (1350 Kb) 4) GGDEF-$EAL^{231-657}$ (1278 Kb); 5) GGDEF-$EAL^{236-657}$ (1263 Kb); 6) GGDEF-$EAL^{241-657}$ (1248 Kb)..... 57

Figura 16 - Purificação da construção GGDEF_$EAL^{242-668}$. (A) Gel SDS-PAGE 15%, onde podemos observar: (A) (1) Marcador molecular, (2) Pré-Indução, (3) Induzido, (4) *Pellet*, (5) Sobrenadante, (6) Eluição da coluna de afinidade, (7) Clivagem com UPL1 e (8) GGDEF_EAL^{231-}

^{657} após cromatografia de exclusão molecular. (B) Cromatograma de exclusão molecular da GGDEF_EAL$^{242-668}$, com a coluna Superdex75 16/60. 58

Figura 17 - Cromatografia de Exclusão Molecular analítica: domínios citoplasmáticos HAMP-GGDEF-EAL$^{184-668}$ (preto), GGDEF-EAL$^{242-668}$ (vermelho), EAL$^{392-668}$ (azul), HAMP-GGDEF$^{181-499}$ (verde) e GGDEF$^{242-399}$ (magenta) da STM3615. A) Curvas de eluição das construções na coluna superdex200 10/300: HAMP-GGDEF-EAL$^{184-668}$: 12,94 mL, GGDEF-EAL$^{242-668}$: 13,42 mL, EAL: 15,23 mL, HAMP-GGDEF$^{181-399}$: 15,64 mL e GGDEF$^{242-399}$: 18,21 mL. B) Calibração da coluna superdex200 10/300 com proteínas-padrão e ajuste dos pontos experimentais na curva de calibração dos domínios. C) Gel SDS 15% de amostras dos domínios HAMP-GGDEF-EAL$^{181-399}$ (1), GGDEF-EAL$^{242-668}$ (2), EAL$^{392-668}$ (3), HAMP-GGDEF$^{181-399}$ (4) e GGDEF$^{242-399}$ (5)..... 60

Figura 18 - Curvas de espalhamento de raios-X a baixo ângulo, curvas de distribuição de distâncias e modelos de SAXS. Podemos observar que para as três construções, HAMP-GGDEF-EAL (A), GGDEF-EAL$^{\mathbf{231-657}}$ (B) e EAL (C), as curvas de experimentais se ajusta bem em relação aos modelos de baixa resolução obtidos e também às estruturas cristalográficas mostrando na inserção as curvas de Guinier[12] que demonstram a qualidade satisfatória dos dados obtidos. Todos os modelos apresentados foram comparados a estrutura cristalográfica dos domínios GGDEF-EAL, obtida neste trabalho, apresentadas em três disposições, rotacionadas de 90° em relação a anterior. A figura D, ilustra a sobreposição das curvas de distribuição de distâncias e dos três modelos obtidos. Observamos também que conforme esperado, as curvas de distribuição de distâncias possuem diâmetros máximos diferentes para cada uma das construções analisadas, sendo maior para a HAMP-GGDEF-EAL e menor para EAL (D)..... 63

Figura 19 - Caracterização do produto da hidrólise do c-di-GMP pela GGDEF-EAL$^{242-668}$. Na figura A apresentamos o cromatograma completo e em B destacamos os picos que foram submetidos para análises de espectrometria de massas. As curvas em preto representam o c-di-GMP, utilizado aqui como controle do experimento (condição inicial), com pico em aproximadamente 9.3 minutos. As curvas em azul, referem-se ao produto da reação enzimática da construção GGDEF-EAL$^{242-668}$ com c-di-GMP, o pGpG, com pico em aproximadamente 9 minutos..... 64

Figura 20 - Espectro de massas dos produtos de reação com a fosfodiesterase com c-di-GMP. Os picos com a relação carga/massa equivalentes a 691.5 e 709.6 correspondem ao c-di-GMP e 5'pGpG, respectivamente. 65

Figura 21 - Caracterização da atividade enzimática das construções HAMP-GGDEF-EAL$^{184-668}$, HAMP-GGDEF$^{181-399}$, GGDEF$^{242-399}$, GGDEF-

EAL$^{242-668}$ e EAL$^{392-399}$. A) As construções que possuem o domínio GGDEF, foram incubadas com GTP e analisadas posteriormente por HPLC. A figura nos mostra que para nenhuma construção foram obtidos picos de c-di-GMP, ou seja, os domínios GGDEF encontram-se na forma inativa. B) As construções que possuem domínio EAL foram submetidas a ensaios com o substrato c-di-GMP e como podemos observar na figura ampliada (C), todas elas apresentaram um pico correspondente a pGpG, mostrando que todas as construções são ativas. 66

Figura 22 - Cristais e padrões de difração da GGDEF_EAL$^{242-668}$. A) Cristais obtidos na condição 15% *tacsimate* pH7,0 (Hampton), 0,1 M HEPES pH7,0 e 4% PEG 3350 e padrão de difração a 3,5 Å. B) Cristais obtidos com a proteína contendo selenometionina na condição 10% m/v PEG4000, 20% v/v Glicerol, 0,03 M Haletos e 0,1 M MES/Imidazol pH 6,5 e padrão de difração a 2,8 Å. C) Cristais obtidos na condição 10% m/v PEG 8000, 20% v/v Etilenoglicol, 0,03 M NPS e 0,1 M MÊS/Imidazol pH 6,5 e padrão de difração a 2,5 Å. 68

Figura 23 - Cristais do DP$^{46-150}$. Cristais obtidos para a construção DP$^{46-150}$ na condição 10% m/v PEG 20000, 20% v/v PEG MME 550, 0,02 M de D-glicose, D-manose, D-galactose, L-fucose, D-Xilose, N-acetil-D-glucosamina, 0,1 M bicina/Trizma base pH8,5. 70

Figura 24 - Estrutura cristalográfica dos domínios GGDEF e EAL da STM3615. A) Estrutura, na forma de um homodímero ativo para fosfodiesterase, com cada unidade em cores diferentes, destacando os domínios GGDEF (azul) e EAL (verde), unidos por uma hélice α (amarelo). B) Identificação da conformação dos domínios EAL (em verde) formado por 8 hélices α e 8 fitas β na conformação de um barril alpha-beta e GGDEF (em azul), na conformação de um sanduíche alpha-beta, composto por uma folha β central de 4 fitas β antiparalelas, envolta por 5 hélices α. 71

Figura 25 - Conservação de resíduos dos domínios GGDEF e EAL da STM3615. A) A figura apresenta a análise de conservação dos resíduos necessários para promover a atividade catalítica do domínio EAL gerada pelo programa WebLogo, sendo eles o E444, N502, E534, E537, D564 ,K588 e E621, respectivamente. B) Sobreposição do domínio EAL da STM3615 com o c-di-GMP da estrutura da fosfodiesterase de Thiobacillus denitrificans (PDB:3N3T), destacando a posição dos resíduos E444, N502, E534, E537, D564 ,K588 e E621, identificados como responsáveis pela hidrólise do c-di-GMP. C) Através do alinhamento múltiplo de sequências e o score dado a conservação dos resíduos pelo programa Consurf, podemos verificar que a região mais conservada dos domínios é a região catalítica do barril do EAL, mostrada ao lado esquerdo da figura. D) Sobreposição do dímero dos domínios GGDEF_EAL da STM3615 com o c-di-GMP do PDB 3N3T, onde observa-se que as moléculas de c-di-GMP estão em

posições relativas diferentes se comparadas entre as unidades monoméricas. 72

Figura 26 - Comparação das hélices entre os domínios GGDEF e EAL: STM3615, MorA, FimX e LapD. Através da sobreposição das estruturas é podemos observar que as quatro estruturas possuem uma hélice α na região que conecta os domínios, porém, as hélices das proteínas com resíduos catalíticos conservados, STM3615 e MorA, tem tamanho proporcionalmente maior que os receptores de c-di-GMP, LapD e FimX, fazendo com que sua estrutura esteja em uma forma mais estendida..... 74

Figura 27 - Conservação dos domínios GGDEF e EAL. Podemos observar que entre as sequências apresentadas, somente a WspR possui o sítio de inibição RxxD e a região catalítica GGEEF conservadas. As demais proteínas estão divididas entre as ativas e não ativas para fosfodiesterase. 75

Figura 28 - Predição de estrutura secundária de sequências com EAL conservado e degenerado. As diversas sequências da região da hélice que conecta os domínios GGDEF e EAL, determinadas pelo alinhamento da sequência conhecida para a STM3615, possuem um padrão onde a hélice sempre está entre duas fitas betas, porém, ao compararmos as hélices do grupo com EAL conservado e não conservado, observa-se um padrão diferente entre elas, onde aquelas com EAL conservado são maiores e com média de 21 resíduos e aquelas com EAL degenerado são menores, ou duplas, em média com 15 resíduos. 76

Figura 29 - Representação do número de sequências com domínios GGDEF-EAL quanto a atividade, predição de estrutura secundária e sítio inibitório do domínio GGDEF. A) Entre as 150 sequências, nota-se que a maior parte encontra-se na porção onde os resíduos fundamentais para a catálise estão conservados e a minoria na parcela em que apenas os resíduos catalíticos do domínio GGDEF não estão conservados. B) O número de resíduos entre as fitas das extremidades onde encontra-se a hélice não apresenta grandes variações, porém, ao analisar o tamanho das hélices, observa-se uma similaridade entre o número de resíduos que formam a hélice que conecta os domínios GGDEF e EAL para a fração em que resíduos catalíticos conservados apenas no domínio EAL ou para ambos os domínios, sugerindo então que de maneira geral estes possuem hélices estendidas, enquanto para o caso em que temos apenas o domínio GGDEF conservado e GGDEF-EAL não conservados, os dados sugerem hélices mais curtas. C) Entre as sequências analisadas, para os domínios GGDEF-EAL conservados, a presença do sítio inibitório RxxD estava presente em pouco mais da metade das sequências, sugerindo dois mecanismos distintos para a regulação da síntese e degradação de c-di-GMP nessa situação. A presença do RxxD no caso onde nenhum dos domínios apresenta os resíduos catalíticos

conservados ou conservados apenas no EAL, possuem baixa taxa de sítio inibitório, provavelmente porque este não tem nenhum papel fundamental nas duas situações, diferentemente do caso onde apenas os resíduos catalíticos do domínio GGDEF estão conservados, em que todas as sequências apresentaram RxxD conservado. 79

Figura 30 - Análise da sequência de aminoácidos e estrutura cristalográfica da MorA de acordo com a atividade catalítica. A) Alinhamento da sequência de MorA e WspR de Pseudomonas aeruginosa, mostrando a ausência do sítio inibitório que é dado pelo motivo RxxD, destacado em azul, cinco resíduos a esquerda do motivo GGDEF. B) Segundo Phippen e colaboradores, a atividade catalítica de diguanilato ciclase ocorre em construções que apresentam somente o domínio GGDEF ou simultaneamente de acordo com a construção que apresenta os domínios GGDEF-EAL, o que não é justificado por (C), já que quando o domínio EAL encontra-se na forma dimérica, os domínios GGDEF estão em disposições opostas, o que não permitem sua atividade catalítica. 80

Figura 31 - Comparação das hélices α5 do monômero e dímero da MorA e STM3615. A) e B) Se comparada às hélices α5 do monômero (lilás) e dímero da MorA (vermelho), a estrutura da STM3615 (azul) apresenta essa hélice mais próxima ao monômero da MorA, mesmo apresentando o dímero na estrutura cristalográfica obtida. C) Se compararmos uma região um pouco maior que apenas as hélices, notamos que os resíduos da STM3615 estão em disposição nas quais poderiam coordenar os dois metais necessários para promover a hidrólise do c-di-GMP e a fenilalanina que coordena o c-di-GMP no sítio catalítico encontra-se em posição intermediária à tirosina do monômero e dímero da MorA. 82

Figura 32 - Proposta de mecanismo para MorA. Em baixos níveis de c-di-GMP, os domínios GGDEF encontram-se na forma dimérica com seus sítios catalíticos com arranjo face-a-face, inativando a fosfodiesterase e contribuindo para o aumento do pool de c-di-GMP local (esquerda). Já em altos níveis de c-di-GMP, a fosfodiesterase toma a forma dimérica, inativando a fosfodiesterase, auxiliando na regulação do pool de c-di-GMP local (direita). 83

Figura 33 - Hipótese de mecanismo de ação da STM3615. A STM3615, através de interações com outras proteínas como diguanilato ciclases (DGC) e fosfodiesterases (FDE) controlam o pool de c-di-GMP constante. Um estímulo no domínio periplasmático é transmitido através do domínio HAMP, alterando o pool de c-di-GMP, através do deslocamento de algumas das proteínas envolvidas no mecanismo, diminuindo a concentração local do nucleotídeo e sinalizando a dispersão do biofilme bacteriano..... 85

LISTA DE TABELAS

Tabela 1 - Construções da STM3615 do domínio periplasmático (DP), HAMP, GGDEF e EAL, massa molecular teórica e vetores de expressão. 48

Tabela 2 - Etapas de reação de PCR para amplificação dos fragmentos de DNA de interesse. 48

Tabela 3 - Resumo dos resultados obtidos para clonagem, expressão, purificação e cristalização das construções da STM3615. 59

Tabela 4 - Volume de eluição (Ve) e massa molecular (MW) obtidas através de cromatografia de exclusão molecular analítica. 60

Tabela 5 - Parâmetros estruturais derivados dos dados de SAXS para os domínios HAMP-GGDEF-EAL$^{184-668}$, GGDEF-EAL$^{242-668}$ e EAL$^{392-668}$ da STM3615. 62

Tabela 6 - Massas moleculares calculadas através das curvas de SAXS das construções HAMP-GGDEF-EAL$^{184-668}$, GGDEF-EAL$^{242-668}$ e EAL$^{392-668}$. 62

Tabela 7 - Estatísticas de coleta, processamento e refinamento dos dados coletados para os cristais obtidos com a construção GGDEF-EAL$^{242-668}$. 69

Tabela 8 - Dados dos domínios GGDEF-EAL analisados. 77

SUMÁRIO

1 INTRODUÇÃO 27
1.1 Vias de sinalização mediadas por c-di-GMP 28
1.2 Bioquímica da síntese e degradação de c-di-GMP 29
1.2.1 Síntese do c-di-GMP: o domínio GGDEF 29
1.2.2 Hidrólise do c-di-GMP: domínios EAL e HD/GYP 30
1.3 Proteínas Dual Domain 31
1.3.1 LapD – *Pseudomonas fluorescens* 32
1.3.2 FimX – *Pseudomonas Aeruginosa* 35
1.3.3 MorA – *Pseudomonas Aeruginosa* 37
1.3.4 STM3615 – *Salmonella enterica* 39
2 OBJETIVOS 43
3 METODOLOGIA 45
3.1 Clonagem das construções da STM3615 45
3.2 Expressão e purificação das construções da STM3615 48
3.2.1 Cromatografia de exclusão molecular analítica 49
3.2.2 Espalhamento de raios-X a baixo ângulo (SAXS) 50
3.3 Ensaios de atividade enzimática 51
3.4 Cristalização e Estrutura de alta resolução dos domínios GGDEF-EAL 52
3.4.1 Cristalização 52
3.4.2 Estrutura cristalográfica dos domínios GGDEF-EAL 53
3.5 Análise de sequência, atividade e estrutura secundária de proteínas dual domain GGDEF-EAL 53
3.5.1 Seleção de sequências 54
3.5.2 Alinhamento e predição de atividade catalítica e estrutura secundária 54
4 Resultados 57
4.1 Clonagem, expressão e purificação das construções da STM3615 57
4.2 Estado Oligomérico dos domínios citoplasmáticos 59
4.2.1 Cromatografia de exclusão molecular analítica 59
4.2.2 Espalhamento de raios-X a baixo ângulo 61
4.3 Ensaios de atividade enzimática 64
4.3.1 Hidrólise do c-di-GMP 64
4.3.2 Ensaios enzimáticos das construções da STM3615 65
4.4 Ensaios de cristalização dos domínios solúveis 66

4.5 Estrutura cristalográfica dos domínios GGDEF_EAL da STM3615 70
4.6 Análises computacionais de sequencias codificantes para proteínas GGDEF-EAL .. 74
4.7 Modelo de mecanismo regulatório para enzimas GGDEF-EAL bifuncionais 79
4.8 Hipótese funcional para STM3615 baseada em sua estrutura e função 84
5 CONCLUSÃO 87
REFERÊNCIAS 89

1 INTRODUÇÃO

Recentemente, muita ênfase tem sido dada à noção de que culturas bacterianas não só existem como uma suspensão de células individuais, mas também aparecem e funcionam como comunidades multicelulares, os chamados biofilmes.[1-2]

As bactérias possuem dois estágios de vida distintos: planctônica, onde as células são individuais e comportam-se independentemente, e biofilmes. Um biofilme é definido por uma comunidade bacteriana que adere a uma superfície e produz uma matriz que envolve as células bacterianas, constituída de polissacarídeos extracelulares, proteínas e DNA (Figura 1).

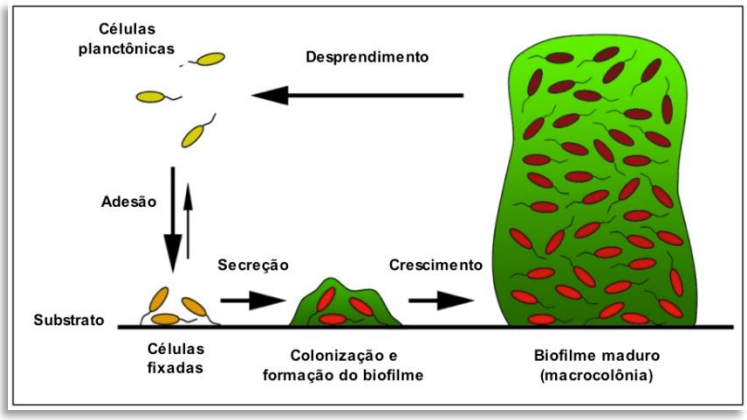


Figura 1 - Representação do ciclo de formação de biofilme. Bactérias podem reversivelmente aderir à superfícies. Expressão de moléculas de adesãofornece uma forte base para que as bactérias se tornem estáticas, levando à colonização. As células então secretam uma matriz extracelular consistente de moléculas poliméricas, induzindo a formação de macrocolônias. Cada estágio coincide com um padrão de expressão de distintos genes, sugerindo que a formação de biofilme é um processo altamente regulado. Por fim, as bactérias podem desprender-se da colônia, novamente através de processos ativos e regulados.[1]

Fonte: Adaptada de O'TOOLE; KAPLAN; KOLTER.[1]

Em particular, a formação de biofilmes desempenha um papel crucial nas infecções hospitalares, na colonização patológica de órgãos, implantes e cateteres, e infecções causadas por patógenos oportunistas.[13] Frequentemente, tais infecções estão além dos tratamentos tradicionais por apresentarem uma maior resistência contra antibióticos, principalmente devido à alta concentração de bactérias embebidas em uma matriz polimérica protetora e pela facilidade de transferência horizontal de genes.[15]

1.1 Vias de sinalização mediadas por c-di-GMP

Ocupando papel central no processo de formação de biofilme, encontra-se uma molécula mensageira encontrada apenas no mundo microbiano, a guanosina monofosfato (3'-5-)-cíclica dimérica (c-di-GMP), responsável pelo controle da secreção, adesão, motilidade celular e aumento na citotoxicidade[16] (Figura 2). De forma geral, níveis aumentados de c-di-GMP resultantes de alta atividade das ciclases GGDEF levam a um aumento na formação de biofilme e inibem a mobilidade mediada por flagelos.[17-18] Reciprocamente, baixos níveis do nucleotídeo, associados com a atividade fosfodiesterase das proteínas contendo o domínio EAL, suprimem a adesão de adesinas extracelulares, promovendo a dispersão do biofilme, e aumento da virulência bacteriana. Embora tais efeitos gerais tenham sido demonstrados em vários estudos de super-expressão,[17-18] a atividade redundante entre diferentes proteínas contendo domínios GGDEF e EAL no mesmo genoma é somente aparente.

Frequentemente, proteínas com organização de domínios e atividades catalíticas similares apresentam respostas fisiológicas distintas.[19-20] Esse é exatamente o caso para a família de proteínas contendo GGDEF e/ou EAL em *P. aeruginosa*. Um estudo em escala genômica demonstrou que apenas algumas proteínas apresentavam a tendência mencionada, correlacionando ciclases e fosfodiesterases com aumento e diminuição da formação de biofilme, respectivamente. Outras, no entanto, mostraram relação inversa ou nenhuma, sugerindo que essas proteínas tenham funções específicas, atividades distintas e/ou diferentes modos de regulação.[14]

Os domínios proteicos responsáveis pela síntese e degradação de c-di-GMP, as enzimas diguanilato ciclases (DGC) da família GGDEF e as fosfodiesterases (PDE) das famílias EAL ou HD-GYP, respectivamente, tem sido identificados em grande número e em quase todos os genomas bacterianos sequenciados até hoje.[17] Tais domínios são encontrados sozinhos ou associados a outros módulos sinalizadores. Também podem fazer parte de sistemas de sinalização dois-componentes, sugerindo uma correlação com a detecção de estímulos ambientais. Apesar de recentes progressos começarem a elucidar o funcionamento de algumas

enzimas chaves dessa via,[21-22] pouco se sabe sobre os mecanismos enzimáticos e regulatórios que controlam os níveis celulares de c-di-GMP.

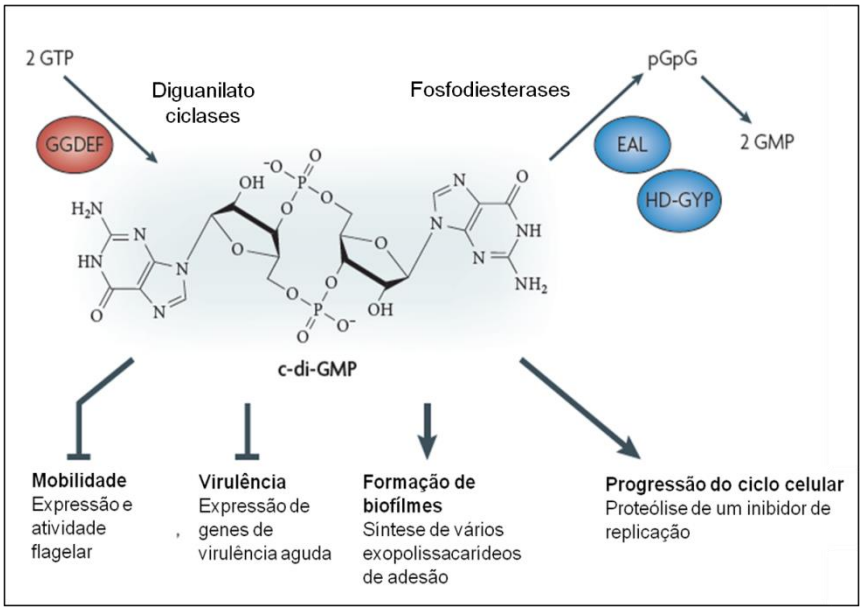


Figura 2 - Estrutura e funções fisiológicas de c-diGMP. A nível celular, c-di-GMP é controlado por diguanilato ciclases que possuem domínios GGDEF (vermelho) e fosfodiesterases específicas que apresentam domínios EAL ou HD-GYP (azul). c-di-GMP pode, por exemplo, reduzir a mobilidade por baixa expressão flagelar em *Pseudomonas aeruginosa* ou interferir com a função motor de flagelos em *Escherichia coli* e *C. crescentus*. Baixos níveis de c-di-GMP são necessários para a expressão de genes de virulência aguda (por exemplo, em *Vibrio cholerae*).[11] Em todas as bactérias testadas, altas concentrações de c-di-GMP estimularam a variação das funções associadas a biofilmes, como a formação de fímbrias e várias outras matrizes de exopolissacarídeos. [14]

Fonte: Adaptada de HENGGE.[14]

1.2 Bioquímica da síntese e degradação de c-di-GMP

1.2.1 Síntese do c-di-GMP: o domínio GGDEF

A evidência genética apresentada por algumas pesquisas, como a de Ausmees e colaboradores, sugere que o domínio GGDEF pode ser suficiente para atividade de diguanilato ciclases.[20, 23-25] Análises de bioinformática da sequência e estrutura do domínio GGDEF realizadas por Pei e Grishin relacionaram os domínios à atividade de ciclases.[26] Os autores descobriram que o domínio GGDEF possui função de adenilato/guanilato ciclases[27-28] e propuseram que o domínio GGDEF é

uma diguanilato ciclase predito por um loop envolvendo um motivo muito conservado, GG(D/E)EF, sendo parte do sitio de ligação do substrato (GTP). Posteriormente pesquisadores mostraram que a forma fosforilada da PleD é capaz de converter GTP em c-di-GMP *in vitro.*[29-31]

Benziman e colaboradores[23] revelaram o mecanismo de catálise dos domínios GGDEF para formação de c-di-GMP a partir de duas moléculas de GTP em uma reação de dois passos, com 5'-pppGpG como intermediário e duas moléculas de pirofosfato como subproduto. A natureza do dinucleotídeo c-di-GMP implica na necessidade dos domínios GGDEF se apresentarem como homodímeros ativos, onde dois monômeros se juntam para formar um sitio ativo na interface dimérica.[32] Cada monômero do domínio GGDEF contribui com uma molécula de GTP para a formação de uma ligação fosfodiester intermolecular. Nessa configuração, o motivo GG(D/E)EF forma um *hairpin*,[26] sendo que as duas primeiras glicinas estão envolvidas na ligação de GTP, enquanto o terceiro resíduo do motivo (asparagina/glutâmico) é indispensável para a catálise e também tem o papel de coordenar íons metálicos juntamente com o quarto resíduo do motivo, o ácido glutâmico.[33-34]

1.2.2 Hidrólise do c-di-GMP: domínios EAL e HD/GYP

Desde a elucidação funcional do domínio GGDEF quanto à síntese de c-di-GMP, o domínio EAL passou a ser visto como possível responsável pela hidrólise do dinucleotídeo. Benziman e colaboradores purificaram uma fosfodiesterase de *G. xylinus* e mostraram que essas proteínas hidrolisam c-di-GMP em di-GMP linear, ou seja, 5'-pGpG. Para que haja atividade específica das fosfodiesterases com c-di-GMP, é necessária a presença de Mn^{2+} ou Mg^{2+}, enquanto que na presença de Ca^{2+} a reação é fortemente inibida.[35] A atividade do domínio EAL depende da estrutura dos dois cátions metálicos, coordenados por duas moléculas de água, envolvidas com o ataque hidrolítico na ligação fosfodiester do c-di-GMP.

