UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

ANDREA RODRIGUES CARDOSO

Mapeamento global de interações proteicas nas vias de sinalização mediadas por
c-di-GMP de *Pseudomonas aeruginosa*

São Carlos
2016

ANDREA RODRIGUES CARDOSO

Mapeamento global de interações proteicas nas vias de sinalização mediadas por
c-di-GMP de *Pseudomonas aeruginosa*

Dissertação 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 Mestre em Ciências.

Área de concentração: Física Aplicada

Opção: Física Biomolecular

Orientador: Prof. Dr. Marcos Vicente de Albuquerque Salles Navarro.

Versão Corrigida

(Versão original disponível na Unidade que aloja o Programa)

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)

Cardoso, Andrea Rodrigues

Mapeamento global de interações proteicas nas vias de sinalização mediadas por c-di-GMP de *Pseudomonas aeruginosa* / Andrea Rodrigues Cardoso; orientador Marcos Vicente de Albuquerque Salles Navarro - versão corrigida -- São Carlos, 2016.
107 p.

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

1. Biofilmes bacterianos. 2. Guanosina monofosfato (3'-5')-cíclica dimérica. 3. Interação proteica. I. Navarro, Marcos Vicente de Albuquerque Salles, orient. II. Título.

Aos meus pais, Míriam e Pedro, não somente pelo
aconselhamento, apoio e carinho constantes, mas também
por admirá-los imensamente.

AGRADECIMENTOS

Agradeço, por tudo e antes de tudo, a Deus.

Agradeço aos meus pais, Míriam e Pedro, por todo o amor e carinho de sempre, pelo apoio incansável, por serem minha fonte de força, alegria e inspiração nos momentos de dificuldade. Por acreditarem em mim, mesmo quando eu não acredito. Por serem pessoas extraordinárias. Não há como expressar o quanto sou grata por tê-los como pais.

Agradeço ao meu irmão, Marcelo, pelo apoio, às vezes velado, mas sempre presente, principalmente nos momentos de dificuldade. Agradeço também por ter me deixado ficar na sua casa sempre que precisei para concluir este trabalho.

Agradeço à minha tia Márcia por todo o apoio, carinho e cumplicidade. Por ser tia, mas valer por mais uma mãe!

Agradeço às minhas avós Myrna e Fernanda. À minha avó Myrna, por sua espontaneidade, que sempre me diverte, e à minha avó Fernanda, pelo cuidado e preocupação. Agradeço ao meu avô Pedro, que, acredito, intercede pela nossa família sempre que algo nos aflige.

Agradeço às minhas primas Daniele e Sabrina pelo apoio, carinho e compreensão. Por isso também, agradeço aos meus tios Antônio e Telma.

Agradeço infinitamente aos grandes amigos que fiz no laboratório, que estiveram sempre do meu lado, que me ajudaram e me animaram. Com eles, tantas vezes, o estresse virou risada e o desalento foi curado pelo companheirismo. Muito obrigada, Bruno, Éverton, Flávio, Helton, Juliana, Naiara e Nathalya!

Agradeço à minha amiga Raissa, agora irmã de coração, pelo ano incrível que passamos juntas e por ter me estimulado a encarar os eventos da vida com mais leveza, alegria e amor. Sou realmente muito grata por ter não ter desperdiçado a chance de dividir esse ano com você.

Tenho que agradecer novamente às minhas grandes amigas Naiara e Raissa por me acudirem nas horas de aperto! Incondicionalmente.

À Naiara, agradeço muito ainda por ter me ajudado e me acolhido logo quando entrei neste grupo de pesquisa.

Agradeço aos amigos de longa data Bruno Luan, Caio, Evandro, Luísa e Priscilla, com quem compartilhei grandes momentos e sempre me apoiaram, apesar da distância.

Agradeço ao meu orientador, Prof. Dr. Marcos Navarro, por ter me orientado neste trabalho. Agradeço, em especial, por sua compreensão e por sua paciência neste último ano e, mais ainda, nesta etapa final.

Agradeço ao Ricardo, do Serviço de Pós-Graduação, por sua imensa ajuda por tantas vezes! Agradeço também por ser sempre tão simpático, prestativo e assertivo ao nos atender.

Agradeço aos amigos Atílio e Laís. Obrigada por ter cedido seus móveis, Atílio! E, Laís, muito obrigada por ter cedido a sua casa para que eu pudesse continuar os meus experimentos!

Agradeço aos amigos que fiz entre 2014 e 2015 na Braskem por todos os momentos de alegria, descontração e diversão, que foram tão importantes e marcantes para mim. Obrigada, Davi, Denise, Diógenes, Ellen Paula, Ellen Aquino, Fernando, Flávia, Hellen, Johana, João, Julianna, Kelin, Marcos, Mariana, Marilene, Mateus, Monique, Nemailla, Paola, Rafael Leite, Rafael Melo, Rafaela, Renata, Tálisson e Verônica. À Verônica, agradeço também a compreensão, a ajuda e a tutoria.

Agradeço também aos amigos Débora, Leandro e Marcela pelo apoio e carinho.

Agradeço às técnicas Andressa e Isabel por todo o apoio, auxílio e disponibilidade. Agradeço, em especial, à Andressa pela enorme ajuda com os inúmeros sequenciamentos.

Agradeço à Neusa e à Cristina pela simpatia de sempre e por terem revisado esta dissertação em curto prazo para que eu pudesse depositá-la. Agradeço muito também ao pessoal da gráfica, Italo, Denise e Cristovão.

Agradeço ao colega Raphael pelo grande auxílio na etapa final dos experimentos. Agradeço também ao colega Wesley pelo auxílio no início deste trabalho.

Agradeço ao pesquisador Dr. César Moisés Camilo, cuja colaboração foi imprescindível para este trabalho, e também ao Prof. Dr. Igor Polikarpov por ter cedido seu laboratório para parte dos experimentos.

Agradeço aos funcionários da USP e, especialmente, aos funcionários do IFSC por seu trabalho, pela simpatia quase universal e, muitas vezes, até mesmo zelo por nós, alunos.

Agradeço aos professores do IFSC, por sua disposição em passar adiante o seu conhecimento e experiência.

Agradeço aos membros da Comissão de Pós-Graduação.

Agradeço à FAPESP (processo 2012/25313-9) e à CAPES pelo apoio financeiro.

Agradeço a todos os que, de alguma forma, colaboraram para o projeto.

Por fim, agradeço, de antemão, à banca examinadora pela disponibilidade para avaliar este trabalho.

“A experiência, a sabedoria, não é recebida; é preciso descobri-la por nós mesmos, depois de uma jornada que ninguém pode fazer por nós e da qual ninguém pode nos poupar.”

Marcel Proust

RESUMO

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

A persistência bacteriana correlacionada à formação de biofilmes bacterianos é, há algum tempo, fonte de grande preocupação médica em virtude de sua ampla associação com a dificuldade de tratamento de infecções crônicas. Por outro lado, as perspectivas de utilização de biofilmes bacterianos em novas aplicações biotecnológicas e até mesmo para fins terapêuticos são promissoras. Há, portanto, grande interesse em compreender os mecanismos que levam as células bacterianas a deixar o estado planctônico, de vida livre, e associarem-se nesses conglomerados celulares altamente complexos. Ao longo das últimas décadas, o segundo mensageiro c-di-GMP – em conjunto com as moléculas que catalisam sua síntese (diguanylate ciclases) e sua degradação (fosfodiesterases) e seus receptores – estabeleceu-se como um elemento central de regulação de uma série de respostas celulares que determinam a formação ou a dispersão de biofilmes. Curiosamente, as proteínas que participam do metabolismo deste segundo mensageiro estão, frequentemente, codificadas múltiplas vezes em um mesmo genoma bacteriano. Em vista dessa observação, estudos mais recentes apontam que, para reger paralelamente uma variedade tão ampla de fenótipos, este sistema opera em modo de alta especificidade de sinalização e que, portanto, o sinal metabolizado por determinados conjuntos de diguanylate ciclases e fosfodiesterases tem alvos celulares específicos. Evidências robustas, porém isoladas até o momento, apontaram que um dos meios pelo qual ocorre a segregação entre sinal produzido e alvo específico é a interação direta entre as proteínas componentes das vias de sinalização. Mais, demonstrou-se que, em algumas vias, a transmissão de sinal ocorre exclusivamente via interação proteica, dispensando a intermediação do sinalizador em si. Para avaliar a validade e relevância global deste mecanismo, propôs-se, neste estudo, a investigação da rede total de interações entre as proteínas tipicamente associadas às vias de sinalização por c-di-GMP em *Pseudomonas aeruginosa*, utilizando ensaios de duplo-híbrido bacteriano. Para tanto, foram construídas duas bibliotecas de DNA direcionadas e foram feitos testes de interação de forma estratégica para possibilitar o esgotamento e averiguação de todas as possíveis interações entre as proteínas alvo identificadas. O resultado obtido, um mapa inicial, porém abrangente, da rede

de interações proteicas em *P. aeruginosa*, indica uma grande probabilidade de que os mecanismos previamente descritos sejam realmente recorrentes e relevantes para o intermédio da sinalização nesse organismo. Algumas das interações mais robustas encontradas são bastante interessantes e serão, em estudos futuros, mais extensivamente estudadas.

Palavras-chave: Biofilmes bacterianos. Guanosina monofosfato (3'-5')-cíclica dimérica. Interação proteica.

ABSTRACT

CARDOSO, A. R. **Construction of a global map of protein-protein interactions in c-di-GMP signalling pathways of *Pseudomonas aeruginosa***. 2016. 107 p. Dissertação (Mestrado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2016.