O produto da hidrólise do c-di-GMP, 5'-pGpG é posteriormente degradado para um monômero de pG aparentemente por enzimas diferentes que apresentam atividade independente de Ca^{2+}.[36] Assim como o domínio GGDEF, os domínios EAL com atividade de fosfodiesterase se caracterizam também como dímeros ou outros oligômeros *in vitro.*[22,37-38] O estado dimérico é crítico para a ativação da fosfodiesterase,[39; 40] sendo compatível para promover a hidrólise quando a unidade funcional do domínio EAL está ligada com c-di-GMP.[35,38]

1.3 Proteínas Dual Domain

Frequentemente os domínios GGDEF e EAL aparecem de forma isolada ou *in tandem* na mesma cadeia. Essa última classe de proteínas representa um enigma, dada a potencial presença de atividades antagônicas na mesma enzima. No entanto, a presença dessas proteínas é recorrente nos genomas sequenciados até o momento e representam uma grande parcela das proteínas contendo GGDEF ou EAL, colocando-as em um papel central nas vias de sinalização mediadas por c-di-GMP. Quando presentes na mesma cadeia polipeptídica, os domínios GGDEF e EAL podem se apresentar ambos ativos ou inativos, ou ainda um inativo e outro ativo, sendo que os inativos podem ser receptores de c-di-GMP em interações proteína-proteína ou proteína-RNA. Em geral esses domínios estão presentes, por exemplo, em proteínas modulares em conjunto com domínios REC, GAF, HAMP ou PAS.[41-42]

Dentre essas diversas proteínas GGDEF-EAL envolvidas nas vias de sinalização mediadas por c-di-GMP, uma grande parcela, aproximadamente 70%, possui hélices transmembranares (TM), responsáveis pelo ancoramento dessas proteínas à membrana ou conexão de domínios citoplasmáticos a domínios do periplasma (DP). O domínio N-terminal DP geralmente é responsável pelo reconhecimento de estímulos ambientais e celulares[17,43], como oxigenação do meio, luminosidade, falta de nutrientes, antibióticos ou até moléculas sinalizadoras, que podem ativar os domínios citoplasmáticos com atividade diguanilato ciclase ou

fosfodiesterase para então regular os níveis de c-di-GMP e consequentemente a fisiologia e comportamento bacteriano.[14]

Proteínas que apresentam os domínios GGDEF e EAL *in tandem* ainda são pouco descritas na literatura, com informações e estruturas apenas de algumas disponíveis como a LapD de Pseudomonas *fluorescens*[6,8,44] MorA[10,45] e FimX[46-48] de Pseudomonas aeruginosa e STM3615 de Salmonella entérica,[5,49] sendo a última, a proteína de interesse do presente trabalho.

1.3.1 LapD – *Pseudomonas fluorescens*

Entre as proteínas transmembranares citadas, uma arquitetura (figura 4) de particular interesse encontra-se de forma muito recorrente nos genomas bacterianos sequenciados (704 sequências na base PFAM). Em comum, essa sub-classe apresenta na porção C-terminal os domínios EAL e GGDEF seguido do domínio HAMP, um módulo sinalizador muito encontrado em proteínas bacterianas responsável por mediar dimerização e ativação através de mudanças conformacionais.[3] Conectado ao domínio HAMP, em direção ao N-terminal, aparece uma única hélice transmembranar que liga a porção citoplasmática da proteína ao domínio periplasmático, este geralmente bem variante, e de função desconhecida, exceto por um único caso. Somente uma proteína com tal arquitetura foi estudada até o momento, LapD de *P. fluorescens*, sendo que seu mecanismo de sinalização do citoplasma para o periplasma abriu um novo paradigma nas vias de sinalização mediadas por c-di-GMP.[6,9,44]

Figura 3 - Representação da predição e disposição das hélices transmembranares e dos domínios periplasmático, HAMP, GGDEF e EAL da LapD de *Pseudomonas fluorescens.* TM: Hélice transmembranar; DP: Domínio periplasmático.

Fonte: Elaborada pelo autor.

Resumidamente, em condições de baixos níveis de c-di-GMP intracelular, a proteína LapD encontra-se auto-inibida, permitindo que LapG, uma cisteíno protease periplasmática, degrade especificamente a porção N-terminal de LapA, liberando-a da superfície celular e promovendo a desestabilização do biofilme. A medida que os níveis de c-di-GMP aumentam, a ligação do nucleotídeo na parte citoplasmática de LapD promove um aumento da afinidade de seu domínio periplasmático por LapG, sequestrando-a para a membrana interna. Dessa forma, a adesão celular é mantida através da localização de LapA na superfície da bactéria (figura 4).[9]

Quando não ligada a c-di-GMP, LapD assume uma forma auto-inibida onde o domínio EAL no C-terminal interage com uma hélice α (denominada de hélice S) da continuação do domínio HAMP. Nessa conformação o sítio de ligação à c-di-GMP encontra-se parcialmente obstruído na interface de interação e o domínio periplasmático de LapD é incapaz de ligar-se à LapG, e, portanto, essa protease tem acesso e cliva a adesina LapA, liberando-a a superfície celular (Figura 5). Um aumento nos níveis de c-di-GMP intracelular promove um rompimento das interações inibitórias através da ligação do nucleotídeo no domínio EAL, acompanhado provavelmente de grandes mudanças conformacionais que permitem a homodimerização do mesmo. O acoplamento da dimerização e ligação de c-di-GMP pode contribuir adicionalmente para a eficiência da ativação, prevenindo a reversão para o estado auto-inibido. Várias mutações no módulo citoplasmático e no domínio HAMP levam a ativação constitutiva de LapD.[6] Esses dados sugerem que de fato as interações auto-inibitórias intrínsecas são necessárias para que o sistema adote uma conformação ativa constitutivamente. Nesse sentido, a liberação do domínio EAL complexado com c-di-GMP da interação com a hélice S pode relaxar o

receptor, transmitindo as mudanças conformacionais para o domínio periplasmático através do domínio HAMP, resultando no sequestro de LapG (figura 5).[44]

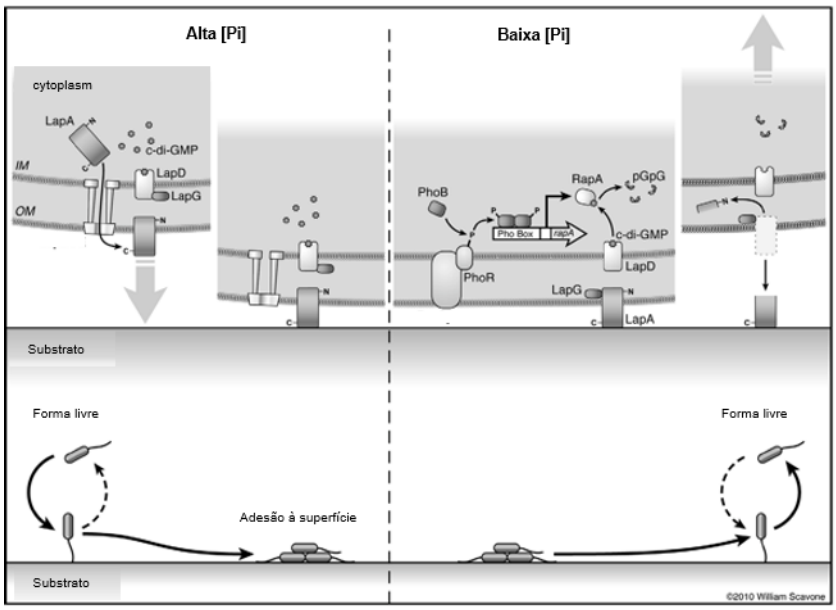


Figura 4 - Modelo de como concentrações de Pi controlam a formação de biofilmes de *P. fluorescens*. Esquerda: Em altas concentrações de Pi, c-di-GMP se acumula nas células. LapD liga-se a c-di-GMP e sequestra a protease LapG na membrana interna, promovendo a manutenção da adesina LapA na superfície celular. Nessas condições, as bactérias se aderem irreversivelmente em um substrato permitindo a progressão do biofilme. Direita: Quando baixas concentrações extracelulares de Pi são sentidas pelo sistema dois-componentes PhoR/Pst, uma cascata de sinalização é disparada resultando na transcrição da fosfodiesterase da família EAL RapA. A atividade de RapA na conversão de c-di-GMP para pGpG, abaixa os níveis intracelulares dessa molécula levando à dissociação de c-di-GMP da LapD. Na forma apo, LapD não consegue interagir com LapG, liberando-a para clivar a parte N-terminal de seu substrato natural no periplasma, LapA, promovendo sua remoção da superfície celular. A liberação de LapA de bactérias no biofilme leva a sua desadesão da superfície, sendo que as bactérias planctônicas sem LapA ficam impedidas de se aderirem novamente à uma superfície.

Fonte: Adaptada de NEWELL.[9]

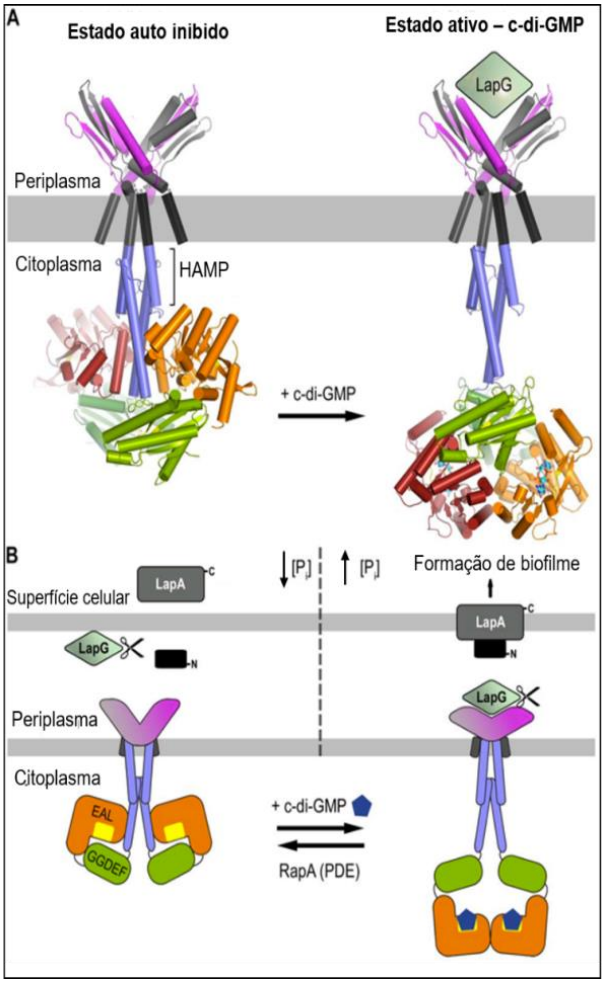


Figura 5 - A) Modelos estruturais para a proteína LapD completa na forma auto-inibida e ativa ligada a c-di-GMP baseados nas estruturas cristalinas e estudos *in vitro* e *in vivo*. Somente a forma ligada com c-di-GMP é capaz de ligar-se a LapG no periplasma. O domínio HAMP foi modelado por homologia com base em alinhamentos sequenciais e informações estruturais disponíveis.[3-4] B) Modelo para o controle da formação de biofilme mediado por LapD. O desenho representa o modelo atual baseado nas análises estruturais e outros resultados. [6-9]

Fonte: Adaptada de NAVARRO.[44]

1.3.2 FimX – *Pseudomonas Aeruginosa*

A FimX de *P. aeruginosa* também apresenta domínios GGDEF-EAL *in tandem* (Figura 6). Tal proteína está envolvida diretamente no comportamento celular de *P. aeruginosa*, sendo um receptor de c-di-GMP de alta afinidade, regulando a motilidade e formação de biofilme.[46-47]

Figura 6 - Representação da predição e disposição dos domínios REC, PAS, GGDEF e EAL da FimX de *Pseudomonas aeruginosa.*
Fonte: Elaborada pelo autor.

Estudos iniciais com esta proteína levaram a duas propostas de mecanismos moleculares para a FimX e suas consequências no funcionamento celular. Em um modelo, o c-di-GMP se ligaria ao domínio EAL da FimX com alta afinidade e formaria um complexo com parceiros de ligação que seriam regulados através da ligação do c-di-GMP ao domínio EAL. Em um modelo alternativo, um outro domínio EAL ou proteína com domínios GGDEF-EAL poderiam formar complexos específicos com a FimX através interações semelhantes aos homodímeros formados pelo domínio EAL da FimX. Neste processo, o heterodímero faria com que a atividade catalítica do domínio EAL (ou GGDEF no caso de dual domains), das proteínas que formaram interações com a FimX, fossem reprimidas (Figura 7).[48,50] Posteriormente, novas pesquisas sobre a FimX confirmaram alguns dos pontos propostos e identificaram que a FimX pode mudar a conformação dos domínios GGDEF-EAL através da interação com a proteína PilZ e c-di-GMP (Figura 7). [54]

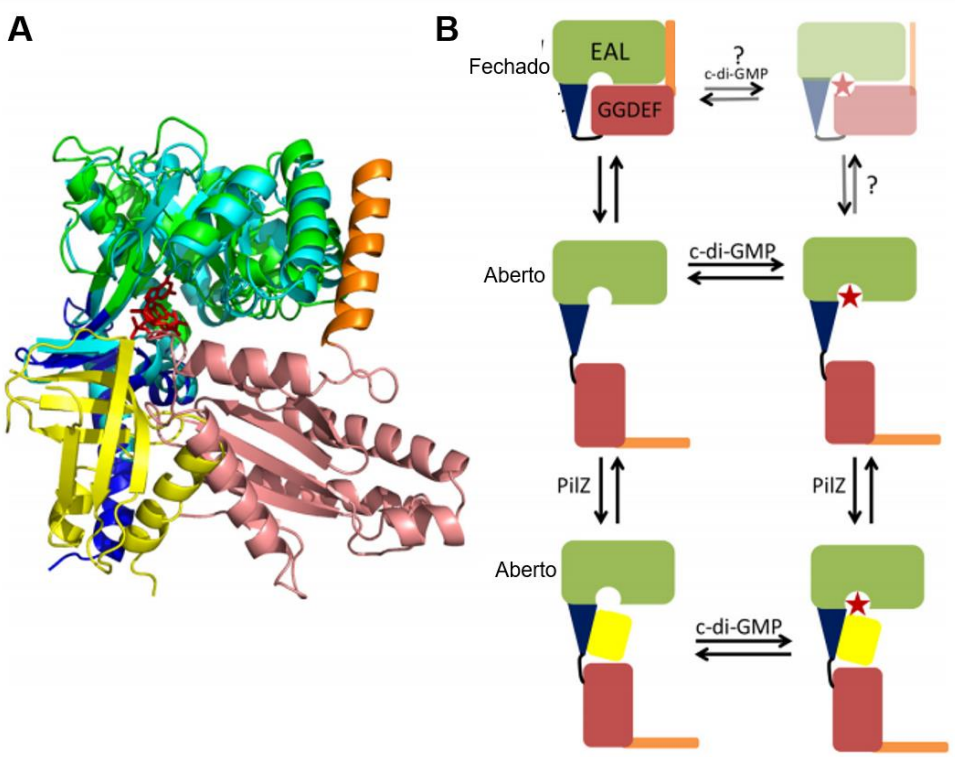


Figura 7 - A) Alinhamento estrutural do complexo PilZ-FimXEAL e LapD GGDEF-EAL apo. As estruturas estão coloridas de forma a identificar as diferentes regiões, sendo a PilZ (amarelo), FimXEAL (verde e azul) e GGDEF (salmão), EAL (ciano) e hélice S (laranja) da LapDGGDEF-EAL. A ligação do c-di-GMP na FimX está disposta na cor vermelha. O barril TIM do domínio EAL da LapD e FimX foram usados como base do alinhamento estrutural. B) A figura ilustra a proposta de mecanismo para as duas possíveis conformações dos domínios GGDEF-EAL da FimX dadas através da interação com PilZ e c-di-GMP. Na linha superior, podemos observar os domínios GGDEF-EAL em conformação fechada na ausência de c-di-GMP e PilZ, de forma similar ao observado com a apo-LapDGGDEF-EAL.[44] Essa conformação é estabilizada pela interação do barril TIM do domínio EAL com a hélice S (superior esquerda), onde na LapDGGDEF-EAL, essa conformação impede a interação da proteína com c-di-GMP. Como os domínios GGDEF-EAL da XacFimX e P. fluorescens LapD apresentam, respectivamente, apenas 25% e 39% de identidade e similaridade de sequência, não foi possível elucidar se o c-di-GMP poderia se ligar à FimX em conformação fechada (como indicada na representação transparente – superior direita). Nas figuras na região central e abaixo, observamos que os domínios GGDEF e EAL podem se dissociar, passando então para a conformação aberta, que permite a ligação de c-di-GMP (estrela vermelha) e também da PilZ.[50]

Fonte: Adaptada de GUZZO.[50]

1.3.3 MorA – *Pseudomonas Aeruginosa*

Phippen e colaboradores determinaram a estrutura cristalográfica dos domínios GGDEF-EAL proteína reguladora de motilidade A (MorA), conservada em espécies de Pseudomonas.[45] Além dos domínios GGDEF e EAL, a porção

citoplasmática contém quatro domínios sensores PAS (*Per-Arnt-SIM homology*)[51], como mostra a predição de domínios na figura 8. A proteína é localizada na membrana com duas hélices transmembranares no N-terminal, podendo afetar a motilidade da bactéria[45], além de estar relacionada com formação de flagelos e fímbrias, que são requisitos necessários para dispersão e formação de biofilme bacteriano[52-53]. Os motivos GGDEF e EAL estão conservados para a MorA, indicando atividades de fosfodiesterase e diguanilato ciclases na mesma proteína.[10]

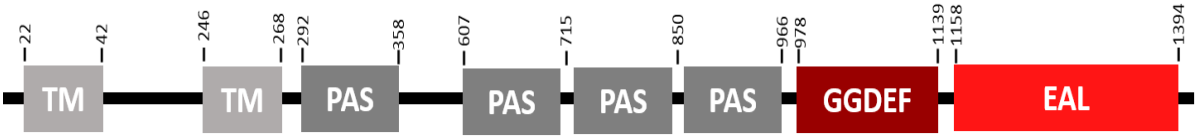


Figura 8 – Representação Da predição e disposição das hélices transmembranares e domínios PAS, GGDEF e EAL da MorA de *Pseudomonas aeruginosa.* TM: Hélice transmembranar.

Fonte: Elaborada pelo autor.

Neste mesmo estudo, Phippen e colaboradores observaram que a MorA (porção citoplasmática) encontra-se no estado monomérico na ausência de c-di-GMP e dimérico na presença do nucleotídeo. Observando a estrutura cristalográfica, precisamente no domínio EAL, notamos que a dimerização proporciona mudanças estruturais na hélice α5, de forma com que quando no monômero, esta hélice encontra-se mais estendida que no dímero, e um loop maior nesta região permite que os resíduos catalíticos e aqueles que coordenam o c-di-GMP no sítio de ligação estejam em posição adequada para receber o nucleotídeo (Figura 9).

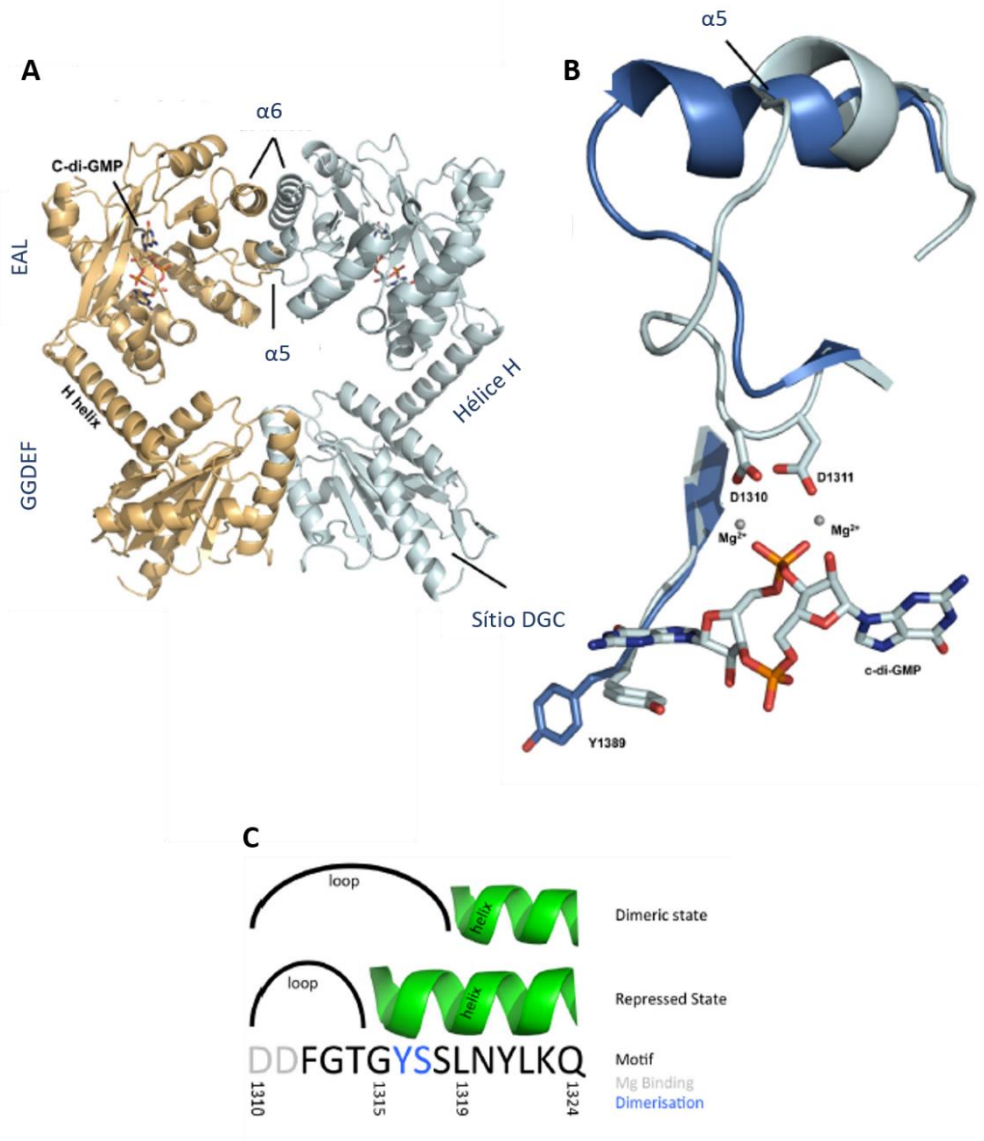


Figura 9 – Estrutura Cristalográfica dos domínios GGDEF-EAL da MorA. A) Na presença de c-di-GMP, a estrutura da MorA apresenta a formação de um dímero. B) Com a sobreposição das estruturas nos dois estados oligoméricos, monômero (azul claro) na ausência de c-di-GMP e dímero (azul escuro) na presença de c-di-GMP. A hélice repressora, α5 destacada também em (A), muda a orientação e tamanho (C) de acordo com o estado oligomérico da enzima, afetando estruturalmente a posição dos aspárticos que coordenam a ligação do Mg^{2+}, onde no monômero não há atividade pois os aspárticos ficam afastados do sítio ativo, porém, a extensão do *loop* na forma dimérica permite que os resíduos se posicionem de forma a coordenar os metais e o substrato de forma adequada para promover a hidrólise.[10]

Fonte: Adaptada de PHIPPEN. [10]

1.3.4 STM3615 – *Salmonella enterica*

Outro caso interessante refere-se ao organismo *Salmonella entérica serovar Typhimurium*. Análises genômicas mostram que esse organismo possui apenas uma

proteína transmembranar com a mesma arquitetura da LapD, a STM3615, no entanto seu domínio periplasmático é totalmente diferente e nenhum homólogo de LapG foi encontrado e seu domínio EAL possui todos dos resíduos essenciais para catálise conservados.

Ahmad e colaboradores mostraram que cepas mutantes, com a STM3615 deletada, foram encontradas severamente atenuadas na colonização do intestino em camundongos tratados com estreptomicina (Figura 10).[5] Além disso, essa proteína foi identificada recentemente como uma das grandes responsáveis por regular a formação do biofilme bacteriano em *Salmonella*. Anwar e colaboradores demonstraram *in vivo* que a deleção das proteínas dsbA e dsbB (dissulfeto oxidase periplasmática e dissulfeto oxidoredutase citoplasmática, respectivamente) causam aumento na formação de biofilme, mas curiosamente, quando a proteína STM3615 é deletada, os efeitos causados pela deleção das dsb são anulados. Esses pontos demonstram a possibilidade de variações no sistema estudado de *P. fluorescens* e levanta a hipótese de diferentes funcionalidades e parceiros de interação para os domínios periplasmáticos dessa classe de proteínas.[49]

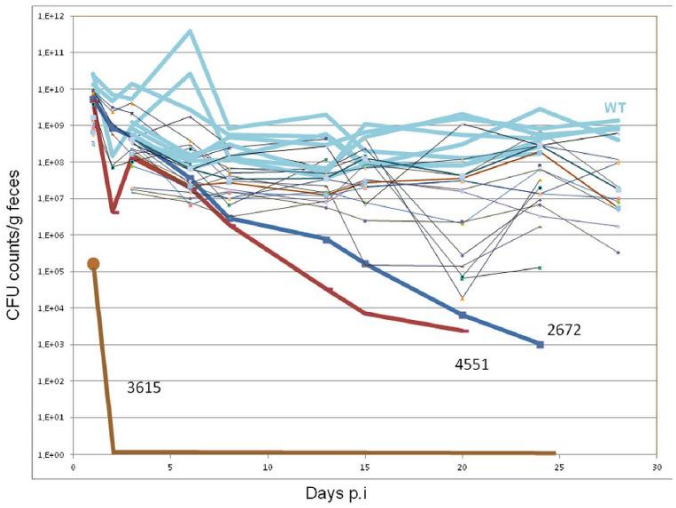


Figura 10 – Presença de mutantes de GG(D/E)EF/EAL de *Salmonella typhimurium* UMR1 no bolo fecal de ratos tratados com estreptomicina. Ratos tratados com estreptomicina foram infectados com quatro grupos de cepas (selvagem e três mutantes de GG(D/E)EF/EAL). As variações dos mutantes no intestino dos ratos foram avaliadas pela contagem de unidades de colônias formadas (CFU) por grama de fezes até 30 dias depois da infecção. Três dos vinte mutantes (STM2672, STM3615 e STM4551) sumiram de forma mais rápida do intestino dos camundongos.[5]

Fonte: AHMAD. et al. [5]

É importante também destacar que várias outras sequências contendo a mesma arquitetura possuem um alto grau de conservação dos resíduos putativos dos domínios EAL e/ou GGDEF necessários para atividade enzimática. Dessa forma, podemos especular que mecanismos semelhantes ao descrito para LapD possam também ser aplicáveis nesses sistemas, nos quais mudanças conformacionais causadas por interações proteicas no periplasma possam ser transmitidas através do domínio HAMP para os domínios catalíticos EAL ou GGDEF em um mecanismo de sinalização de “fora-para-dentro”, regulando a atividade fosfodiesterase ou diguanilato ciclase, e portanto, levando a mudanças nos níveis de c-di-GMP intracelular.

2 OBJETIVOS

Como objetivo geral, este trabalho buscou desvendar os mecanismos moleculares da proteína STM3615 de *Salmonella enterica serovar thyphimurium*, com a finalidade de compreender seu papel na formação de biofilme em *Salmonella* e outras bactérias. Além da importância fisiológica já observada dessa proteína em *S. enterica*, sua arquitetura de domínios contendo um GGDEF degenerado e um EAL conservado (potencialmente ativo) provê um modelo ideal para o estudo de um dos tipos mais representativos e ainda pouco conhecido de proteínas *dual domain.*

Para isso, os objetivos específicos foram:

- Clonagem das regiões codificantes da STM3615 e de seus diferentes domínios solúveis isolados, a partir do DNA genômico de *Salmonella enterica;*

- Determinação de protocolos de expressão e purificação das diferentes construções obtidas;

- Ensaios enzimáticos para determinação de atividade e produtos de hidrólise;

- Estudos bioquímicos e biofísicos em solução para determinação dos estados oligoméricos dos diferentes domínios da STM3615;

- Ensaios cristalográficos envolvendo diversas construções dos domínios solúveis da STM3615;

- Propostas de mecanismo para a STM3516.

- Análises de sequências de aminoácidos estruturas para elucidação de proteínas que apresentam domínios GGDEF-EAL.

3 METODOLOGIA

3.1 Clonagem das construções da STM3615

Para desenhar os primers necessários para amplificar a região codificante do gene de interesse e obtermos construções adequadas para a realização do trabalho, buscamos algumas informações relevantes sobre a STM3615, baseadas na predição de hélices hidrofóbicas (servidor TMHMM[54]) (Figura 11), pela delimitação dos domínios (servidor PFAM[55]), bem como uma análise computacional das estruturas secundárias da proteína (servidor PSIPRED[56]) (Figuras 12 e 13).

.

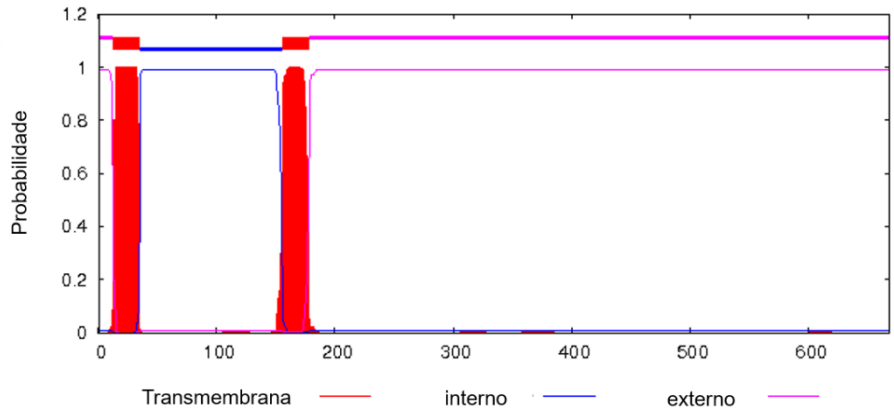
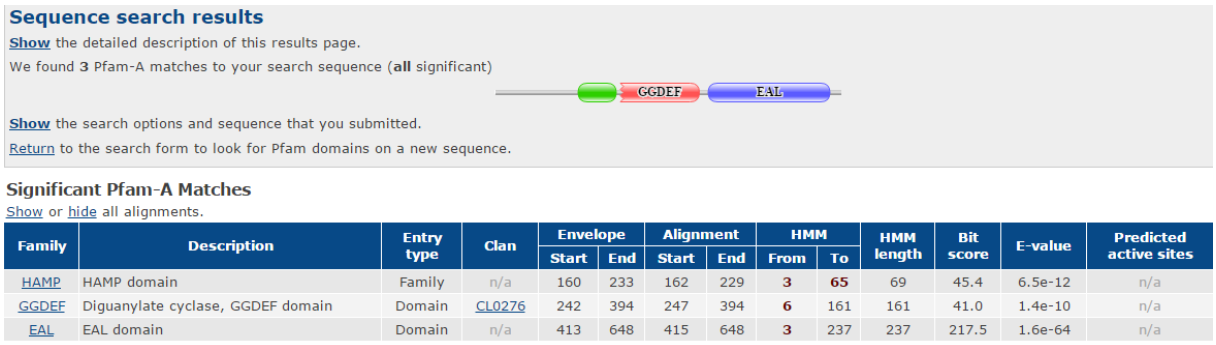


Figura 11 – Predição de hélices transmembranares da STM3615 realizada pelo servidor TMHMM. As regiões vermelhas demonstram a alta probabilidade de regiões transmembranares no N-terminal (peptídeo sinal) e logo antes do domíno HAMP.

Fonte: Elaborada pelo autor



Sequence search results

Show the detailed description of this results page.

We found **3** Pfam-A matches to your search sequence (**all** significant)

Show the search options and sequence that you submitted.

Return to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches

Show or hide all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites
				Start	End	Start	End	From	To				
HAMP	HAMP domain	Family	n/a	160	233	162	229	**3**	**65**	69	45.4	6.5e-12	n/a
GGDEF	Diguanylate cyclase, GGDEF domain	Domain	CL0276	242	394	247	394	**6**	161	161	41.0	1.4e-10	n/a
EAL	EAL domain	Domain	n/a	413	648	415	648	**3**	237	237	217.5	1.6e-64	n/a

Figura 12 – Análise da delimitação dos domínios HAMP (verde), GGDEF (vermelho) e EAL (azul) realizada pelo servidor PFAM.

Fonte: Elaborada pelo autor.

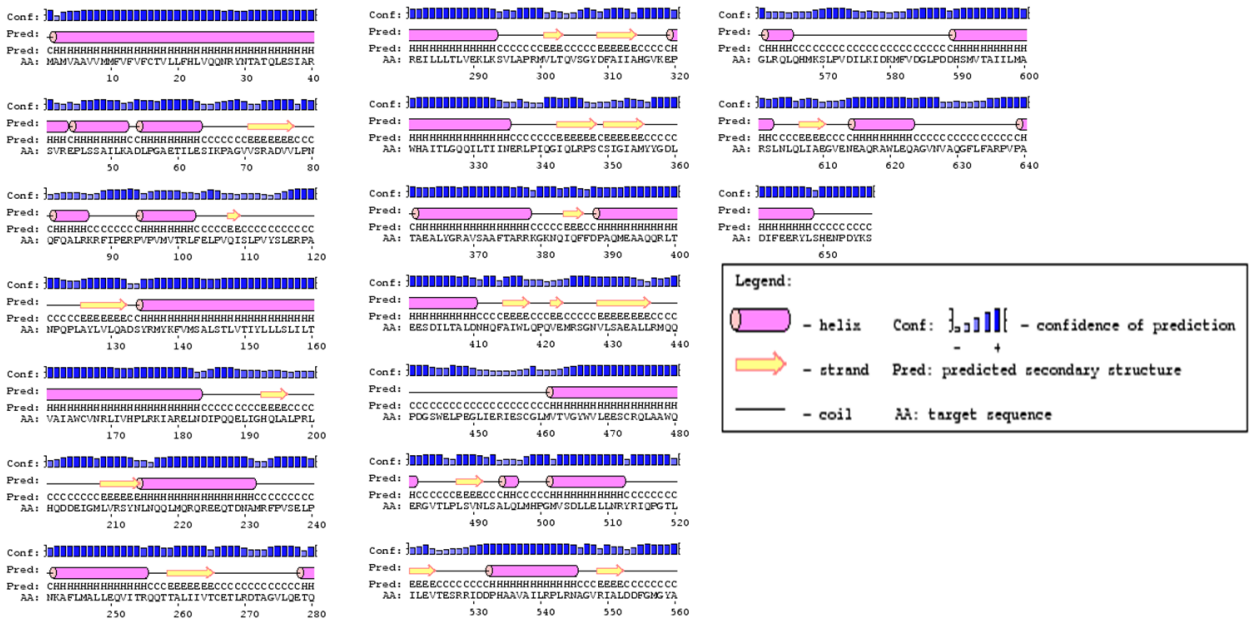


Figura 13 - Análise computacional para predição de estruturas secundárias da STM3615 pelo servidor PSIPRED.

Fonte: Elaborada pelo autor.

Segundo as predições realizadas, observamos portanto a presença duas hélices transmembrana (TM) entre o domínio periplasmático (DP), seguido por três domínios citoplasmáticos: HAMP, GGDEF e EAL (Figura 14).

F

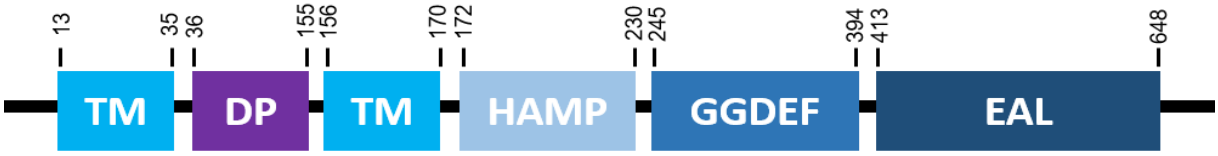


Figura 14 - Representação da predição e disposição dos domínios periplasmático, HAMP, GGDEF, EAL e hélices transmembranares da STM3615 de Salmonella enterica. TM: Hélice Transmembranar, DP: domínio periplasmático.

Fonte: Elaborada pelo autor.

Considerando as duas hélices transmembranares e os quatro domínios da STM3615, uma série de oligonucleotídeos foram desenhados para amplificação por PCR a partir do gene completo, posteriormente clonados em vetor de propagação (pGEM) e transferidos para vetores de expressão. As informações da predição de estrutura secundária foram utilizadas para determinas os limites das diferentes

construções de domínios isolados (tabela 1), as quais foram clonadas para os estudos em solução desse trabalho, mas também visando aumentar a probabilidade de cristalização.

O DNA genômico, cedido gentilmente pelo Prof. Michael Hensel (Universitaet Osnabrueck, Alemanha), foi utilizado para isolar as regiões codificantes de interesse da STM3615. A amplificação foi feita por PCR utilizando primes específicos *forward* e *reverse* . Visando a sub-clonagem no vetor de expressão pET28a ou pSMT3, cuja proteína de interesse é expressa em fusão com a proteína modificada 6xHis-SUMO (proteína Small Ubiquitin Modifier de levedura, com uma cauda de 6xHis no N-Terminal).[57] sítios de restrição para as endonucleases BamHI, XhoI e NdeI foram introduzidos nos oligonucleotídeos. Clones recombinantes contendo os vetores pSMT3_DP, pSMT3_HAMP-GGDEF-EAL e pSMT3_GGDEF-EAL, pSMT3_EAL, pET28_HAMP-GGDEF e pET28_GGDEF foram selecionados e confirmados por sequenciamento. Células de *E. coli* BL21 (DE3) foram transformadas com os plasmídeos de cada uma das construções confirmadas e os testes de expressão mostraram bandas compatíveis com a massa esperada das proteínas em SDS-PAGE.

Para as reações de PCR foram utilizados 200 ng de DNA genômico de *Salmonela enterica serovar Typhimurium LT2* como molde e 100 pmol de cada primer, nos sentidos *forward* e *reverse,* específicos para as construções da STM3615 em reações de 50 µl, utilizando um mix de enzimas com alta fidelidade (Fermentas - *High Fidelity*) conforme protocolo do fabricante (Tabela 2).

Os produtos de PCR foram purificados por eletroforese em gel de agarose 0,8% utilizando o kit "*Wizard SV Gel and PCR Clean-Up System"* (Promega). As amostras purificadas foram submetidas a dupla digestão com as enzimas de restrição e após uma nova purificação, os insertos foram subclonados no vetor de expressão pET28a ou pSMT3. As construções foram clonadas nesse vetor utilizando protocolos convencionais de ligação por reação usando T4 DNA ligase (relação molar inserto:vetor = 3:1 ou 6:1). As reações de ligação foram transformadas em células de *Escherichia coli* DH5α competentes por tratamento com cloreto de cálcio para fins de propagação e extração de DNA plasmidial. As células transformadas foram plaqueadas em meio Lúria-Bertani (LB) contendo canamicina e os clones

positivos foram confirmados por dupla digestão do DNA plasmidial, examinados por eletroforese em gel de agarose 0,8% e sequenciados (3130 Genetic Analyzer - Applied Byosystems).

Tabela 1 - Construções da STM3615 do domínio periplasmático (DP), HAMP, GGDEF e EAL, massa molecular teórica e vetores de expressão.

Construção	Resíduo Inicial	Resíduo Final	Massa Molecular (kDa)	Vetor de expressão
DP	36	145	12,1	pSMT3
DP	36	150	12,8	pSMT3
DP	36	155	13,3	pSMT3
DP	41	145	11,5	pSMT3
DP	41	150	12,2	pSMT3
DP	41	155	12,8	pSMT3
DP	46	145	10,8	pSMT3
DP	46	150	11,6	pSMT3
DP	46	155	12,2	pSMT3
HAMP-GGDEF-EAL	179	668	53,6	pSMT3
HAMP-GGDEF-EAL	184	668	54,7	pSMT3
GGDEF-EAL	218	668	50,6	pSMT3
GGDEF-EAL	242	668	47,7	pSMT3
GGDEF-EAL	247	668	47,2	pSMT3
GGDEF-EAL	252	668	46,6	pSMT3
EAL	392	668	31,1	pSMT3
HAMP-GGDEF	181	388	26,1	pET28a
GGDEF	242	388	16,8	pET28a

Fonte: Elaborada pelo autor.

Tabela 2 – Etapas de reação de PCR para amplificação dos fragmentos de DNA de interesse.

Etapa do PCR	Desnaturação Inicial	Desnaturação	Hibridização	Extensão	Extensão Final
Temperatura / Tempo	95°C / 3 min	95°C / 30 seg	60°C / 1 min	72°C/ 2 min	72°C/10min

Fonte: Elaborada pelo autor.

3.2 Expressão e purificação das construções da STM3615

Os vetores de expressão contendo as construções de interesse, foram transformados em *E. coli* BL21(DE3) (Novagen) e plaqueadas em ágar LB. Colônias isoladas foram inoculadas em meio LB líquido a 37°C, contendo 50 µg/ml de canamicina, nos quais foram realizados ensaios de expressão recombinante através da indução com 0,1 mM de IPTG, adicionado em $D.O._{600nm}$=0,6. As células foram incubadas sob agitação constante a 150 rpm, *overnight* a 18°C, e ao final da

indução, isoladas por centrifugação, ressuspendidas em tampão 20 mM Tris-HCl pH 8,0 contendo 300 mM de NaCl e lisadas sob banho de gelo por 20 pulsos de ultra-som de 30 segundos com amplitude de 20% e intervalos de descanso de 40 segundos. A fração solúvel, isolada por centrifugação, foi filtrada e aplicada em resina TALON® (Clontech) previamente equilibrada no mesmo tampão, para imobilização por afinidade das proteínas produzidas em fusão com o 6xHis SUMO na extremidade amino-terminal com o cobalto. A coluna foi então lavada com 20 volumes de coluna do mesmo tampão com adição de 20 mM de imidazol, para eliminação dos contaminantes que tenham se ligado por interações inespecíficas. A proteína de fusão foi então eluída da coluna com um gradiente crescente de imidazol no tampão de lise. Para clivagem do 6xHis SUMO, a proteína foi incubada com uma protease específica, ULP1, *overnight* a 10 °C e isolada após uma segunda cromatografia por afinidade em resina TALON. O imidazol foi removido previamente utilizando uma coluna HiTrap Desalting (GE Healthcare). Dessa forma, a proteína de interesse clivada não se liga à coluna devido a ausência da cauda de histidina, enquanto que a 6xHis SUMO permanece retida.

Por fim uma etapa adicional de purificação é realizada através de cromatografia de exclusão molecular (HiLoad 16/60 Superdex 75) para assegurar alto grau de pureza e eliminação de possíveis agregados solúveis, possibilitando o início dos ensaios experimentais. A qualidade da purificação e do protocolo foi aferida por análises em SDS-PAGE corados com *Comassie blue*.

3.2.1 Cromatografia de exclusão molecular analítica

Para determinação do estado oligomérico das construções HAMP-GGDEF-EAL$^{184-668}$, HAMP-GGDEF$^{181-410}$, GGDEF-EAL$^{242-668}$, GGDEF$^{242-399}$ e EAL$^{392-668}$ realizamos ensaios de exclusão molecular analítica utilizando o Kit *Gel Filtration Calibration* (GE Helthcare), seguindo suas instruções para a calibração da coluna de exclusão molecular Superdex75 10/300 (GE Helthcare). Calibramos a coluna com um conjunto de proteínas recomendadas pelo kit, cobrindo de 6,5 kDa a 75 kDa (Aprotinina 6.5 kDa, Ribonuclease 13.7 kDa, Anidrase carbônica 29.0 kDa, Ovoalbumina 44.0 kDa e Conalbumina 75.0 kDa), além do Blue Dextran, utilizado para cálculo do *void* da coluna. Com a coluna calibrada, realizamos corridas com

cada uma das construções, verificando o volume de eluição de cada proteína no máximo de absorção para o cálculo da massa molecular através da curva de calibração. Em cada cromatografia foi aplicado 100 µL de proteína (1-2 mg/mL) para uma corrida com fluxo contínuo de 0,5 mL/min em tampão 20 mM Tris-HCl pH 8,0 contendo 300 mM de NaCl.

3.2.2 Espalhamento de raios-X a baixo ângulo (SAXS)

Com a finalidade de comparar os resultados obtidos por cromatografia de exclusão molecular, a técnica de espalhamento de raios-X a baixo ângulo também foi utilizada para determinar o estado oligomérico dos domínios citoplasmáticos em solução. Adicionalmente, esses resultados foram utilizados para determinar alguns modelos tridimensionais de baixa resolução das construções de STM3615 isoladas.

Diversas concentrações de cada construção foram utilizadas para a coleta de dados de SAXS na linha D02A-SAXS2 do Laboratório Nacional de Luz Síncrotron (LNLS) utilizando comprimento de onda λ = 1.488 Å, detector bidimensional e distância amostra-detector de 1028.4 mm. Levando-se em conta o ruído gerado pelo CCD, intensidade do feixe, absorção da amostra e espalhamento da solução tampão de cada uma das diferentes amostras, foram realizadas análises sistemáticas das imagens obtidas e os padrões bidimensionais foram integrados para obtenção das curvas de espalhamento.

O raio de giro (R_g) foi obtido utilizando a curva de Guinier,[58] pelo programa AutoRg[59] e através do método de transformada inversa de Fourier implementado no programa Gnom,[60] onde também obtive-se a função de distribuição de distâncias P(r) e diâmetro máximo (Dmax) do envelope.

A determinação da massa molecular e do estado oligomérico das proteínas foram calculadas por três métodos: proteínas padrão (lisozima e BSA),[61-62] programa SAXSMoW[63] e curvas em escala absoluta.[60]

Para cada construção diferente, 25 estruturas de baixa resolução foram obtidas a partir dos dados experimentais através do programa Dammin.[64] O

programa Damsel[59] foi utilizado para selecionar os modelos similares e excluir os discrepantes. Os modelos selecionados pelo Damsel foram alinhados pelo programa Damsup[59] para então gerar um modelo médio através do programa Damaver.[59]

Sobreposição dos envelopes de SAXS com estruturas cristalográficas foram realizadas com o programa Massha.[65] Os valores de Rg, Dmax, a curva de espalhamento simulado das coordenadas atômicas e as discrepâncias entre os modelos propostos e os dados experimentais foram calculados utilizando o programa Crysol.[66]

3.3 Ensaios de atividade enzimática

Para o ensaio enzimático com finalidade de verificar a atividade dos domínios GGDEF e EAL nas diferentes construções, realizamos ensaios contendo como substrato 100 µM de c-di-GMP ou GTP. Os substratos foram incubados com 10 µM de cada construção em tampão 50 mM Tris pH8.0, 300mM NaCl, 5 mM $MgCl_2$ por diferentes tempos (10-120 min). As reações foram paradas através da desnaturação da enzima a 100 °C por 5 minutos. O precipitado de proteínas foi removido por centrifugação (16000 g por 10 minutos). Para detecção do substrato c-di-GMP e do produto pGpG, o sobrenadante foi submetido a cromatografia de fase reversa em coluna C18 acoplada ao HPLC (Waters Alliance). Nessa análise utilizamos um tampão base contendo 100 mM de Acetato de Amônio pH 5.5 e um tampão apolar, contendo 30% do tampão base e 70% de metanol. No HPLC, realizamos, para cada amostra aplicada, um gradiente com 70% de metanol buscando isolar os componentes das amostras.[12,67] Os picos identificados foram submetidos a análises de espectrometria de massas para identificação dos produtos.

3.4 Cristalização e estrutura de alta resolução dos domínios GGDEF-EAL

3.4.1 Cristalização

Para que o modelo de reconhecimento de c-di-GMP da STM3615 seja desvendado, iniciamos os ensaios de cristalização com a finalidade de resolver a estrutura atômica dos domínios solúveis da STM3615. As triagens iniciais de cristalização foram feitas pelos métodos padrões de gotas sentadas (*sitting drop*) e suspensas (*hanging drop*), a 18 °C, utilizando soluções precipitantes comerciais disponíveis com os kits Index, Crystal Screen Cryo, PEG/Íon, Memb Fac e PEGRx, da empresa Hampton Research, totalizando 480 condições. Também utilizamos o kit Morpheus da empresa Molecular Dimensions [68] que apresenta 96 soluções que baseiam-se em formulações de acordo com análises dos componentes que se apresentam com maior frequência no PDB.[69]

As soluções precipitantes que apresentaram formações cristalinas, tiveram as condições refinadas, variando principalmente o pH, concentração do agente precipitante e concentração da proteína, buscando melhorias na formação e crescimento dos cristais, a fim de torná-los mais adequados para experimentos de difração de raios X.

Como a identidade das proteínas depositadas no PDB em relação à construção GGDEF_EAL$^{242\text{-}668}$ são baixas (no máximo 35%) e não permitiram resolver a estrutura da proteína através de substituição molecular, realizamos dois métodos para resolver a estrutura da proteína por dispersão anômala em um comprimento de onda (SAD), um deles baseado na incorporação de selenometionina e outro através de *soaking* com solução idêntica à que ele foi cristalizado, contendo 500 mM iodeto de sódio.

Para garantir a incorporação da selenometionina, as condições gerais de expressão foram as mesmas utilizadas para a proteína nativa, porém utilizamos o meio de cultura mínimo M9. Após atingir a D.O. = 0,6 adicionamos um mix de aminoácidos contendo 100 mg de L-lisina, 100 mg de L-fenilalanina, 100mg de Treonina, 50 mg de L-isoleucina, 50 mg de L-Leucina, 50 mg de L-Valina e 60 mg de

L-selenometionina, em seguida a adição dos aminoácidos, aguardamos cerca de 15 minutos para homogeneização e posteriormente adicionamos 0,1 mM de IPTG.

Previamente a coleta de dados, os cristais foram colocados em uma solução crioprotetora com a mesma composição da solução em que o cristal foi formado, contendo 20% de etilenoglicol ou glicerol. Os cristais obtidos foram submetidos para difração de raios X e tiveram seus dados coletados na linha MX-2 no Laboratório Nacional de Luz Síncrotron (LNLS, Campinas – SP).[70] Todos os conjuntos de dados foram processados com o programa *X-ray Detector Software* (XDS).[71]

3.4.2 Estrutura cristalográfica dos domínios GGDEF-EAL

Para a resolução das estruturas utilizando sinal anômalo de Selênio, os sítios de átomos pesados foram encontrados através do programa ShelXD[72] e a estrutura foi faseada com o programa *Python-based Hierarchical Environment for Integrated Xtallography* (PHENIX).[73] As estruturas finais foram construídas manualmente e refinadas com o programa *Crystallographic Object-Oriented Toolkit* (Coot)[74] e refinadas com o programa PHENIX.

3.5 Análise de sequência, atividade e estrutura secundária de proteínas dual domain GGDEF-EAL.

Com o intuito de generalizar as análises feitas a partir da estrutura dos domínios GGDEF-EAL de STM3615 e comparações com estruturas das proteínas LapD de *P. fluorescens* e MorA de *P. aeruginosa* fizemos uma seleção de 150 sequencias representativas de proteínas contendo GGDEF-EAL em bancos de dados. Os critérios para seleção das sequencias, bem como os tipos de análises realizados são descritos a seguir.

3.5.1 Seleção de sequências

Para selecionar um número significativo de sequências, utilizamos o banco de dados disponibilizados por Galperin e colaboradores[42,75-76] que possui uma versão atualizada disponível na internet (http://www.ncbi.nlm.nih.gov/Complete_Genomes/SignalCensus.html). Com base nestes dados, selecionamos 16 espécies diversas do filo proteobactéria (alfa, beta, gama, delta e épsilon). O número de espécies selecionadas em cada sub-divisão seguiu a proporção geral das mesmas, seguindo como critério para seleção a presença de no mínimo 7 sequencias GGDEF-EAL codificada pelos genomas. No total, foram selecionadas 164 sequencias dos seguintes micro-organismos: *Caulobacter crescentus, Rhodobacter sphaeroides, Bukhoderia vietnamiensis, Nitrosospira multiformis, Cellvrio japonicus, Escherichia coli, Pseudomonas aeruginosa, Salmonella enterica serovar typhimuium LT2, Vibrio Cholerae, Xanthomonas axonopodis, Geobacter sulfurre*ducens, *Arcobacter butzleri, Photobacterium profundum 559, Enterobacter sp 638, Halornodospira holophila SL1* e *Nitrosococcus oceani ATCC19707.*

.

3.5.2 Alinhamento e predição de atividade catalítica e estrutura secundária

As sequências selecionadas foram observadas individualmente quanto aos seus domínios preditos pelo *Protein Families database* (Pfam)[55] para verificar se todas estavam de fato dentro do padrão imposto, ficando com 150 sequências ao final da avaliação. Definidas as sequências e a delimitação dos domínios das proteínas, essas foram divididas em três partes, definindo três grupos. O primeiro apenas com os resíduos do domínio GGDEF, o segundo composto apenas pelos resíduos do domínio EAL e o terceiro com os resíduos da hélice que conecta od domínios GGDEF e EAL. Como no terceiro grupo temos uma sequência relativamente pequena, para garantir a observação posterior da hélice em questão adicionamos 30 resíduos antes e 30 depois do predito pelo Pfam.[55]

Os alinhamentos das sequências foram feitos para cada um dos grupos, onde verificamos cada um dos resíduos fundamentais para a atividade enzimática,[42] nos casos dos domínios GGDEF e EAL, e também para identificar corretamente as sequências das hélices entre os domínios.

Identificadas as sequências das hélices, utilizamos o servidor Psipred[56] para obter a predição de estruturas secundárias de cada uma delas e comparamos individualmente com o alinhamento feito para esse grupo e finalmente verificar o número de resíduos que compõe a hélice de cada uma das proteínas.

4 Resultados

4.1 Clonagem, expressão e purificação das construções da STM3615

Os genes de cada uma das construções da STM3615 (conforme tabela 1) foram amplificados por PCR e após digestão com enzimas de restrição adequadas, foram subclonados no vetor de expressão com sucesso. Clones recombinantes contendo os vetores pSMT3_DP, pSMT3_HAMP-GGDEF-EAL e pSMT3_GGDEF-EAL, pSMT3_EAL, pET28_GGDEF, pET28_HAMP-GGDEF foram selecionados e confirmados por análises de restrição e por sequenciamento do DNA (Figura 15).

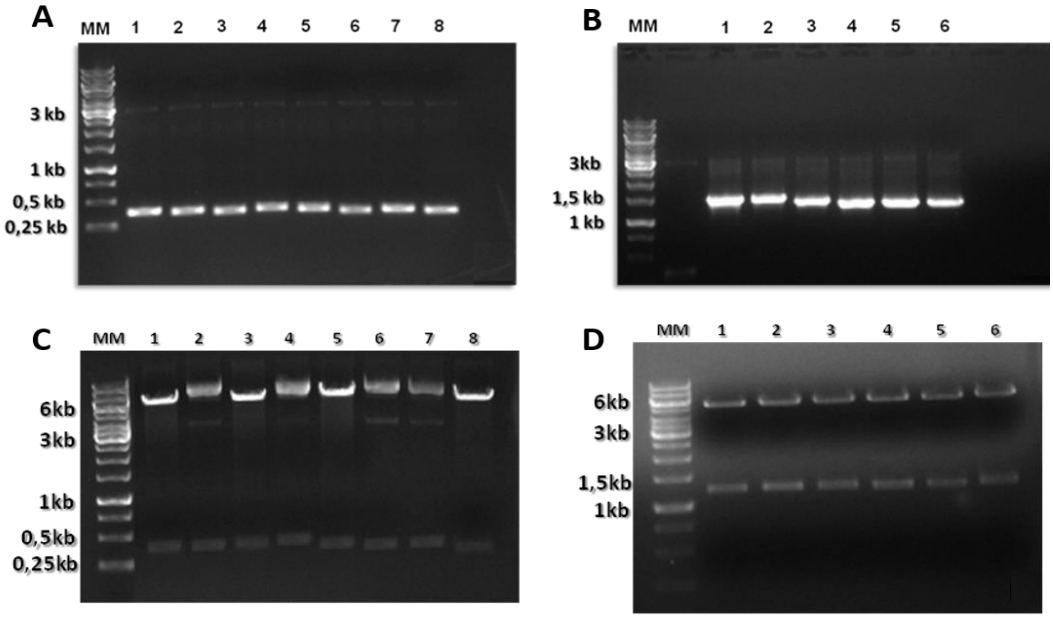
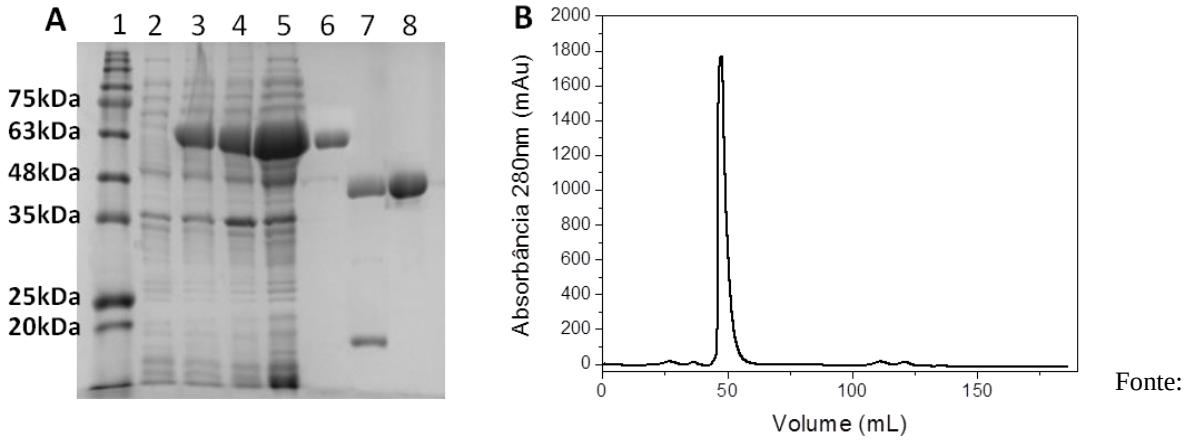


Figura 15 - Clonagem das construções da STM3615: Análise eletroforética do produto de PCR e de restrição em gel de agarose 0,8%. (A) e (C) MM) Padrão de massa molecular, 1) DP^{36-145} (327 Kb); 2) DP^{36-150} (342 Kb); 3) DP^{36-155} (357 Kb); 4) DP^{41-150} (312Kb); 5) DP^{41-155} (327Kb); 6) DP^{46-145} (342 Kb); 7) DP^{46-150} (297 Kb); 8) DP^{46-145} (312Kb). (B) e (D) MM) Padrão de massa molecular, 1) HAMP-GGDEF-$EAL^{168-657}$ (1467 Kb); 2) HAMP-GGDEF-$EAL^{173-657}$ (1452 Kb); 3) GGDEF-$EAL^{207-657}$ (1350 Kb) 4) GGDEF-$EAL^{231-657}$ (1278 Kb); 5) GGDEF-$EAL^{236-657}$ (1263 Kb); 6) GGDEF-$EAL^{241-657}$ (1248 Kb).

Fonte: Elaborada pelo autor.

A expressão e purificação das construções dos domínios citoplasmático e algumas do domínio periplasmático foram realizadas com sucesso, obtendo um

protocolo de expressão e purificação satisfatório para realização dos ensaios previstos, como mostra a tabela 3. A figura 16 ilustra alguns dos passos do protocolo de purificação dessas construções, com o respectivo, incluindo o perfil de eluição da etapa de cromatografia por exclusão molecular.



Fonte:

Figura 16 – Purificação da construção GGDEF_EAL$^{242-668}$. (A) Gel SDS-PAGE 15%, onde podemos observar: (A) (1) Marcador molecular, (2) Pré-Indução, (3) Induzido, (4) *Pellet*, (5) Sobrenadante, (6) Eluição da coluna de afinidade, (7) Clivagem com UPL1 e (8) GGDEF_EAL$^{231-657}$ após cromatografia de exclusão molecular. (B) Cromatograma de exclusão molecular da GGDEF_EAL$^{242-668}$, com a coluna Superdex75 16/60.

Fonte: Elaborada pelo autor.

Através da tabela 3, podemos justificar o grande número de construções dos diferentes domínios que foram clonadas. Todos os domínios citoplasmáticos foram obtidos na fração solúvel, mas obtivemos cristais em apenas dois deles. Para as construções do domínio periplasmático, apenas as construções que começam no resíduo 46 foram obtidas na fração solúvel, com exceção da DP$^{41-155}$. Provavelmente as demais construções se mostraram insolúveis por ainda apresentarem resíduos da região transmembrana da STM3615.

Tabela 3 – Resumo dos resultados obtidos para clonagem, expressão, purificação e cristalização das construções da STM3615.

Construção	Expressão	Solubilidade	Purificação	Cristalização
DP$^{36\text{-}145}$	Sim	Insolúvel	-	-
DP$^{36\text{-}150}$	Sim	Insolúvel	-	-
DP$^{36\text{-}155}$	Sim	Insolúvel	-	-
DP$^{41\text{-}145}$	Sim	Insolúvel	-	-
DP$^{41\text{-}150}$	Sim	Insolúvel	-	-
DP$^{41\text{-}155}$	Sim	Solúvel	Sim	-
DP$^{46\text{-}145}$	Sim	Solúvel	Sim	-
DP$^{46\text{-}150}$	Sim	Solúvel	Sim	Sim
DP$^{46\text{-}155}$	Sim	Solúvel	Sim	-
HAMP-GGDEF-EAL$^{179\text{-}668}$	Sim	Solúvel	Sim	-
HAMP-GGDEF-EAL$^{184\text{-}668}$	Sim	Solúvel	Sim	-
GGDEF-EAL$^{218\text{-}668}$	Sim	Solúvel	Sim	Sim
GGDEF-EAL$^{242\text{-}668}$	Sim	Solúvel	Sim	Sim
GGDEF-EAL$^{247\text{-}668}$	Sim	Solúvel	Sim	-
GGDEF-EAL$^{252\text{-}668}$	Sim	Solúvel	Sim	-
HAMP-GGDEF$^{181—388}$	Sim	Solúvel	Sim	-
GGDEF$^{252\text{-}388}$	Sim	Solúvel	Sim	-
EAL$^{392\text{-}668}$	Sim	Solúvel	Sim	-

Fonte: Elaborada pelo autor.

4.2 Estado Oligomérico dos domínios citoplasmáticos

4.2.1 Cromatografia de exclusão molecular analítica

Como primeiro passo para verificar o estado oligomérico dos domínios citoplasmáticos da STM3615, analisamos as amostras por cromatografia de exclusão molecular analítica (SEC) (Figura 17A), visando estimar a massa molecular das proteínas em solução (Tabela 4).

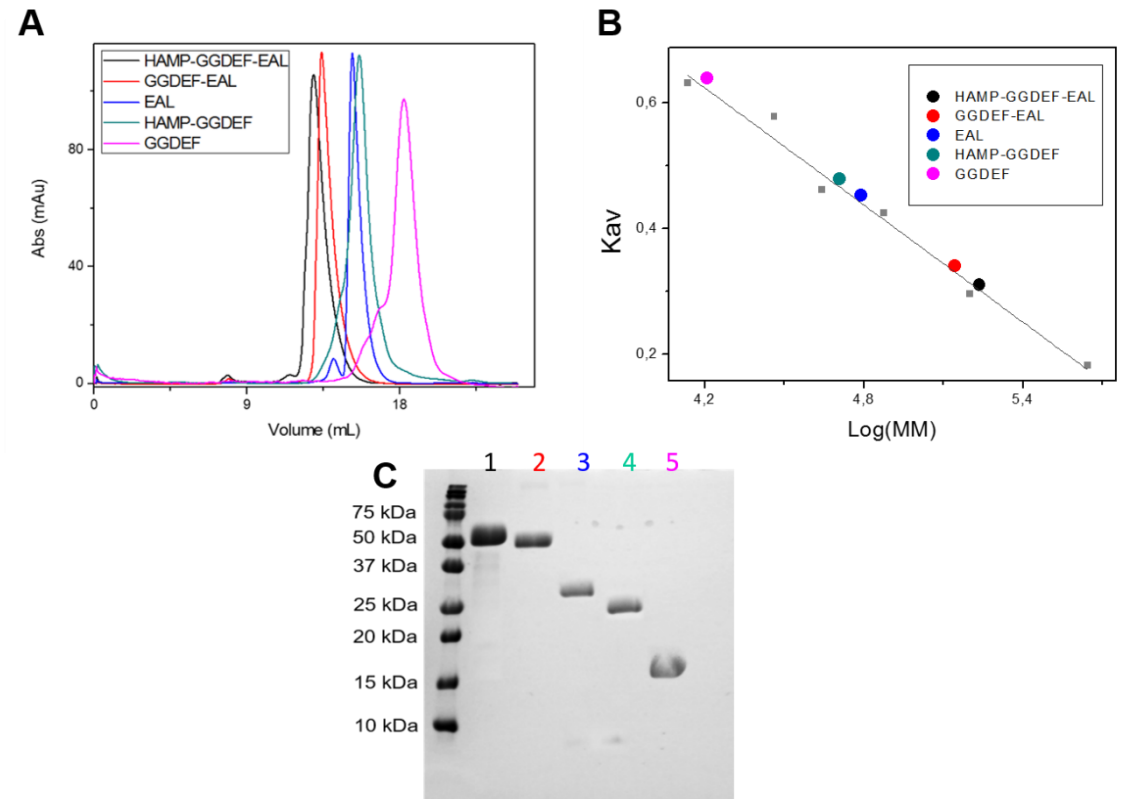


Figura 17 - Cromatografia de Exclusão Molecular analítica: domínios citoplasmáticos HAMP-GGDEF-EAL$^{184-668}$ (preto), GGDEF-EAL$^{242-668}$ (vermelho), EAL$^{392-668}$ (azul), HAMP-GGDEF$^{181-499}$ (verde) e GGDEF$^{242-399}$ (magenta) da STM3615. A) Curvas de eluição das construções na coluna superdex200 10/300: HAMP-GGDEF-EAL$^{184-668}$: 12,94 mL, GGDEF-EAL$^{242-668}$: 13,42 mL, EAL: 15,23 mL, HAMP-GGDEF$^{181-399}$: 15,64 mL e GGDEF$^{242-399}$: 18,21 mL. B) Calibração da coluna superdex200 10/300 com proteínas-padrão e ajuste dos pontos experimentais na curva de calibração dos domínios. C) Gel SDS 15% de amostras dos domínios HAMP-GGDEF-EAL$^{181-399}$ (1), GGDEF-EAL$^{242-668}$ (2), EAL$^{392-668}$ (3), HAMP-GGDEF$^{181-399}$ (4) e GGDEF$^{242-399}$ (5).

Fonte: Elaborada pelo autor.

Tabela 4 – Volume de eluição (Ve) e massa molecular (MW) obtidas através de cromatografia de exclusão molecular analítica.

Construção	Ve (mL)	MW teórica (kDa)	MW SEC (kDa)
HAMP-GGDEF-EAL$^{184-668}$	12,94	54,70	171,73
GGDEF-EAL$^{242-668}$	13,42	47,77	138,48
EAL$^{392-668}$	15,23	31,21	61,51
HAMP-GGDEF$^{181-399}$	15,64	26,07	51,18
GGDEF$^{242-399}$	18,21	16,80	16,17

Fonte: Elaborada pelo autor.

Através da equação da reta ajustada para os pontos determinados experimentalmente da curva de calibração (Figura 17B), e observando os valores apresentados na tabela 4, concluímos que as construções HAMP-GGDEF-EAL$^{184-668}$, GGDEF-EAL$^{242-668}$, EAL$^{392-668}$ e HAMP-GGDEF$^{181-399}$ possuem um padrão de eluição compatíveis com dímeros, enquanto apenas o domínio GGDEF$^{242-399}$ isolado é compatível com um padrão de eluição de um monômero.

4.2.2 Espalhamento de raios-X a baixo ângulo

Alternativamente, realizamos experimentos de SAXS na linha SAXS2 do LNLS com as construções HAMP-GGDEF-EAL$^{184-668}$, GGDEF-EAL$^{242-668}$, EAL$^{392-668}$ para confirmar os dados apresentados pela SEC. A figura 18 apresenta as curvas de espalhamento das amostras, dos modelos de átomos dummy (DAM) e dos modelos de corpo rígido, além das curvas de Guinier para análise dos dados obtidos e curvas de distribuição de distâncias (P(r)). Os modelos de átomos dummy obtidos para cada uma das construções foi sobreposto ao modelo de corpo rígido dos domínios GGDEF-EAL da STM3615 (Figura 18).

As curvas de Guinier mostram a qualidade da amostra e das curvas obtidas, que nesse caso foram satisfatórias. Os dados obtidos para a massa molecular das construções e os envelopes de SAXS obtidos apresentaram um bom ajuste em relação à massa teórica de dímeros e à estrutura cristalográfica do dímero dos domínios GGDEF-EAL da STM3615, corroborando com os dados obtidos por SEC.

Os dados apresentados na tabela 5 e na figura 18 D nos mostram, através das curvas de distribuição de distâncias e diâmetros máximos obtidos, que as construções analisadas possuem tamanhos diferentes, sendo HAMP-GGDEF-EAL a maior e EAL o menor. As massas moleculares aparentes calculadas por diferentes métodos a partir desses experimentos corroboram os estados oligoméricos observado nos experimentos de SEC. Interessante destacar que todas as construções estudadas por SAXS encontram-se diméricas em solução independentemente da presença de co-fatores, c-di-GMP ou até mesmo da

presença do HAMP, o qual é normalmente fundamental para oligomerização de proteínas de mesma arquitetura.

A formação dimérica do domínio EAL isolado observada por diferentes métodos nos permite comparar e verificar que a STM3615 apresenta diferenças relevantes em relação as proteínas já estudadas LapD de *P. fluorescens e* MorA e FimX de *P. aeruginosa*. Em todas essas proteínas, o domínio EAL isolado encontra-se no estado monomérico, na ausência de c-di-GMP.[44; 48] Interessantemente, enquanto que a ligação de c-di-GMP com alta afinidade ao domínio EAL de FimX não promove mudança oligomérica, a interação desse nucleotídeo com os domínios EAL de LapD e MorA leva a dimerização dos mesmos através da interface comumente encontrada em fosfodiesterases ativas, sejam elas dímeros transientes ou constitutivos. Enquanto que em LapD essa mudança oligomérica advém de um grande rearranjo conformacional interdomínios, o mesmo não é esperado para MorA e será discutido adiante.

Tabela 5 – Parâmetros estruturais derivados dos dados de SAXS para os domínios HAMP-GGDEF-EAL$^{184-668}$, GGDEF-EAL$^{242-668}$ e EAL$^{392-668}$ da STM3615.

	HAMP-GGDEF-EAL$^{184-668}$			**GGDEF-EAL$^{242-668}$**			**EAL$^{392-668}$**		
	Exp†	**DAM‡**	**RBM$^{\phi}$**	**Exp†**	**DAM‡**	**RBM$^{\phi}$**	**Exp†**	**DAM‡**	**RBM$^{\phi}$**
D_{max} (Å)	140,0	162.3	108,3	120,0	138,7	108,3	90,0	102,9	92,31
Rg (Å)	49,2	48.03	32,99	45,1	44.38	32,99	31,06	33,43	27,1
Resolução (Å)	19,7	-	-	19,7	-	-	19,7	-	-

† Calculados a partir dos dados experimentais.
‡ Parâmetros para modelos de átomos dummy.
$^{\phi}$ Parâmetros calculados para o modelo de corpo rígido.
Fonte: Elaborada pelo autor.

Tabela 6 – Massas moleculares calculadas através das curvas de SAXS das construções HAMP-GGDEF-EAL$^{184-668}$, GGDEF-EAL$^{242-668}$ e EAL$^{392-668}$.

Construção	**Teórica**	**SAXS Mow**	**Proteína Padrão****	**Escala Absoluta**
HAMP-GGDEF-EAL$^{184-668}$	54,7	176,1	118,8	160,5
GGDEF-EAL$^{242-668}$	47,8	163,5	98,6	118,6
EAL$^{392-668}$	31,2	63,2	69,3	66,1

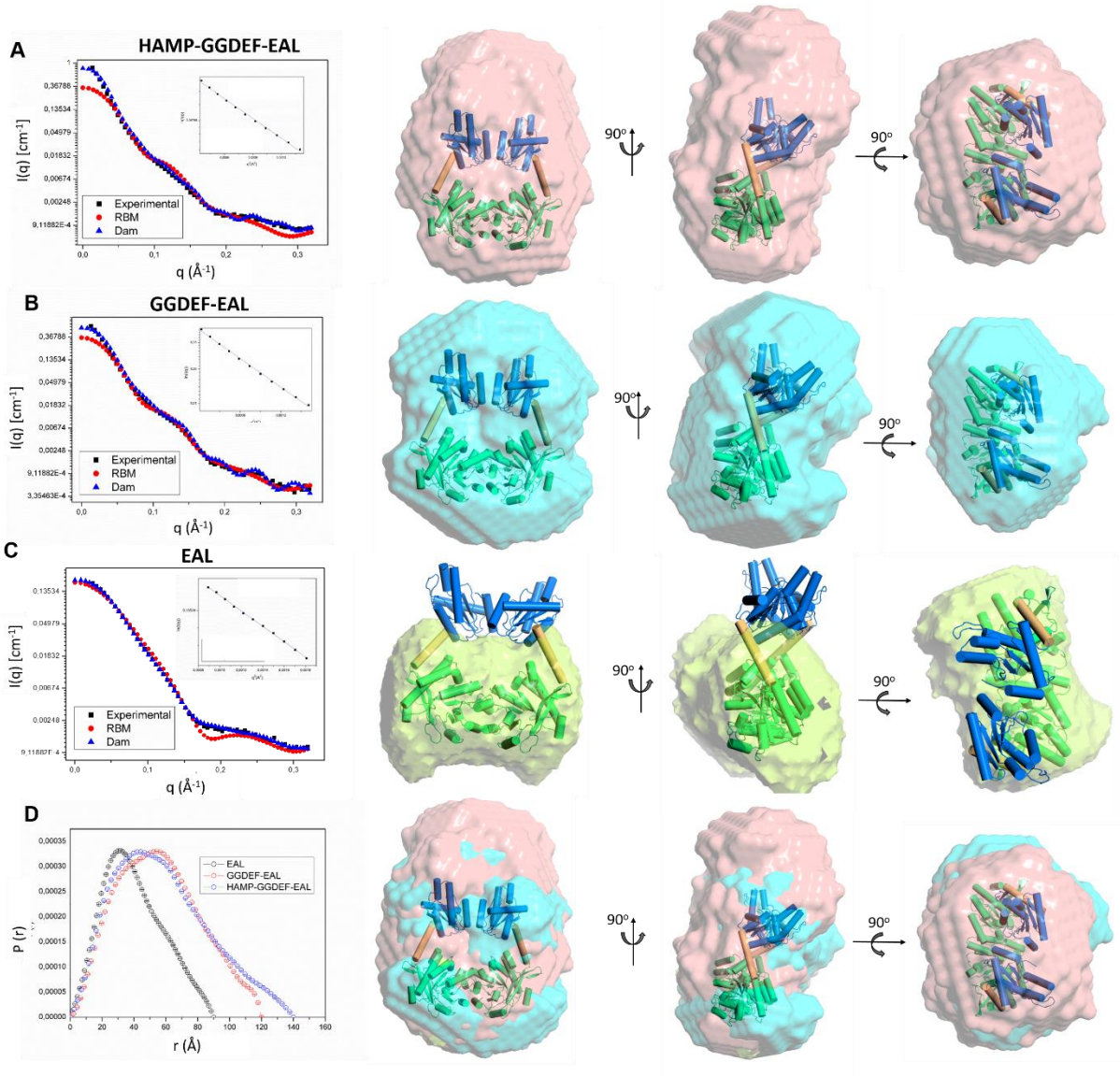


Figura 18 - Curvas de espalhamento de raios-X a baixo ângulo, curvas de distribuição de distâncias e modelos de SAXS. Podemos observar que para as três construções, HAMP-GGDEF-EAL (A), GGDEF-EAL$^{231-657}$ (B) e EAL (C), as curvas de experimentais se ajusta bem em relação aos modelos de baixa resolução obtidos e também às estruturas cristalográficas mostrando na inserção as curvas de Guinier[12] que demonstram a qualidade satisfatória dos dados obtidos. Todos os modelos apresentados foram comparados a estrutura cristalográfica dos domínios GGDEF-EAL, obtida neste trabalho, apresentadas em três disposições, rotacionadas de 90° em relação a anterior. A figura D, ilustra a sobreposição das curvas de distribuição de distâncias e dos três modelos obtidos. Observamos também que conforme esperado, as curvas de distribuição de distâncias possuem diâmetros máximos diferentes para cada uma das construções analisadas, sendo maior para a HAMP-GGDEF-EAL e menor para EAL (D).

Fonte: Elaborada pelo autor.

4.3 Ensaios de atividade enzimática

Os dados apresentados de oligomerização das construções sugerem que todas as construções analisadas estão na forma dimérica em solução, com exceção do domínio GGDEF isolado. Esse resultado condiz com a predição de atividade enzimática para STM3615, onde os resíduos essenciais para atividade somente são conservados no domínio EAL, uma vez que a interface dimérica observada em proteínas contendo EAL é essencial para atividade. Para verificar as atividades enzimáticas de STM3615 realizamos uma série de ensaios enzimáticos, separando os produtos por HPLC e identificados posteriormente por espectrometria de massas.

4.3.1 Hidrólise do c-di-GMP

A construção GGDEF-EAL$^{242-668}$ foi utilizada como padrão para determinação de protocolo de atividade e posterior obtenção e identificação do produto da fosfodiesterase 5'-pGpG através de cromatografia de fase reversa em coluna C18 (Figura 19).

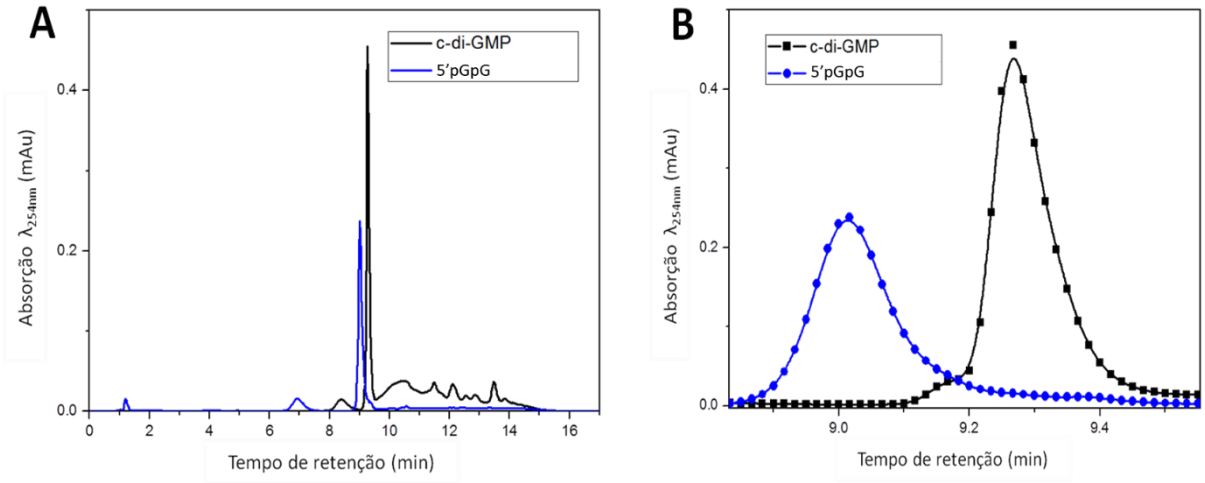


Figura 19 - Caracterização do produto da hidrólise do c-di-GMP pela GGDEF-EAL$^{242-668}$. Na figura A apresentamos o cromatograma completo e em B destacamos os picos que foram submetidos para análises de espectrometria de massas. As curvas em preto representam o c-di-GMP, utilizado aqui como controle do experimento (condição inicial), com pico em aproximadamente 9.3 minutos. As curvas em azul, referem-se ao produto da reação enzimática da construção GGDEF-EAL$^{242-668}$ com c-di-GMP, o pGpG, com pico em aproximadamente 9 minutos.

Fonte: Elaborada pelo autor.

Através das curvas obtidas pela cromatografia de fase reversa, podemos verificar que o domínio EAL estava ativo, porém, para confirmar o produto formado, as amostras relativas aos picos observados foram analisadas por espectrometria de massas (Figura 20), onde é possível confirmar que a construção GGDEF-EAL$^{242\text{-}668}$ hidrolisou todo o c-di-GMP (MM 690.1) em pGpG (MM 708.1).

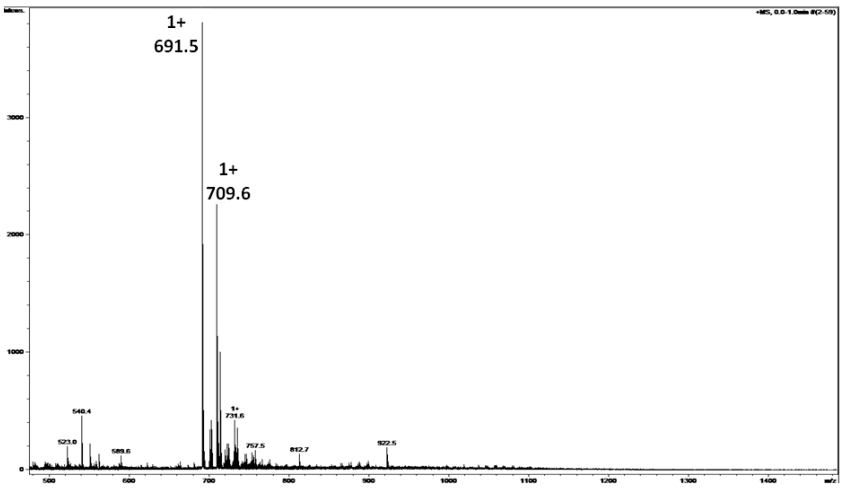


Figura 20 - Espectro de massas dos produtos de reação com a fosfodiesterase com c-di-GMP. Os picos com a relação carga/massa equivalentes a 691.5 e 709.6 correspondem ao c-di-GMP e 5'pGpG, respectivamente.

Fonte: Elaborada pelo autor.

4.3.2 Ensaios enzimáticos das construções da STM3615

Com os protocolos de ensaio enzimático estabelecidos e com o produto identificado, realizamos experimentos nas mesmas condições com cinco construções contendo os diferentes domínios estudados, HAMP-GGDEF-EAL$^{184\text{-}668}$, HAMP-GGDEF$^{181\text{-}399}$, GGDEF$^{242\text{-}399}$, GGDEF-EAL$^{242\text{-}668}$ e EAL$^{392\text{-}668}$. Os ensaios forma realizados incubando as proteínas com GTP ou c-di-GMP, para verificar a atividade de diguanilatociclase ou fosfodiesterase, respectivamente (Figura 21).

Como já esperado, o domínio GGDEF realmente é inativo na STM3615, enquanto o domínio EAL, independente do domínio HAMP ou GGDEF em sua construção, encontra-se na forma ativa, o que corrobora com os dados obtidos para

o estado oligomérico obtido nos resultados apresentados anteriormente e também diferencia a STM3615 das demais proteínas dual domain até o momento estudadas.

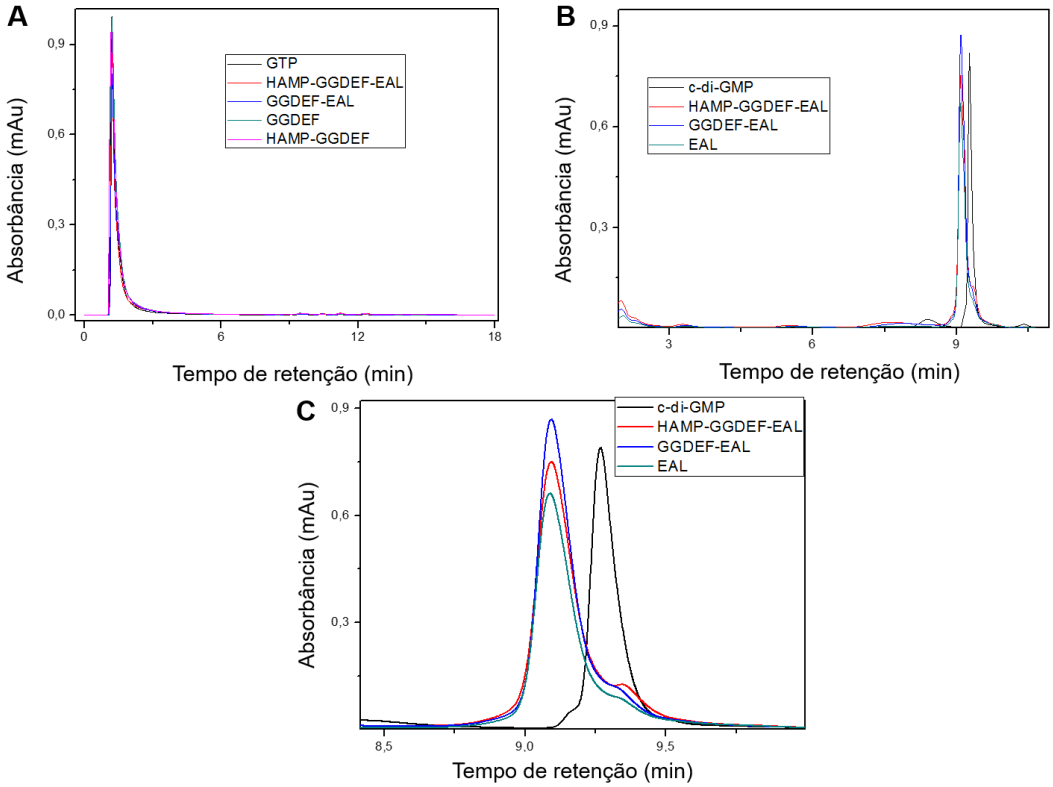


Figura 21 - Caracterização da atividade enzimática das construções HAMP-GGDEF-EAL$^{184-668}$, HAMP-GGDEF$^{181-399}$, GGDEF$^{242-399}$, GGDEF-EAL$^{242-668}$ e EAL$^{392-399}$. A) As construções que possuem o domínio GGDEF, foram incubadas com GTP e analisadas posteriormente por HPLC. A figura nos mostra que para nenhuma construção foram obtidos picos de c-di-GMP, ou seja, os domínios GGDEF encontram-se na forma inativa. B) As construções que possuem domínio EAL foram submetidas a ensaios com o substrato c-di-GMP e como podemos observar na figura ampliada (C), todas elas apresentaram um pico correspondente a pGpG, mostrando que todas as construções são ativas.

Fonte: Elaborada pelo autor.

4.4 Ensaios de cristalização dos domínios solúveis

Das construções obtidas em frações solúveis e submetidas a ensaios iniciais de cristalização com os kits Hampton, foram obtidas condições favoráveis para cristalização apenas das construções GGDEF-EAL$^{218-668}$ e GGDEF-EAL$^{242-668}$. Para

a construção GGDEF-EAL$^{218-668}$, nove condições de cristalização foram identificadas, e para GGDEF-EAL$^{242-668}$,15 condições, contendo em geral *tacsimate,* tartarato de sódio e potássio, citrato de lítio, PEG de tamanhos entre 3350 e 6000, e tampões com pH entre 7.0 e 8.0.

Os cristais obtidos nas condições mencionadas não apresentaram resolução adequada para resolver a estrutura, com imagens a aproximadamente 6Å de resolução.

Com o objetivo de melhorar a qualidade dos cristais para aumentar a resolução dos conjuntos, realizamos alguns refinamentos das condições de cristalização obtidas. A solução contendo 15% *tacsimate* pH7,0 (Hampton), 0.1M HEPES pH7,0 e 4% PEG 3350 foi a mais adequada, apresentou cristais maiores quando utilizamos proteína a 25mg/mL em uma proporção de 1:3 de solução de cristalização e proteína. Os cristais apareciam em cerca de 24 horas e cresciam por cerca de 3 ou 4 dias. Os conjuntos obtidos com estes cristais chegaram a 3,5 Å (Figura 22A), que mesmo após diversas tentativas não foi suficiente para resolver a estrutura cristalográfica nestas condições.

Paralelamente à busca de cristais melhores com novos testes de refinamento da condição de cristalização, realizamos novos *screenings* com o kit Morpheus (Molecular Dimensions). Com este kit, obtivemos cristais em 16 condições para a construção GGDEF_EAL$^{242-668}$, apresentando em geral condições com PEG de tamanhos entre 4000 e 8000, glicerol ou etilenoglicol, haletos, NPS e ácidos carboxílicos (Haletos: fluoreto de sódio, brometo de sódio e iodeto de sódio; NPS: nitrato de sódio, fosfato de sódio dibásico e sulfato de amônio; Ácidos Carboxílicos: formato de sódio, acetato de amônio, citrato trisódico, tartarato de sódio e potássio e oxamato de sódio).

A melhor condição testada para os cristais nativos coletados foi 10% m/v PEG 8000, 20% v/v etilenoglicol, 0,03 M de NPS e 0,1 M MES/imidazol pH 6.5, em que obtivemos um conjunto com resolução de 2,5 Å (Figura 22C).

Mesmo com dados a essa resolução, não foi possível resolver a estrutura cristalográfica por Substituição Molecular. Buscando outros métodos, realizamos algumas tentativas de obter as fases com conjuntos obtidos de cristais em que foram submetidos a soaking com iodo. Nesse caso, as fases foram determinadas, mas o mapa obtido apresentou baixa qualidade, sendo muito desconexo. Ainda como alternativa, buscamos obter cristais da proteína incorporada com selenometionina.

Com a proteína pura, reproduzimos o *screening* com o kit Morpheus, onde obtivemos cristais na maioria das condições já encontradas anteriormente, sendo a mais adequada para coleta a condição 10% m/v PEG4000, 20% v/v Glicerol, 0.03 M Haletos e 0.1 M MES/Imidazol pH 6,5 em que coletamos um conjunto a 2,8 Å de resolução (Figura 22B).

Como obtivemos um conjunto nativo com resolução maior que os que continham selênio, posteriormente a resolver a estrutura, utilizamos o programa *Density Modification* (DM)[77] para estender as fases, chegando a 2.5 Å de resolução. Os resultados da coleta de dados, processamento e refinamento estão resumidos na tabela 7.

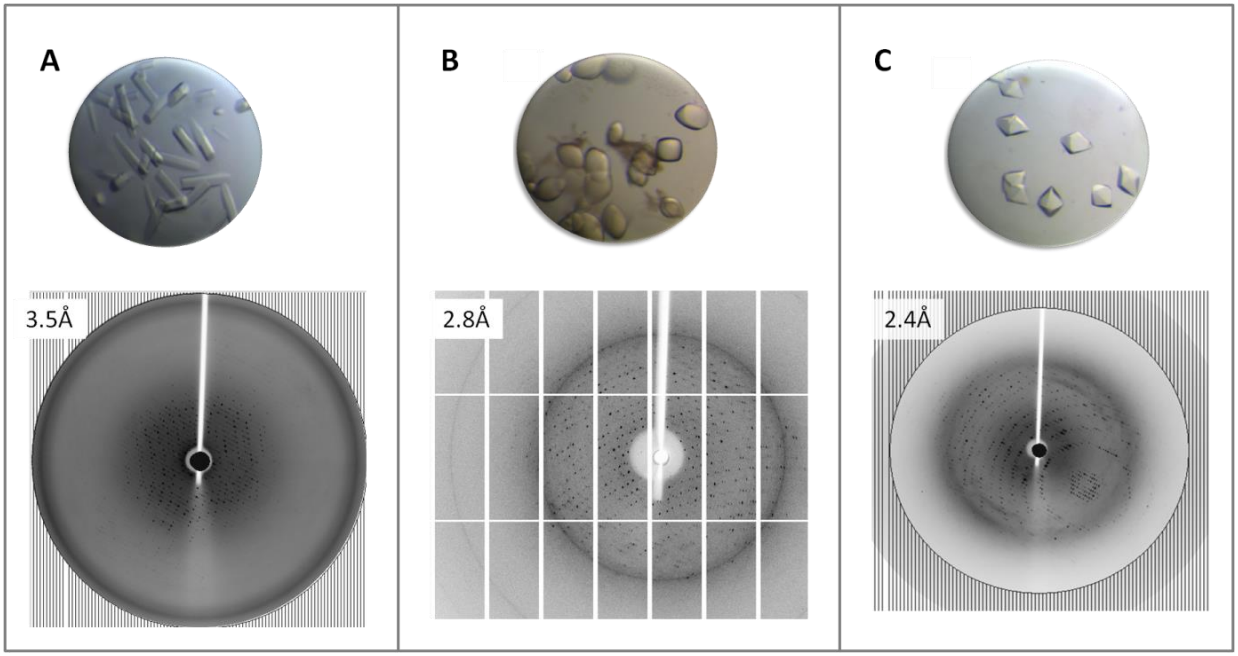


Figura 22 - Cristais e padrões de difração da GGDEF_EAL$^{242\text{-}668}$. A) Cristais obtidos na condição 15% *tacsimate* pH7,0 (Hampton), 0,1 M HEPES pH7,0 e 4% PEG 3350 e padrão de difração a 3,5 Å. B) Cristais obtidos com a proteína contendo selenometionina na condição 10% m/v PEG4000, 20% v/v Glicerol, 0,03 M Haletos e 0,1 M MES/Imidazol pH 6,5 e padrão de difração a 2,8 Å. C) Cristais obtidos na condição 10% m/v PEG 8000, 20% v/v Etilenoglicol, 0,03 M NPS e 0,1 M MÊS/Imidazol pH 6,5 e padrão de difração a 2,5 Å.

Fonte: Elaborada pelo autor.

A estrutura tridimensional dos domínios GGDEF_EAL$^{242\text{-}668}$ foi então determinada a resolução de 2,5 Å. A proteína foi cristalizada no grupo espacial $P4_32_12$ e através do programa Matthews_Coef,[78] verificamos que o cristal apresentava porcentagem de solvente de aproximadamente 52%, contendo apenas uma molécula na unidade assimétrica.

Além dos domínios GGDEF e EAL, obtivemos também alguns cristais do domínio periplasmático com a construção DP$^{46-150}$ (Figura 23) na condição 10% m/v PEG 20000, 20% v/v PEG MME 550, 0,02 M de D-glicose, D-manose, D-galactose, L-fucose, D-xilose, N-acetil-D-glucosamina, 0,1 M bicina/Trizma base pH8,5, porém, os cristais eram muito pequenos, extremamente frágeis e não obtivemos nenhum conjunto de dados adequado para a determinação da estrutura cristalográfica.

Tabela 7 – Estatísticas de coleta, processamento e refinamento dos dados coletados para os cristais obtidos com a construção GGDEF-EAL$^{242-668}$.

	GGDEF_EAL ***Nativo***	**GGDEF_EAL** ***Se***	**GGDEF_EAL** ***Nativo***
Comprimento de onda (Å)	1,458	0,9789	1,458
Grupo Espacial	P4	$P4_32_12$	$P4_32_12$
Parâmetros da Cela Unitária			
a (Å),	186,7	83,5	82.25
b (Å),	186,7	83,5	82.25
c (Å),	134,0	144,0	143.90
α (°) ,	90,0	90	90,0
β (°),	90,0	90	90,0
γ (°)	120,0	90	90,0
Faixa de Resolução (Å)	46,6 (4,2)	48,1, (2,8)	29,3 (2,5)
Número de Reflexões			
Total	54974 (7916)	332078 (44785)	211498 (32732)
Única	10431 (1643)	44785 (3999)	17995 (2805)
Completeza (%)	98,7(99,6)	99,8 (98,9)	99,3 (99,1)
Redundância	5,27 (4,82)	13,35 (11,20)	11,78 (11,67)
I/σ (I)	12,6 (2,0)	19,18 (2,04)	21,46 (2,89)
R_{meas}(%)	12,4 (3.46)	13,4 (144,6)	10,4 (120,02)
R_{work} / R_{free}	-	-	24,6 / 30,3
R_{rmsd}			
Comprimento da ligação (Å)	-	-	0,009
Ângulo da ligação (º)	-	-	1,36
Ramachandran (%)			
Favoráveis			
Permitidas	-	-	92.2
Não permitidas	-	-	7.8
	-	-	0

Fonte: Elaborada pelo autor.

Figura 23 - Cristais do DP^{46-150}. Cristais obtidos para a construção DP^{46-150} na condição 10% m/v PEG 20000, 20% v/v PEG MME 550, 0,02 M de D-glicose, D-manose, D-galactose, L-fucose, D-Xilose, N-acetil-D-glucosamina, 0,1 M bicina/Trizma base pH8,5.

Fonte: Elaborada pelo autor.

4.5 Estrutura cristalográfica dos domínios GGDEF_EAL da STM3615

A estrutura foi resolvida por Se-SAD, onde foi possível observar densidade eletrônica contínua para os dois domínios, GGDEF e EAL, da cadeia na unidade assimétrica. Aplicando uma operação de simetria do cristal é possível ver o dímero de STM3615, com a interface comumente encontrada entre domínios EAL (Figura 24A). O enovelamento dos dois domínios está conservado em relação a aqueles já depositados no PDB. O domínio GGDEF possui conformação de um sanduíche alfa-beta, composto por uma folha β central de 4 fitas β antiparalelas, envolta por 5 hélices α. O domínio EAL apresenta 8 hélices α e 8 fitas β na forma de um barril TIM (Figura 24B). Os domínios GGDEF e EAL são conectados por uma hélice de 22 resíduos.

Apesar de estruturalmente similar a MorA (veja Figura 30), a estrutura obtida nesse trabalho é inédita entre as estruturas disponíveis atualmente, sendo a primeira e única que apresenta o domínio GGDEF degenerado e EAL ativo.

A fim de verificar a conservação dos resíduos que tornam o domínio EAL ativo para fosfodiesterase, realizamos um alinhamento de sequencias de aminoácidos com 30 proteínas de diferentes organismos e que apresentavam arquitetura próxima a da STM3615.

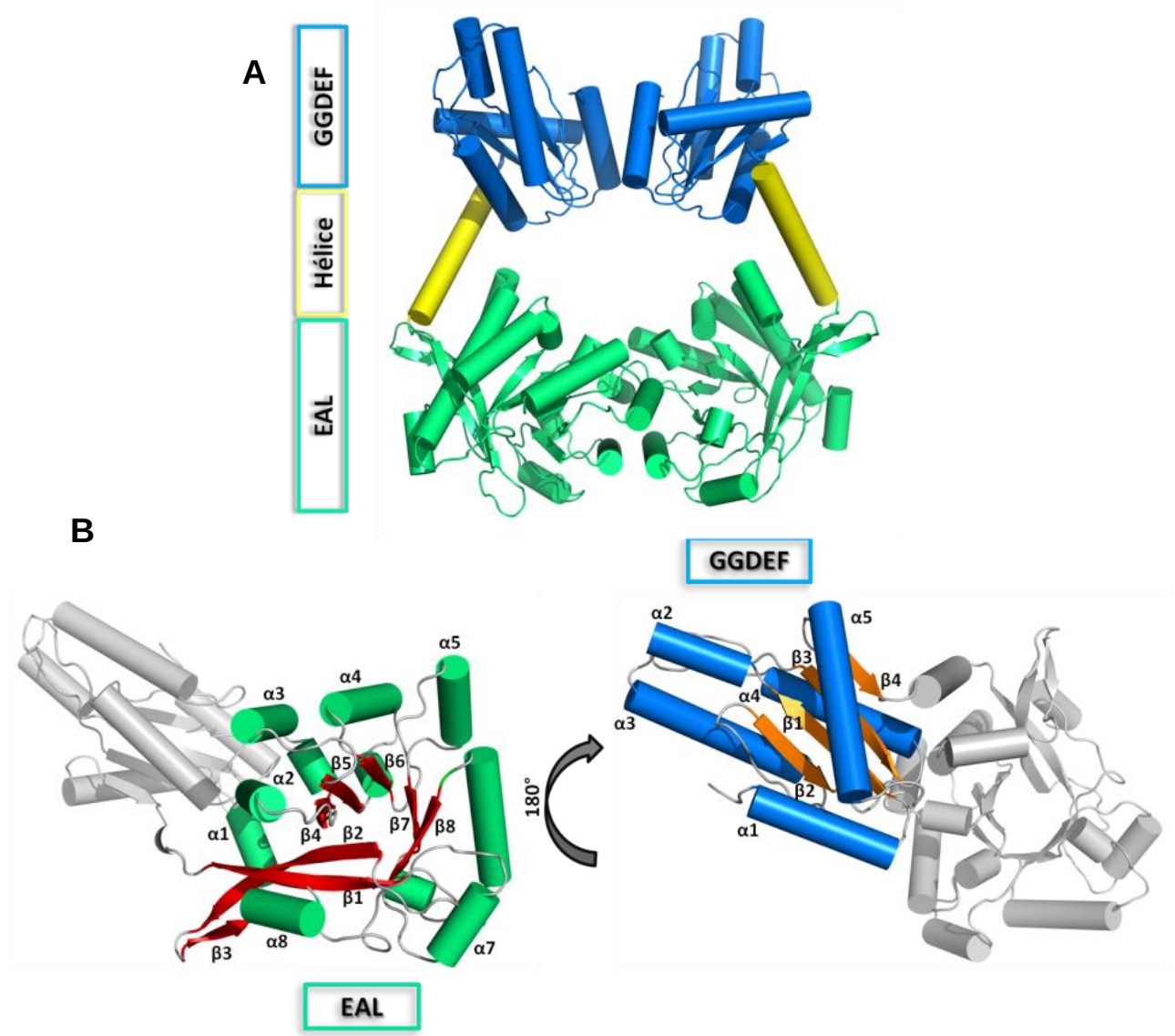


Figura 24 - Estrutura cristalográfica dos domínios GGDEF e EAL da STM3615. A) Estrutura, na forma de um homodímero ativo para fosfodiesterase, com cada unidade em cores diferentes, destacando os domínios GGDEF (azul) e EAL (verde), unidos por uma hélice α (amarelo). B) Identificação da conformação dos domínios EAL (em verde) formado por 8 hélices α e 8 fitas β na conformação de um barril alpha-beta e GGDEF (em azul), na conformação de um sanduíche alpha-beta, composto por uma folha β central de 4 fitas β antiparalelas, envolta por 5 hélices α.

Fonte: Elaborada pelo autor.

Através do alinhamento, podemos verificar que todos os resíduos necessários[42] para o domínio EAL promover a catálise estão conservados, como mostra a figura 25A gerada pelo programa WebLogo.[79] Esses resíduos estão destacados na figura 25B, que mostra a molécula de c-di-GMP do PDB 3N3T. Além disso, o mesmo alinhamento foi utilizado para rodar o programa Consurf,[80] que

gerou um *score* de acordo com a conservação dos resíduos em que podemos analisá-los através da estrutura cristalográfica determinada, em que confirmamos que a região mais conservada é a região catalítica do domínio EAL (Figura 25C). Com a finalidade de verificar a posição das regiões catalíticas do dímero e consequentemente das moléculas de c-di-GMP, utilizamos novamente o c-di-GMP do PDB 3N3T, e assim, podemos observar que o sítios catalíticos estão em orientações diferentes em relação ao dímero (Figura 25D).

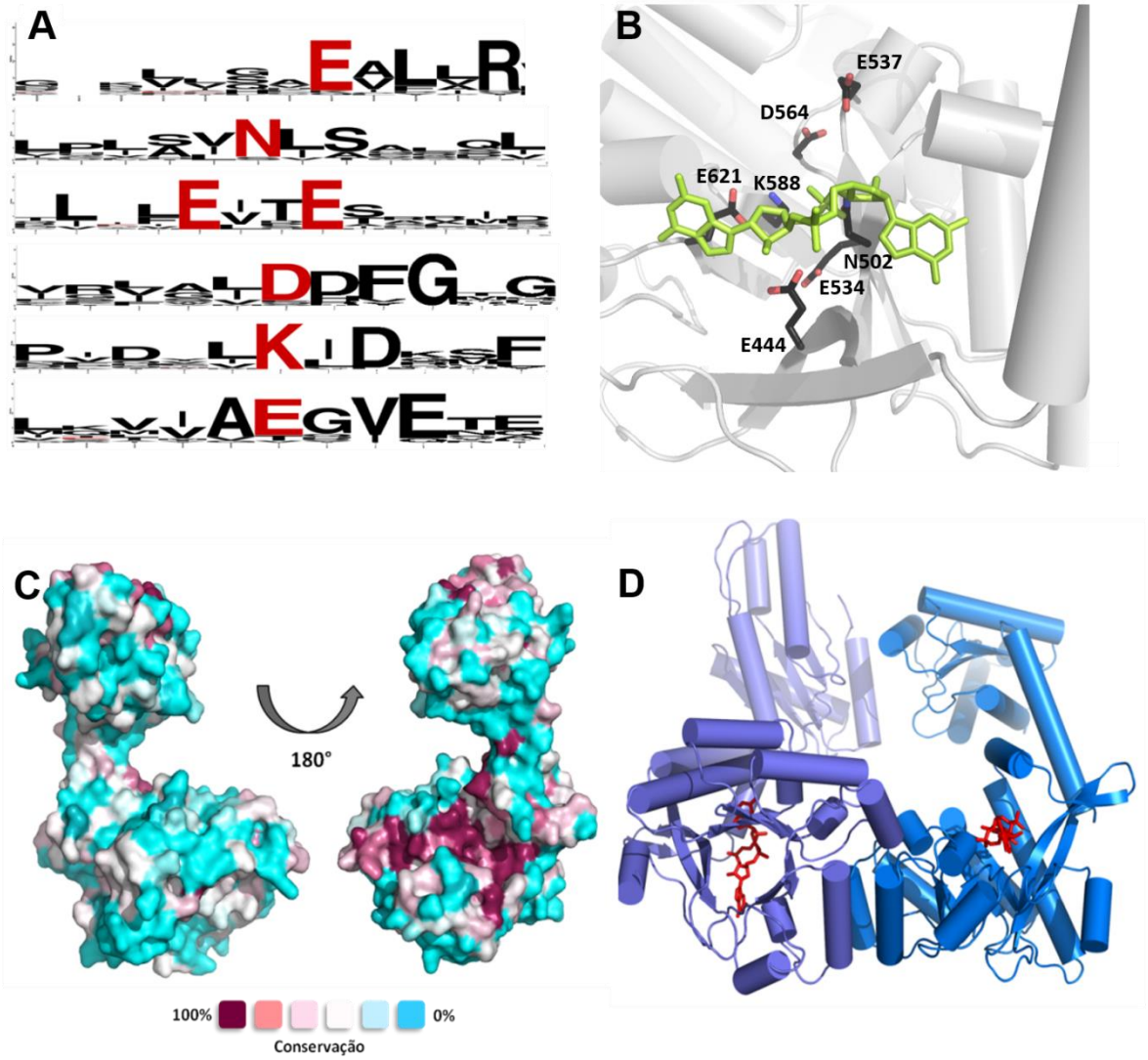


Figura 25 - Conservação de resíduos dos domínios GGDEF e EAL da STM3615. A) A figura apresenta a análise de conservação dos resíduos necessários para promover a atividade catalítica do domínio EAL gerada pelo programa WebLogo, sendo eles o E444, N502, E534, E537, D564 ,K588 e E621, respectivamente. B) Sobreposição do domínio EAL da STM3615 com o c-di-GMP da estrutura da fosfodiesterase de Thiobacillus denitrificans (PDB:3N3T), destacando a posição dos resíduos E444, N502, E534, E537, D564 ,K588 e E621, identificados como responsáveis pela hidrólise do c-di-GMP. C) Através do alinhamento múltiplo de sequências e o score dado a conservação dos resíduos pelo programa Consurf, podemos verificar que a região mais conservada dos domínios é a região catalítica do barril do EAL, mostrada ao lado esquerdo da figura. D) Sobreposição do dímero dos domínios GGDEF_EAL da STM3615 com o c-di-GMP do PDB 3N3T, onde observa-se que as moléculas de c-di-GMP estão em posições relativas diferentes se comparadas entre as unidades monoméricas.

Fonte: Elaborada pelo autor.

Se comparada a estrutura cristalográfica dos domínios GGDEF-EAL da STM3615, o enovelamento da MorA é muito semelhante (Figura 26), possuindo inclusive uma hélice estendida ligando os dois domínios.

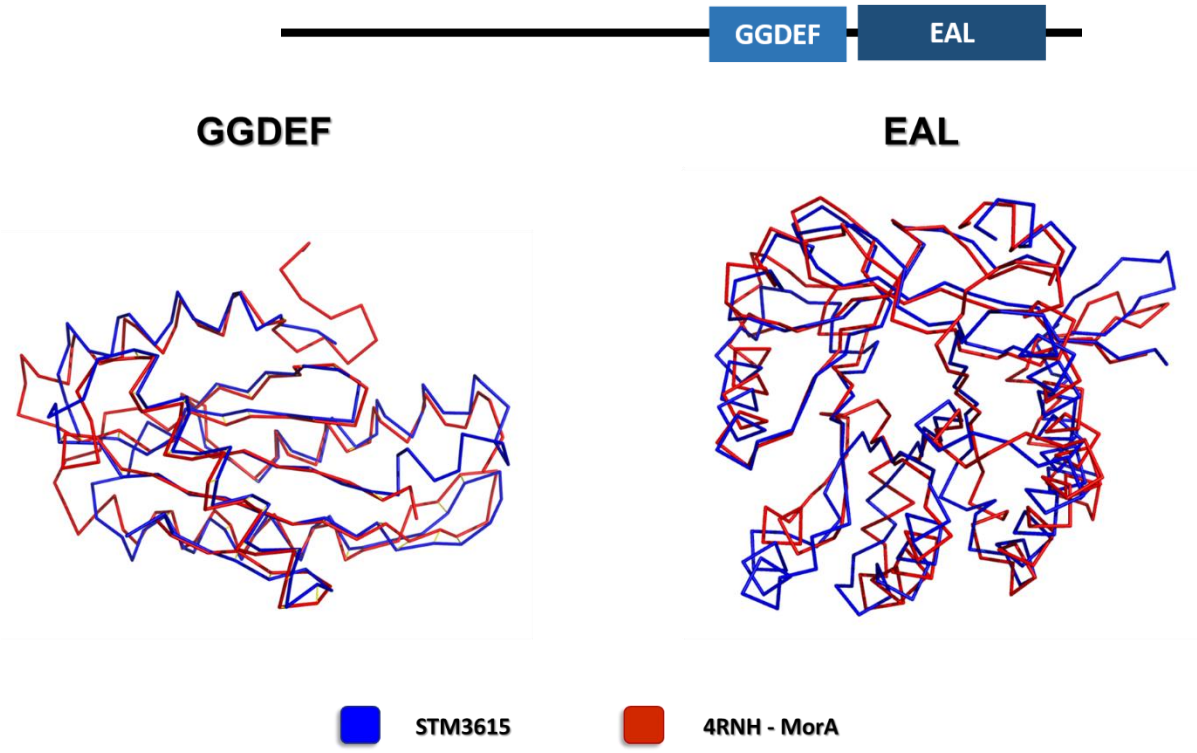


Figura 26 - Sobreposição dos domínios GGDEF-EAL da MorA de *Pseudomonas aeruginosa* e STM3615 de *Salmonella enterica*. As estruturas cristalográficas dos domínios GGDEF-EAL da MorA e STM3615, mostram ter o mesmo enovelamento, porém nota-se que de maneira geral principalmente a estrutura do domínio EAL se sobrepõe parcialmente deslocada, o que justifica o valor de RMS_{EAL} igual a 1,86, enquanto o RMS_{GGDEF} é de 1,07.

Fonte: Elaborada pelo autor.

Observando a disposição dos domínios da estrutura obtida, a forma que se conectam e comparando com a LapD de *Pseudomonas fluorescens*, os domínios GGDEF e EAL encontram-se em disposições relativas diferentes, onde aparentemente a principal diferença se dá pela hélice que conecta os dois domínios. Buscando visualizar esta diferença, alinhamos a estrutura obtida com a LapD (PDB:3PJX),[44] FimX (PDB:4J40),[48] ambas com domínio EAL degenerado, e MorA

(PDB:4RNH),[10] com GGDEF e EAL ativos. Neste caso, verificamos que as estruturas apresentam hélices na mesma região, porém, a hélice da STM3615 e MorA é maior que as hélices da LapD e FimX, deixando sua estrutura mais estendida (Figura 27). Essa alteração na região que conecta os domínios GGDEF e EAL muito provavelmente deixa as estruturas de MorA e STM15 mais rígidas, impossibilitando a mudança conformacional intra-domínios observada em LapD.

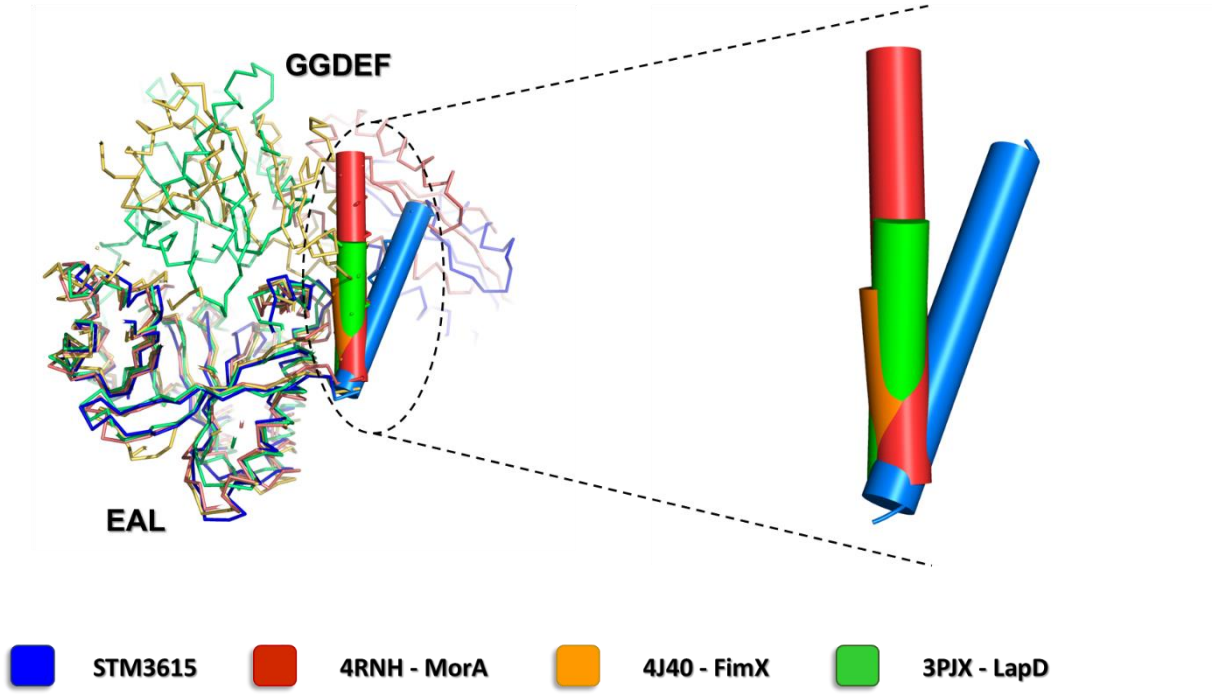


Figura 27 - Comparação das hélices entre os domínios GGDEF e EAL: STM3615, MorA, FimX e LapD. Através da sobreposição das estruturas é podemos observar que as quatro estruturas possuem uma hélice α na região que conecta os domínios, porém, as hélices das proteínas com resíduos catalíticos conservados, STM3615 e MorA, tem tamanho proporcionalmente maior que os receptores de c-di-GMP, LapD e FimX, fazendo com que sua estrutura esteja em uma forma mais estendida.

Fonte: Elaborada pelo autor.

4.6 Análises computacionais de sequencias codificantes para proteínas GGDEF-EAL

Com base nestas observações e para verificar a possibilidade de existir um padrão de hélices longas para domínios GGDEF-EAL ativos e hélices mais curtas para GGDEF-EAL inativos, buscamos sequências de diversos organismos que apresentassem arquitetura próxima a da STM3615, apresentando os domínios GGDEF e EAL. Inicialmente, um pequeno número de sequências foram

selecionadas e divididas entre aquelas que possuem os domínios GGDEF e EAL degenerados, e aquelas com GGDEF degenerado e EAL com os resíduos catalíticos conservados. Com as sequências alinhadas, para verificar os domínios conservados ou degenerados, colocamos também sequências de proteínas modelos para diguanilato ciclases, a WspR e fosfodiesterase, RocR, ambas de *Pseudomonas aeruginosa*, além da FimX e LapD (Figura 28).

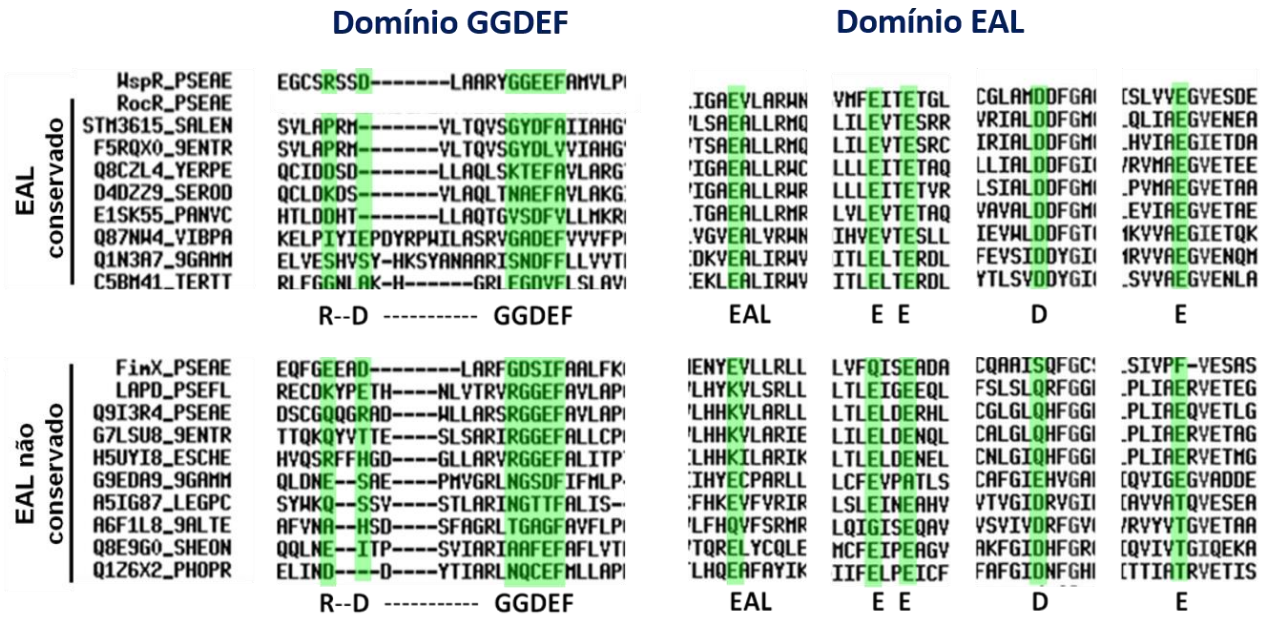


Figura 28 - Conservação dos domínios GGDEF e EAL. Podemos observar que entre as sequências apresentadas, somente a WspR possui o sítio de inibição RxxD e a região catalítica GGEEF conservadas. As demais proteínas estão divididas entre as ativas e não ativas para fosfodiesterase.

Fonte: Elaborada pelo autor.

Identificadas as sequências, separadas entre as diferentes formas de conservação, utilizamos o servidor Psipred para obter a predição da estrutura secundária de cada uma das proteínas selecionadas. Além disso, alinhamos cada uma das sequências com a STM3615 de forma independente, buscando identificar e comparar a região da hélice que conecta os domínios GGDEF e EAL na sequência e consequentemente na predição de estrutura secundária (Figura 29).

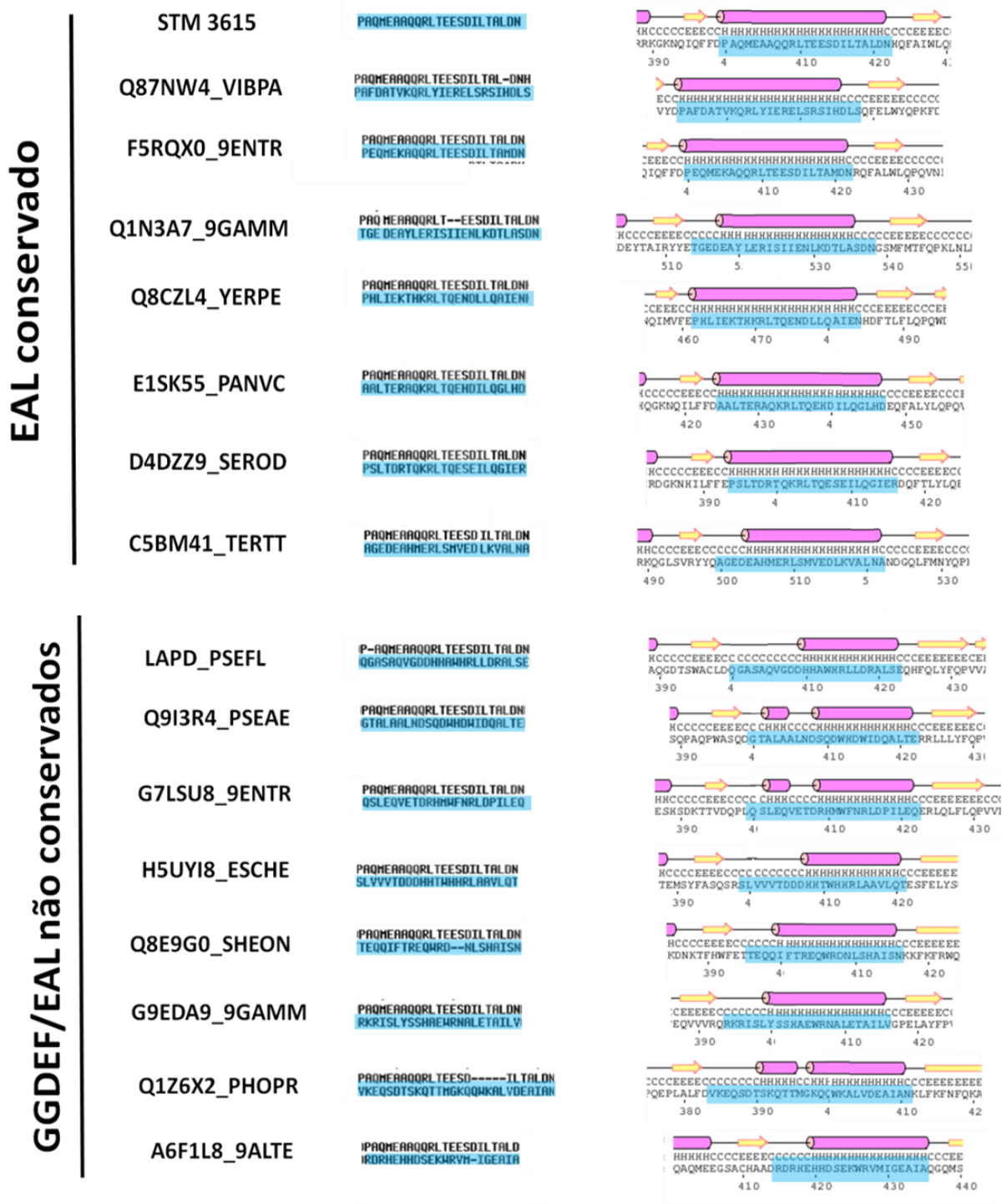


Figura 29 - Predição de estrutura secundária de sequências com EAL conservado e degenerado. As diversas sequências da região da hélice que conecta os domínios GGDEF e EAL, determinadas pelo alinhamento da sequência conhecida para a STM3615, possuem um padrão onde a hélice sempre está entre duas fitas betas, porém, ao compararmos as hélices do grupo com EAL conservado e não conservado, observa-se um padrão diferente entre elas, onde aquelas com EAL conservado são maiores e com média de 21 resíduos e aquelas com EAL degenerado são menores, ou duplas, em média com 15 resíduos.

Fonte: Elaborada pelo autor.

Através dessa análise, notamos que ao compararmos as regiões entre as fitas dos dois grupos, EAL conservado e degenerado, o número de resíduos entre as duas fitas é em média 28,1 ± 0,6 resíduos para os conservados e 27,1 ± 2,0 resíduos para os não conservados, ou seja, em média, temos números de resíduos muito próximos para essas regiões. Se ainda verificarmos o número de resíduos formando loops, observamos 7,0 ± 2,0 e 11,6 ± 2,0 para os conservados e não conservados, respectivamente. As hélices têm em média 21,1 ± 2,0 e 15,5 ± 2,0 resíduos para esses grupos, sugerindo que para a arquitetura analisada, quando o domínio EAL é conservado e o GGDEF é degenerado, temos hélices maiores, como no caso da STM3615, e quando os dois domínios estão degenerados, temos hélices menores conectando os dois domínios, como no caso da LapD.

Como o número de sequências não possuía ainda um número significativo para confirmar de fato essa hipótese, realizamos os mesmos procedimentos, como descrito na metodologia deste trabalho, para 150 sequências, os resultados então obtidos confirmam as previsões obtidas pelo procedimento feito inicialmente, conforme ilustrado na tabela 8 e figura 30. Esse padrão de comprimento da hélice entre os domínios GGDEF e EAL está de acordo com as análises feitas para STM3615, MorA e LapD, provavelmente indicando que as proteínas GGDEF-EAL cataliticamente ativas, EAL-mono ou bifuncionais, não sofram mudanças conformacionais intra-domínios nos seus mecanismos de regulação.

Tabela 8 –Dados dos domínios GGDEF-EAL analisados.

Conservados	Não Conservados	N° de sequências	Resíduos entre fitas	Loops	Hélice	RXXD Conservados
GGDEF-EAL	_	85	30,7 ± 3,4	8,8 ± 3,1	21,9 ± 0,9	44
GGDEF	**EAL**	3	27,0 ± 1,0	8,7 ± 0,6	18,0 ± 1,0	3
EAL	**GGDEF**	45	30,1 ± 4,2	9,18 ± 3,1	21,0 ± 2,2	7
_	**GGDEF-EAL**	17	28,5 ± 2,7	11,3 ± 3,8	17,4 ± 2,1	1

Fonte: Elaborada pelo autor.

Adicionalmente, uma análise de conjunto representativo de proteínas *dual domain* revela uma distribuição de tipos funcionais inédito (Figura 30). Proteínas contendo todos os resíduos essenciais para atividade conservados para ambos GGDEF e EAL são a maioria (~58%), seguido pelo grupo de proteínas GGDEF-EAL com atividade fosfodiesterase predita (grupo da STM3615) com ~30% das sequências. Enquanto que as proteínas com domínios degenerados, e que potencialmente tornaram-se receptores de c-di-GMP com FimX e LapD, ainda representam uma boa parcela dessa classe de proteínas (~11%), as sequencias com atividade predita somente para o domínio GGDEF são apenas 1% desse conjunto. Apesar da conservação dos resíduos em GGDEF nesse último tipo de proteína *dual domain* a atividade catalítica *in vivo* pode ser irrelevante, como já observado em outros casos, indicando talvez que possam na verdade atuar apenas como receptores de c-di-GMP através de seus domínios EAL degenerados. Outra análise interessante refere-se à conservação do sítio inibitório RxxD (sítio I) dos domínios GGDEF. Enquanto que todas as sequencias GGDEF-conservado/EAL-degenerado, embora minoria, apresentam sítios I conservados, essa proporção cai para aproximadamente 5% e 12% nas proteínas em que o GGDEF é degenerado, GGDEF-degenerado/EAL-degenerado e GGDEF-degenerado/EAL-conservado respectivamente. Essa baixa taxa de conservação do sítio I provavelmente ocorre devido a desnecessidade de regulação alostérica por c-di-GMP da atividade desses domínios, os quais uma vez inativos não sofreram uma pressão seletiva funcional para sua manutenção. Já no caso das enzimas bifuncionais a taxa de conservação dos sítios I é surpreendentemente baixa, em torno de 52%, quando comparada a enzimas somente contendo o domínio GGDEF, onde a quase totalidade mantém esse sítio alostérico conservado. Fica claro que nesse tipo de enzima bifuncional a auto-regulação da síntese de c-di-GMP está a cargo de outros mecanismos, incluindo o discutido na seção seguinte.

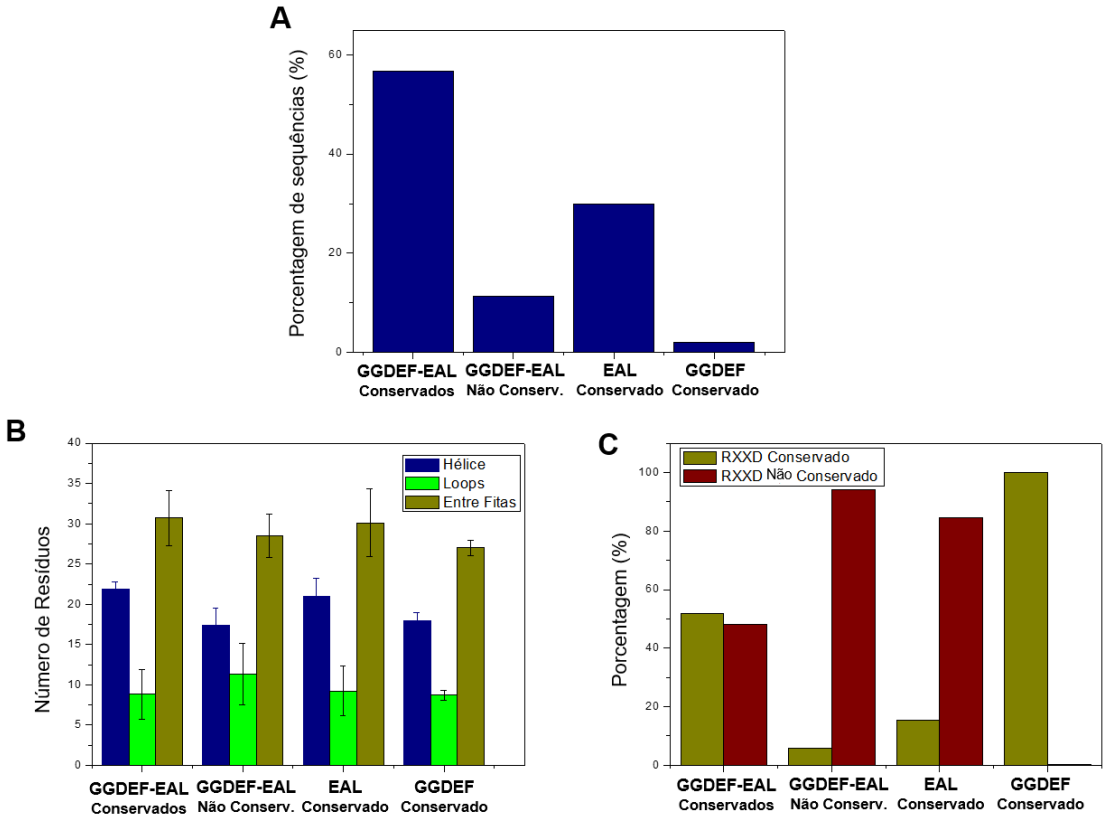


Figura 30 - Representação do número de sequências com domínios GGDEF-EAL quanto a atividade, predição de estrutura secundária e sítio inibitório do domínio GGDEF. A) Entre as 150 sequências, nota-se que a maior parte encontra-se na porção onde os resíduos fundamentais para a catálise estão conservados e a minoria na parcela em que apenas os resíduos catalíticos do domínio GGDEF não estão conservados. B) O número de resíduos entre as fitas das extremidades onde encontra-se a hélice não apresenta grandes variações, porém, ao analisar o tamanho das hélices, observa-se uma similaridade entre o número de resíduos que formam a hélice que conecta os domínios GGDEF e EAL para a fração em que resíduos catalíticos conservados apenas no domínio EAL ou para ambos os domínios, sugerindo então que de maneira geral estes possuem hélices estendidas, enquanto para o caso em que temos apenas o domínio GGDEF conservado e GGDEF-EAL não conservados, os dados sugerem hélices mais curtas. C) Entre as sequências analisadas, para os domínios GGDEF-EAL conservados, a presença do sítio inibitório RxxD estava presente em pouco mais da metade das sequências, sugerindo dois mecanismos distintos para a regulação da síntese e degradação de c-di-GMP nessa situação. A presença do RxxD no caso onde nenhum dos domínios apresenta os resíduos catalíticos conservados ou conservados apenas no EAL, possuem baixa taxa de sítio inibitório, provavelmente porque este não tem nenhum papel fundamental nas duas situações, diferentemente do caso onde apenas os resíduos catalíticos do domínio GGDEF estão conservados, em que todas as sequências apresentaram RxxD conservado.

Fonte: Elaborada pelo autor.

4.7 Modelo de mecanismo regulatório para enzimas GGDEF-EAL bifuncionais

Analisando a estrutura primária e as estruturas cristalográficas da enzima bifuncional MorA (Figura 31B), observamos primeiramente que esta pertence exatamente ao grupo das ciclases GGDEF ativas, com sítio inibitório RxxD degenerado (Figura 31A). Na estrutura cristalográfica apresentada para o dímero, que seria a forma fosfodiesterase ativa, notamos que os sítios catalíticos e motivos conservados do domínio GGDEF encontram-se em sentidos opostos, o que não justifica a atividade catalítica da diguanilato ciclase. Essa conformação mostra que atividade ciclase não pode ocorrer concomitantemente com atividade fosfodiesterase (Figura 31).

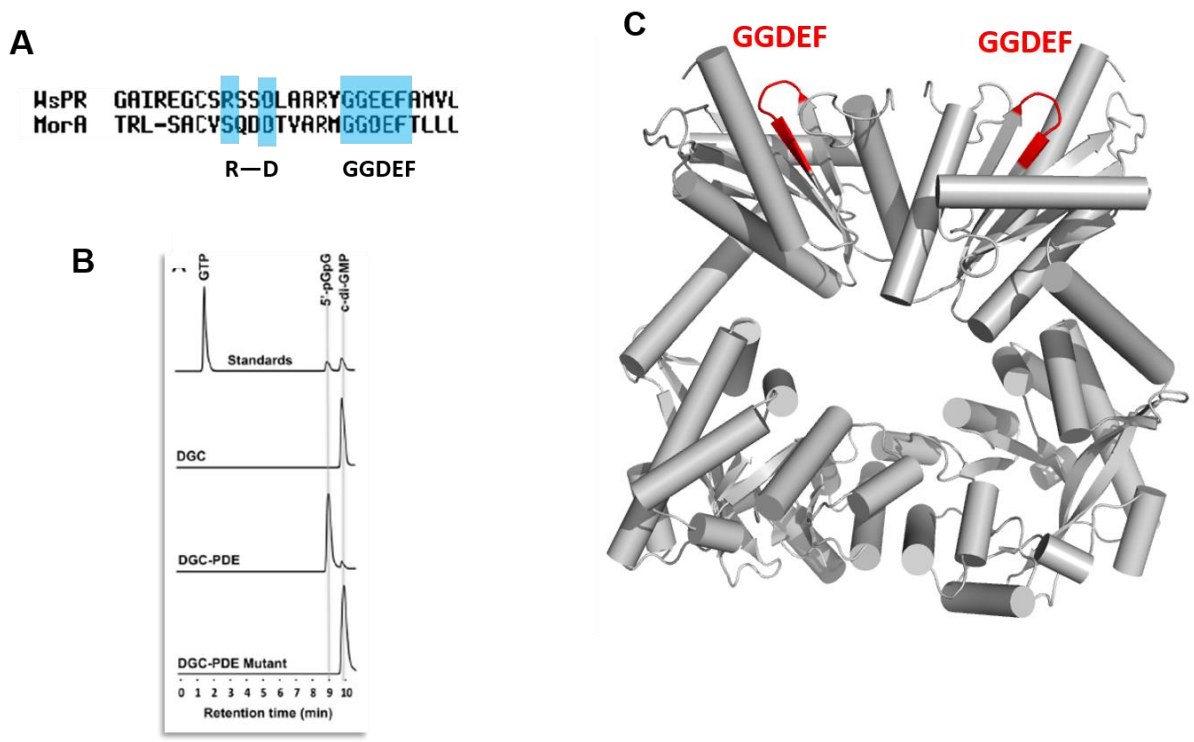


Figura 31 - Análise da sequência de aminoácidos e estrutura cristalográfica da MorA de acordo com a atividade catalítica. A) Alinhamento da sequência de MorA e WspR de Pseudomonas aeruginosa, mostrando a ausência do sítio inibitório que é dado pelo motivo RxxD, destacado em azul, cinco resíduos a esquerda do motivo GGDEF. B) Segundo Phippen e colaboradores, a atividade catalítica de diguanilato ciclase ocorre em construções que apresentam somente o domínio GGDEF ou simultaneamente de acordo com a construção que apresenta os domínios GGDEF-EAL, o que não é justificado por (C), já que quando o domínio EAL encontra-se na forma dimérica, os domínios GGDEF estão em disposições opostas, o que não permitem sua atividade catalítica.

Fonte: Elaborada pelo autor (A e C); (B) Adaptado de PHIPPEN.[11]

Phippen e colaboradores mostraram em seu trabalho que construções solúveis de MorA eram monoméricas em solução. Na presença de c-di-GMP todas as construções contendo o domínio EAL dimerizavam, mostrando que a interação desse nucleotídeo com o domínio EAL induz a formação do dímero cataliticamente ativo. Essa estabilização induzida pela interação com c-di-GMP já foi vista em outras fosfodiesterases ativas e também em LapD, que no caso assume uma conformação dimérica remanescente de seus homólogos, mas incapaz de hidrolisar o nucleotídeo.

No caso de MorA, a transição monômero-dímero mediada por c-di-GMP promove mudanças conformacionais na hélice α5[10], estabilizando a forma dimérica ao mesmo tempo que aproxima do sítio ativo os resíduos de ácido aspártico essências para coordenação dos co-fatores metálicos (Figura 32). Comparações com a hélice α5 do monômero e dímero da MorA, mostram que na estrutura da STM3615 esta hélice tem uma orientação mais parecida com a do monômero inativo de MorA. O fato de STM3615 encontrar-se na forma dimérica em solução e a estrutura mostrar um sítio ativo pronto para interagir e hidrolisar c-di-GMP (Figura 32C) demonstra que a regulação da atividade fosfodiesterase de MorA não aplica-se em STM3615, sugerindo um mecanismo alternativo ao até mesmo ausência de controle regulatório intrínseco em STM3615.

Com base na análise funcional de MorA e os resultados apresentados nesse trabalho, propomos um mecanismo de regulação enzimática para proteínas GGDEF-EAL bifuncionais. Tomando MorA como exemplo, quando a concentração local de c-di-GMP é baixa ao redor da enzima, os domínios EAL do dímero (hipoteticamente estabilizado por regiões intra-membrana e domínios PAS) não estão interagindo. Nesse estado, os sítios ativos dos domínios GGDEF não estariam travados em uma conformação inativa, como vista nas estruturas, e poderiam interagir de forma competente para síntese de c-di-GMP, como em PleD ou WspR, por exemplo. Como consequência, os níveis locais de c-di-GMP devem aumentar com o tempo, aumentando também a chance de interação do nucleotídeo com os domínios EAL. Isso levaria automaticamente a dimerização desses promovendo a forma dimérica cataliticamente ativa de EAL, ao mesmo tempo de induz uma conformação inativa para os domínios GGDEF da proteína. Nesse mecanismo é interessante notar que as atividades ciclase e fosfodiesterase são excludentes e devem manter os níveis de

c-di-GMP locais sob estrita regulação. Nossa observação que as enzimas GGDEF-EAL bifuncionais apresentam conservação do sítio I do domínio GGDEF muito menor que em enzimas contendo somente GGDEF é coerente com essa proposta de mecanismo, uma vez que ela impõe automaticamente uma regulação da produção de c-di-GMP sem a necessidade desse sítio inibitório alostérico (Figura 33).

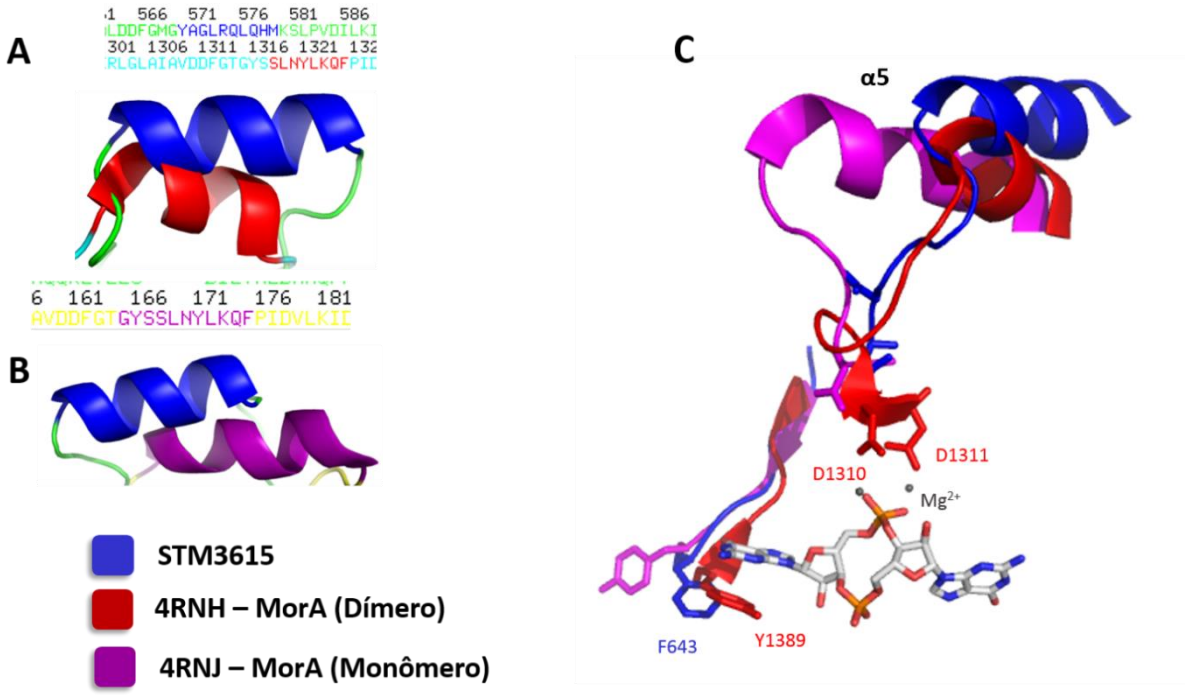


Figura 32 - Comparação das hélices α5 do monômero e dímero da MorA e STM3615. A) e B) Se comparada às hélices α5 do monômero (lilás) e dímero da MorA (vermelho), a estrutura da STM3615 (azul) apresenta essa hélice mais próxima ao monômero da MorA, mesmo apresentando o dímero na estrutura cristalográfica obtida. C) Se compararmos uma região um pouco maior que apenas as hélices, notamos que os resíduos da STM3615 estão em disposição nas quais poderiam coordenar os dois metais necessários para promover a hidrólise do c-di-GMP e a fenilalanina que coordena o c-di-GMP no sítio catalítico encontra-se em posição intermediária à tirosina do monômero e dímero da MorA.

Fonte: Elaborada pelo autor.

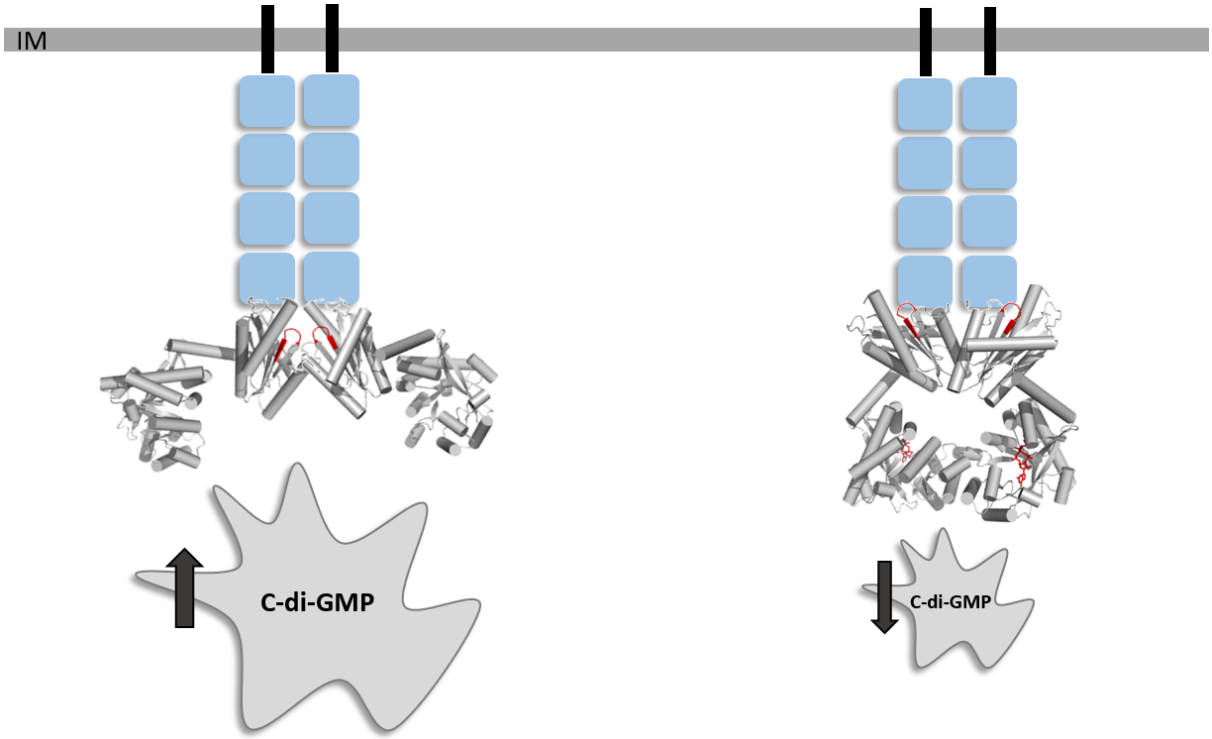


Figura 33 - Proposta de mecanismo para MorA. Em baixos níveis de c-di-GMP, os domínios GGDEF encontram-se na forma dimérica com seus sítios catalíticos com arranjo face-a-face, inativando a fosfodiesterase e contribuindo para o aumento do pool de c-di-GMP local (esquerda). Já em altos níveis de c-di-GMP, a fosfodiesterase toma a forma dimérica, inativando a fosfodiesterase, auxiliando na regulação do pool de c-di-GMP local (direita).

Fonte: Elaborada pelo autor.

4.8 Hipótese funcional para STM3615 baseada em sua estrutura e função

Diferentemente de MorA e potencialmente da maioria das enzimas bifuncionais, STM3615 encontra-se na forma dimérica com os domínios EAL pronto para catálise. Por outro lado, como LapD, essa proteína possui um domínio periplasmático dimérico e de função desconhecida. É honesto supor que esse domínio receba algum sinal para promover mudanças morfológicas, como visto experimentalmente, através da modulação de níveis internos de c-di-GMP. No entanto, nosso trabalho mostrou que STM3615 não sofre mudanças conformacionais intra-domínios GGDEF-EAL, como as observadas para LapD quando esta interage com LapG no periplasma, levantando a pergunta de como STM3615 exerce sua função regulatória? A presença domínio de sinalização HAMP logo após sua porção transmembranar de STM3615, de forma bem similar a LapD, levanta outra pergunta a respeito de como um eventual sinal no periplasma poderia propagar para mudanças conformacionais no domínio HAMP e qual o efeito disso?

Considerando-se que o c-di-GMP tem alvos locais, ou seja, o c-di-GMP produzido por uma proteína não se espalha por toda a célula, mas atinge alvos específicos, e que a STM3615 deve ser uma das principais responsáveis por manter o *pool* de c-di-GMP em uma determinada região, que em hipótese, outras proteínas podem interagir com a STM3615, para conjuntamente controlar o *pool* local de c-di-GMP. Esse tipo de interação tem sido cada vez mais observada em diferentes organismos e contextos, por exemplo, Basu e Sauer[81] mostraram que uma determinada diguanilato ciclase precisa da interação de uma fosfodiesterase inativa, uma proteína sensora, uma chaperona e uma protease para estabelecer seu mecanismo funcional, onde um estímulo recebido pela ciclase transmembranar leva a sinalização para as demais proteínas, variando então o *pool* de c-di-GMP.[81]

Em um trabalho realizado pela aluna Andrea Rodrigues Cardoso, no grupo do Prof. Marcos Navarro, verificou-se através de técnicas de duplo híbrido que duas proteínas transmembranares contendo domínios GGDEF ativos interagem através do domínio HAMP de uma e uma porção N-terminal conservada de outra. [82] Nesse trabalho foi proposto que o sinal recebido pela proteína contendo HAMP através de seu domínio periplasmático alteraria a conformação dos domínios HAMP, sem

alteração de atividade catalítica, quebrando a interação com a proteína parceira. Como essa parceira também é uma ciclase ativa, esse evento seria capaz de mudar as concentrações locais de c-di-GMP, sem alterar atividades catalíticas. Dentro deste contexto e das análises mostradas neste trabalho, podemos estabelecer uma hipótese de mecanismo funcional análoga para a STM3615 (Figura 34), no qual inicialmente a STM3615, junto a outras diguanilato ciclases e fosfodiesterases manteriam um *pool* de c-di-GMP local estável. O recebimento de algum estímulo esterno através do domínio periplasmático propagaria o sinal através do domínio HAMP, resultando no recrutamento ou dissociação de outras enzimas. Esse evento fatalmente alteraria o *pool* local de c-di-GMP, desencadeando a observada dispersão de células que se encontram no biofilme, um processo reconhecidamente regulado por c-di-GMP.[49]

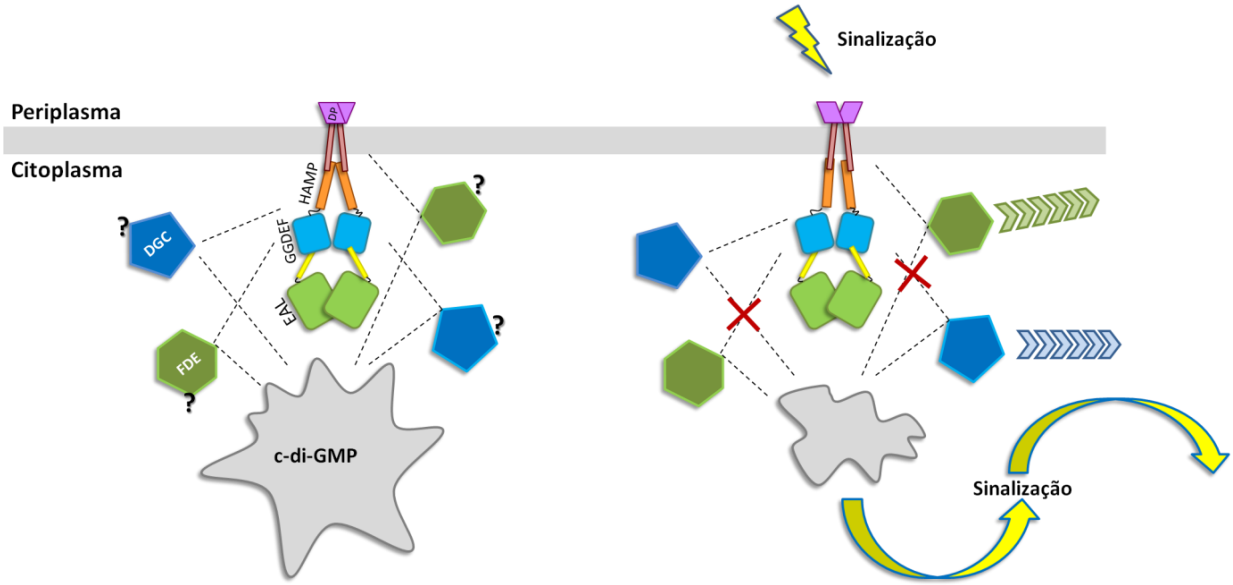


Figura 34 - Hipótese de mecanismo de ação da STM3615. A STM3615, através de interações com outras proteínas como diguanilato ciclases (DGC) e fosfodiesterases (FDE) controlam o pool de c-di-GMP constante. Um estímulo no domínio periplasmático é transmitido através do domínio HAMP, alterando o pool de c-di-GMP, através do deslocamento de algumas das proteínas envolvidas no mecanismo, diminuindo a concentração local do nucleotídeo e sinalizando a dispersão do biofilme bacteriano.

Fonte: Elaborada pelo autor.

5 CONCLUSÃO

As construções clonadas e protocolos de expressão e purificação determinados com sucesso, possibilitaram os ensaios de cristalização. Através de diversas tentativas e refinamentos de condições de cristalização, obtivemos cristais adequados para determinar a estrutura da construção GGDEF_EAL$^{242-668}$, em que o domínio GGDEF encontra-se em um enovelamento do tipo sanduíche α-β de duas camadas e o domínio EAL do tipo barril α-β. A estrutura com os dois domínios apresentou-se na forma dimérica e ativa, como sugerido pelos experimentos de cromatografia de exclusão molecular analítica e SAXS. Como era esperado para o domínio EAL, os resíduos catalíticos foram identificados e estavam totalmente conservados e os ensaios de atividade corroboram com estes dados, mostrando que o domínio EAL encontra-se em sua forma ativa, tendo 5'pGPG como produto da hidrólise do c-di-GMP.

Comparativamente, proteínas com arquitetura próxima a STM3615, com domínios GGDEF degenerados e EAL degenerados ou conservados se diferenciam pela hélice que liga esses domínios, onde aquelas que tem o EAL ativo apresentam hélices maiores que aquelas com EAL degenerados, levando a uma conformação mais estendida da proteína.

Com estes dados, um mecanismo foi proposto, onde a STM3615, juntamente com outras proteínas controlam o *pool* de c-di-GMP, que pode ser alterado por um estímulo no domínio periplasmático que transmite através do HAMP a sinalização para as demais proteínas, diminuindo o *pool* de c-di-GMP, levando a dispersão do biofilme bacteriano.

Poderíamos ainda extrapolar os dados apresentados para o mecanismo do único *dual domain* que não temos disponível dados na literatura ou estrutura, o caso em que temos o domínio EAL degenerado e GGDEF ativo. Nesta hipótese, o domínio GGDEF, na ausência do motivo inibitório RxxD e hélice entre os domínios GGDEF-EAL na forma mais curta, que possibilita maior flexibilidade, pode ter um mecanismo similar ao de proteínas como a LapD, em que um aumento nos níveis de c-di-GMP local poderia promover a ligação do nucleotídeo ao receptor EAL, gerando

mudanças conformacionais, regulando a atividade do domínio GGDEF e o *pool* de c-di-GMP.

Os dados e conclusões deste trabalho trazem uma grande contribuição para os estudos de proteínas com GGDEF-EAL *in tandem*, com a estrutura inédita dos domínios solúveis da STM3615 e propostas de mecanismos para os grupos onde apenas o EAL ou GGDEF são ativos, fechando o arco de proteínas com características semelhantes, aumentando as chances futuras de controle da formação de biofilme bacteriano e consequentemente de infecções causadas por patógenos oportunistas.

REFERÊNCIAS

1 O'TOOLE, G.; KAPLAN, H. B.; KOLTER, R. Biofilm formation as microbial development. **Annual Review of Microbiology**, v. 54, p. 49-79, Oct. 2000. doi: 10.1146/annurev.micro.54.1.49

2 HALL-STOODLEY, L.; COSTERTON, J. W.; STOODLEY, P. Bacterial biofilms: from the natural environment to infectious diseases. **Nature Reviews Microbiology**, v. 2, n. 2, p. 95-108, Feb. 2004

3 HULKO, M. et al. The HAMP domain structure implies helix rotation in transmembrane signaling. **Cell**, v. 126, n. 5, p. 929-940, Sept. 2006.

4 AIROLA, M. V. et al. Structure of concatenated HAMP domains provides a mechanism for signal transduction. **Structure**, v. 18, n. 4, p. 436-448, Mar. 2010.

5 AHMAD, I. et al. Complex c-di-GMP signaling networks mediate transition between virulence properties and biofilm formation in Salmonella enterica serovar Typhimurium. **PLoS One**, v. 6, n. 12, p. e28351, Dec. 2011. doi: 10.1371/journal.pone.0028351.

6 NEWELL, P. D.; MONDS, R. D.; O'TOOLE, G. A. LapD is a bis-(3',5')-cyclic dimeric GMP-binding protein that regulates surface attachment by Pseudomonas fluorescens Pf0-1. **Proceedings of the National Academy of Sciences of the United States of America**, v. 106, n. 9, p. 3461-3466, Feb. 2009.

7 HINSA, S. M. et al. Transition from reversible to irreversible attachment during biofilm formation by Pseudomonas fluorescens WCS365 requires an ABC transporter and a large secreted protein. **Molecular Microbiology**, v. 49, n. 4, p. 905-918, Aug. 2003.

8 HINSA, S. M.; O'TOOLE, G. A. Biofilm formation by Pseudomonas fluorescens WCS365: a role for LapD. **Microbiology**, v. 152, Pt 5, p. 1375-1383, May 2006.

9 NEWELL, P. D. et al. A c-di-GMP effector system controls cell adhesion by inside-out signaling and surface protein cleavage. **PLoS Biology**, v. 9, n. 2, p. e1000587, Feb. 2011.

10 PHIPPEN, C. W. et al. Formation and dimerization of the phosphodiesterase active site of the Pseudomonas aeruginosa MorA, a bi-functional c-di-GMP regulator. **FEBS Letters**, v. 588, n. 24, p. 4631-4636, Dec. 2014.

11 TAMAYO, R.; PRATT, J. T.; CAMILLI, A. Roles of cyclic diguanylate in the regulation of bacterial pathogenesis. **Annual Review of Microbiology**, v. 61, p. 131-148, Oct. 2007. doi: 10.1146/annurev.micro.61.080706.093426.

12 RAO, F. et al Catalytic mechanism of cyclic di-GMP-specific phosphodiesterase: a study of the EAL domain-containg RocR from Pseudomonas aeruginosa. **Journal of Bacteriology**, v. 190, n. 10, p. 3622-3631, May 2008.

13 PARSEK, M. R.; SINGH, P. K. Bacterial biofilms: an emerging link to disease pathogenesis. **Annual Reviews Microbiology**, v. 57, p. 677-701, 2009. doi: 10.1146/annurev.micro.57.030502.090720.

14 HENGGE, R. Principles of c-di-GMP signalling in bacteria. **Nature Reviews Microbiology**, v. 7, n. 4, p. 263-273, Apr. 2009.

15 FUX, C. A. et al. Survival strategies of infectious biofilms. **Trends in Microbiology**, v. 13, n. 1, p. 34-40, Jan. 2005.

16 ROSS, P.; MAYER, R.; BENZIMAN, M. Cellulose biosynthesis and function in bacteria. **Microbiology Reviews**, v. 55, n. 1, p. 35-58, Mar. 1991.

17 JENAL, U.; MALONE, J. Mechanisms of cyclic-di-GMP signaling in bacteria. **Annual Review of Genetics**, v. 40, p. 385-407, Dec. 2006. doi: 10.1146/annurev.genet.40.110405.090423.

18 WOLFE, A. J.; VISICK, K. L. Get the message out: cyclic-Di-GMP regulates multiple levels of flagellum-based motility. **Journal of Bacteriology**, v. 190, n. 2, p. 463-475, Jan. 2008.

19 KIM, Y. K.; MCCARTER, L. L. ScrG, a GGDEF-EAL protein, participates in regulating swarming and sticking in Vibrio parahaemolyticus. **Journal of Bacteriology**, v. 189, n. 11, p. 4094-4107, June 2007.

20 SIMM, R. et al. GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. **Molecular Microbiology**, v. 53, n. 4, p. 1123-1134, Aug. 2004.

21 RYJENKOV, D. A. et al. The PilZ domain is a receptor for the second messenger c-di-GMP: the PilZ domain protein YcgR controls motility in enterobacteria. **Journal of Biological Chemistry**, v. 281, n. 41, p. 30310-30314, Oct. 2006.

22 RAO, F. et al. Catalytic mechanism of cyclic di-GMP-specific phosphodiesterase: a study of the EAL domain-containing RocR from Pseudomonas aeruginosa. **Journal of Bacteriology**, v. 190, n. 10, p. 3622-3631, May 2008. doi: 10.1128/JB.00165-08.

23 AUSMEES, N. et al. Genetic data indicate that proteins containing the GGDEF domain possess diguanylate cyclase activity. **FEMS Microbiology Letters**, v. 204, n. 1, p. 163-167, Oct. 2001.

24 TISCHLER, A. D.; CAMILLI, A. Cyclic diguanylate regulates Vibrio cholerae virulence gene expression. **Infection and Immunity**, v. 73, n. 9, p. 5873-5882, Sept. 2005.

25 ALDRIDGE, P. et al. Role of the GGDEF regulator PleD in polar development of Caulobacter crescentus. **Molecular Microbiology**, v. 47, n. 6, p. 1695-1708, Mar. 2003.

26 PEI, J.; GRISHIN, N. V. GGDEF domain is homologous to adenylyl cyclase. **Proteins**, v. 42, n. 2, p. 210-216, Feb. 2001.

27 LINDER, J. U. Class III adenylyl cyclases: molecular mechanisms of catalysis and regulation. **Cellular and Molecular Life Sciences,** v. 63, n. 15, p. 1736-1751, Aug. 2006.

28 SINHA, S. C.; SPRANG, S. R. Structures, mechanism, regulation and evolution of class III nucleotidyl cyclases. **Reviews of Physiology, Biochemistry and Pharmacology**, v. 157, p. 105-140, Sept. 2006. doi: 10.1007/112_0603.

29 HICKMAN, J. W.; TIFREA, D. F.; HARWOOD, C. S. A chemosensory system that regulates biofilm formation through modulation of cyclic diguanylate levels. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, n. 40, p. 14422-14427, Oct. 2005.

30 GALPERIN, M. Y.; NIKOLSKAYA, A. N.; KOONIN, E. V. Novel domains of the prokaryotic two-component signal transduction systems. **FEMS Microbiology Letters**, v. 203, n. 1, p. 11-21, Sept. 2001.

31 RYJENKOV, D. A. et al. Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: insights into biochemistry of the GGDEF protein domain. **Journal of Bacteriology**, v. 187, n. 5, p. 1792-1798, Mar. 2005.

32 PAUL, R. et al. Activation of the diguanylate cyclase PleD by phosphorylation-mediated dimerization. **Journal of Biological Chemistry**, v. 282, n. 40, p. 29170-29177, Oct. 2007.

33 CHAN, C. et al. Structural basis of activity and allosteric control of diguanylate cyclase. **Proceedings of the National Academy of Sciences of the United States of America**, v. 101, n. 49, p. 17084-17089, Dec. 2004.

34 WASSMANN, P. et al. Structure of BeF3- -modified response regulator PleD: implications for diguanylate cyclase activation, catalysis, and feedback inhibition. **Structure**, v. 15, n. 8, p. 915-927, Aug. 2007.

35 BARENDS, T. R. et al. Structure and mechanism of a bacterial light-regulated cyclic nucleotide phosphodiesterase. **Nature**, v. 459, n. 7249, p. 1015-1018, June 2009.

36 ROSS, P. et al. Control of cellulose synthesis in Acetobacter xylinum. A unique guanyl oligonucleotide is the immediate activator of the cellulose synthase. **Carbohydrate Research,** v. 149, p. 101-117, June 1986. doi: 10.1016/S0008-6215(00)90372-0.

37 GALPERIN, M. Y. Structural classification of bacterial response regulators: diversity of output domains and domain combinations. **Journal of Bacteriology**, v. 188, n. 12, p. 4169-4182, June 2006.

38 TCHIGVINTSEV, A. et al. Structural insight into the mechanism of c-di-GMP hydrolysis by EAL domain phosphodiesterases. **Journal of Molecular Biology**, v. 402, n. 3, p. 524-538, Sept. 2010.

39 BAI, Y. et al. Mycobacterium tuberculosis Rv3586 (DacA) is a diadenylate cyclase that converts ATP or ADP into c-di-AMP. **PLoS One**, v. 7, n. 4, p. e35206, Apr. 2012. doi: 10.1371/journal.pone.0035206.

40 JOHNSON, J. G. et al. Type 3 fimbriae and biofilm formation are regulated by the transcriptional regulators MrkHI in Klebsiella pneumoniae. **Journal of Bacteriology**, v. 193, n. 14, p. 3453-3460, July 2011.

41 CHRISTEN, M. et al. Identification and characterization of a cyclic di-GMP-specific phosphodiesterase and its allosteric control by GTP. **Journal of Biological Chemistry**, v. 280, n. 35, p. 30829-30837, Sept. 2005.

42 RÖMLING, U.; GALPERIN, M. Y.; GOMELSKY, M. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. **Microbiology and Molecular Biology Reviews**, v. 77, n. 1, p. 1-52, Mar. 2013.

43 RYAN, R. P. et al. Cyclic di-GMP signaling in bacteria: recent advances and new puzzles. **Journal of Bacteriology**, v. 188, n. 24, p. 8327-8234, Dec. 2006.

44 NAVARRO, M. V. et al. Structural basis for c-di-GMP-mediated inside-out signaling controlling periplasmic proteolysis. **PLoS Biology**, v. 9, n. 2, p. e1000588, Fev. 2011.

45 CHOY, W. K. et al. MorA defines a new class of regulators affecting flagellar development and biofilm formation in diverse Pseudomonas species. **Journal of Bacteriology**, v. 186, n. 21, p. 7221-7228, Nov. 2004. doi: 10.1128/JB.186.21.7221-7228.2004.

46 HUANG, B.; WHITCHURCH, C. B.; MATTICK, J. S. FimX, a multidomain protein connecting environmental signals to twitching motility in Pseudomonas aeruginosa. **Journal of Bacteriology**, v. 185, n. 24, p. 7068-7076, Dec. 2003.

47 KAZMIERCZAK, B. I.; LEBRON, M. B.; MURRAY, T. S. Analysis of FimX, a phosphodiesterase that governs twitching motility in Pseudomonas aeruginosa. **Molecular Microbiology**, v. 60, n. 4, p. 1026-1043, May 2006.

48 NAVARRO, M. V. et al. Structural analysis of the GGDEF-EAL domain-containing c-di-GMP receptor FimX. **Structure**, v. 17, n. 8, p. 1104-1116, Aug. 2009.

49 ANWAR, N. et al. Modulation of biofilm-formation in Salmonella enterica serovar Typhimurium by the periplasmic DsbA/DsbB oxidoreductase system requires the GGDEF-EAL domain protein STM3615. **PLoS One**, v. 9, n. 8, p. e106095, Aug. 2014. doi: 10.1371/journal.pone.0106095.

50 GUZZO, C. R. et al. Structure of the PilZ-FimXEAL-c-di-GMP complex responsible for the regulation of bacterial type IV Pilus Biogenesis. **Journal of Molecular Biology**, v. 425, n. 12, p. 2174-2197, June 2013.

51 MÖGLICH, A.; AYERS, R. A.; MOFFAT, K. Structure and signaling mechanism of Per-ARNT-Sim domains. **Structure**, v. 17, n. 10, p. 1282-1294, Oct. 2009.

52 KUCHMA, S. L.; CONNOLLY, J. P.; O'TOOLE, G. A. A three-component regulatory system regulates biofilm maturation and type III secretion in Pseudomonas aeruginosa. **Journal of Bacteriology**, v. 187, n. 4, p. 1441-1454, Feb. 2005.

53 MEISSNER, A. et al. Pseudomonas aeruginosa cupA-encoded fimbriae expression is regulated by a GGDEF and EAL domain-dependent modulation of the intracellular level of cyclic diguanylate. **Environmental Microbiology**, v. 9, n. 10, p. 2475-2485, Oct. 2007.

54 KROGH, A. et al. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. **Journal of Molecular Biology**, v. 305, n. 3, p. 567-580, Jan. 2001.

55 FINN, R. D. et al. The Pfam protein families database: towards a more sustainable future. **Nucleic Acids Research**, v. 44, n. D1, p. D279-D285, Jan. 2016.

56 BUCHAN, D. W. et al. Scalable web services for the PSIPRED protein analysis Workbench. **Nucleic Acids Research**, v. 41, p. W349-W357, July 2013. doi: 10.1093/nar/gkt381.

57 ZUO, X. et al. Expression and purification of SARS coronavirus proteins using SUMO-fusions. **Protein Expression and Purification**, v. 42, n. 1, p. 100-110, July 2005.

58 GLATTER, O.; KRATKY, O. **Small angle x-ray scattering**. London: Academic Press 1982.

59 PETOUKHOV, M. V. et al. New developments in the ATSAS program package for small-angle scattering data analysis. **Journal Applied Crystallography**, v. 45, Pt 2, p. 342-350, Apr. 2012.

60 SVERGUN, D. I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. **Journal Applied Crystallography**, v. 25, p. 495-503, Aug. 1992. doi: 10.1107/S0021889892001663.

61 PETOUKHOV, M. V.; SVERGUN, D. I. Joint use of small-angle X-ray and neutron scattering to study biological macromolecules in solution. **European Biophysics Journal**, v. 35, n. 7, p. 567-576, Sept. 2006.

62 MYLONASA, E.; SVERGUN, D. I. Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. **Journal of Applied Crystallography**, v. 40, p. 245-249, Apr. 2007. doi: 10.1107/S002188980700252X.

63 FISCHER, H. et al. The molecular weight of proteins in solution can be determined from a singls SAXS measurement on a relative scale. **Journal Applied Crystallography**, v. 43, Pt. 1, p. 101-109, Feb. 2010. doi: 10.1107/S0021889809043076.

64 SVERGUN, D. I. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. **Biophysical Journal**, v. 76, n. 6, p. 2879-2886, June 1999.

65 KONAREV, P.; PETOUKHOV, M.; SVERGUN, D. MASSHA - a graphics system for rigid-body modelling of macromolecular complexes against solution scattering data. **Journal of Applied Crystallography**, v. 34, Pt. 4, p. 527-532, Aug. 2001. doi:10.1107/S0021889801006100.

66 SVERGUN, D. I.; BARBERATO, C.; KOCH, M. H. J. Crysol - a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. **Journal Applied Crystallography**, v. 28, p. 768-773, Dec. 1995. doi:10.1107/S0021889895007047.

67 RAO, F. et al. The functional role of a conserved loop in EAL domain-based cyclic di-GMP-specific phosphodiesterase. **Journal of Bacteriology**, v. 191, n. 15, p. 4722-4731, Aug. 2009.

68 GORREC, F. The MORPHEUS protein crystallization screen. **Journal of Applied Crystallography**, v. 42, Pt 6, p. 1035-1042, Dec. 2009. doi: 10.1107/S0021889809042022.

69 BERMAN, H. M. et al. The protein data bank and the challenge of structural genomics. **Nature Structural & Molecular Biology**, v. 7, p. 957-959, Nov. 2000. doi: 10.1038/80734

70 GUIMARÃES, B. G. et al. The MX2 macromolecular crystallography beamline: a wiggler X-ray source at the LNLS. **Journal Synchrotron Radiation**, v. 16, Pt 1, p. 69-75, Jan. 2009

71 KABSCH, W. XDS. **Acta Crystallographica D**, v. 66, Pt 2, p. 125-132, Feb. 2010.

72 SHELDRICK, G. M. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. **Acta Crystallographica D**, v. 66, Pt 4, p. 479-485, Apr. 2010.

73 ADAMS, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. **Acta Crystallographica D**, v. 66, Pt 2, p. 213-221, Feb. 2010.

74 EMSLEY, P.; COWTAN, K. Coot: model-building tools for molecular graphics. **Acta Crystallographica D**, v. 60, Pt 12, p. 2126-2132, Dec. 2004.

75 CHOU, S. H.; GALPERIN, M. Y. Diversity of Cyclic Di-GMP-binding proteins and mechanisms. **Journal of Bacteriology**, v. 198, n. 1, p. 32-46, Jan. 2016.

76 GALPERIN, M. Y.; HIGDON, R.; KOLKER, E. Interplay of heritage and habitat in the distribution of bacterial signal transduction systems. **Molecular BioSystems**, v. 6, n. 4, p. 721-728, Apr. 2010.

77 COWTAN, K. Joint CCP4 and ESF-EACBM. **Newsletter on Protein Crystallography**. v. 31, p. 34-38, 1994.

78 MATTHEWS, B. W. Solvent content of protein crystals. **Journal of Molecular Biology**, v. 33, n. 2, p. 491-497, Apr. 1968.

79 CROOKS, G. E. et al. WebLogo: a sequence logo generator. **Genome Research**, v. 14, n. 6, p. 1188-1190, June 2004.

80 GOLDENBERG, O. et al. The ConSurf-DB: pre-calculated evolutionary conservation profiles of protein structures. **Nucleic Acids Research**, v. 37, p. D323-D327, Jan. 2009. doi: 10.1093/nar/gkn822.

81 BASU ROY, A.; SAUER, K. Diguanylate cyclase NicD-based signalling mechanism of nutrient-induced dispersion by Pseudomonas aeruginosa. **Molecular Microbiology**, v. 94, n. 4, p. 771-793, Nov. 2014.

82 CARDOSO, A. R. **Mapeamento global de interações proteicas na via de sinalização mediadas por c-di-GMP em Pseudomonas aeruginosa**. 2016. 107 p. Dissertação (Mestrado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2016.