Persister bacteria are correlated to biofilm formation and have been a source of great medical concern due to its close association with the impairment of traditional treatment in combating chronic infections. On the other hand, using bacterial biofilms to create original biotechnological applications or even as a means of therapeutic treatment in medical settings constitutes a promising prospect. There is, therefore, a great interest in understanding the mechanisms that allow bacteria to leave the free-living planktonic lifestyle and associate in these highly complex cellular aggregates. Over the last decades, the second messenger c-di-GMP – and also the molecules involved in its synthesis (diguanylate ciclases) and degradation (phosphodiesterases) along with its receptors – has been established as a key element implicated in regulation of a series of cellular responses that determine biofilm formation or dispersion. Curiously, the proteins that play a part in the metabolism of this second messenger are frequently coded multiple times in single bacterial genomes. Taking this into account, recent studies indicate that, in order to control such a wide range of phenotypes, this system operates via high specificity of signaling – which means that the signal metabolized by a certain set of diguanylate ciclases and phosphodiesterases has specific cellular targets. Robust but yet isolated evidence indicate that a means by which a signal is segregated with its correlated phenotypic response is through direct protein-protein interaction involving the components of these signaling pathways. Even more, there has been strikingly evidence that, in some of these pathways, signal transduction occurs exclusively through protein-protein interaction, entirely dismissing any mediation by the signal molecule. In order to validate and evaluate the global relevance of this type of mechanism, this study proposed the investigation of the entire network of interactions between proteins typically associated with c-di-GMP signaling pathways of *Pseudomonas aeruginosa* by employing bacterial two-hybrid system assays. To make that possible, two DNA libraries were constructed and interaction essays were performed in a strategic way so that all possibilities of interaction between target proteins were explored. The results obtained from these experiments allowed the construction of a broad map of interactions that, although still primitive, indicates that, chances are, the

mechanisms previously described are both recurrent and relevant to signaling regulation in this organism. Some of the interaction partners found are particularly interesting and will be further investigated in future studies.

Keywords: Bacterial biofilms. Bis-(3'-5')-cyclic dimeric guanosine monophosphate. Protein-protein interactions.

LISTA DE FIGURAS

Figura 1	- Modelos da matriz extracelular em diferentes escalas. São ilustrados: a) um biofilme aderido à superfície; b) os principais componentes do EPS; c) interações entre polissacarídeos, que contribuem para a estabilidade da estrutura e d) a modelagem molecular da interação entre um polissacarídeo (alginato) e uma lipase extracelular na matriz.	23
Figura 2	- Etapas de formação de um biofilme e de sua dispersão em células planctônicas.....	24
Figura 3	- A molécula de c-di-GMP..	28
Figura 4	- Estrutura dos sítios ativos de domínios envolvidos no metabolismo da molécula c-di-GMP em complexo com seus respectivos substratos. Em (a), é mostrado um análogo não hidrolisável de GTP (GTP α S) no sítio ativo do domínio GGDEF; em (b), a molécula de c-di-GMP no sítio ativo do domínio EAL. Os átomos de carbono dos substratos estão coloridos em cinza, assim como as glicinas do motivo GGDEF. Átomos catalíticos importantes estão coloridos em amarelo e indicados por setas. As esferas cor-de-rosa representam os íons metálicos Mg ²⁺	30
Figura 5	- Configuração do centro catalítico formado por três íon Fe e modelagem da molécula de c-di-GMP no sítio ativo do domínio HD-GYP da PDE <i>PmGH</i> . Em (a), é dado enfoque à coordenação dos íons Fe G e M e, em (b), à coordenação dos íons Fe M e H. Os íon Fe estão representados como esferas cor de laranja e foram rotulados G, H e M devido à sua posição no sítio: G e H devido à proximidade aos motivos GYP e HD respectivamente e M por estar no meio. As moléculas cujos átomos de carbono estão coloridos em amarelo são oriundas do tampão de cristalização e as esferas em vermelho representam moléculas do solvente. Átomos de carbono coloridos em cinza pertencem à cadeia lateral dos resíduos envolvidos nas interações. A imagem (c), que representa a ligação do c-di-GMP ao sítio ativo, foi gerada a partir de uma superposição de estruturas. Nela é possível ver a conformação do c-di-GMP no sítio e a provável interação com um dos íons Fe. ...	30
Figura 6	- Representação simplificada de um sistema de dois componentes.	32
Figura 7	- Metabolismo e módulo de controle fenotípico básico da molécula sinalizadora c-di-GMP, cuja estrutura é mostrada no centro da imagem. As formas em verde representam os sinais extracelulares percebidos pelas enzimas que sintetizam e degradam o sinalizador.	33
Figura 8	- Superposição de conformações adotadas pela molécula de c-di-GMP em complexo com diferentes proteínas, mostrada de ângulos distintos. Os códigos 3PJT, 3GFX, 3HV8, 4F3H, 3QYY, 2RDE e 4F5D correspondem às estruturas depositadas no PDB (<i>Protein Data Bank</i>) das quais foram resgatadas as imagens..	34

- Figura 9 - Gráficos de correlação entre formação de biofilme e concentração de c-di-GMP produzido por sete DGCs de *V. cholerae*. Em (a) é mostrada a reta obtida para a correlação global e, em (b), as retas obtidas para as correlações específicas para cada DGC. Em (a), o coeficiente de correlação r^2 está indicado no gráfico; em (b), os índices são mostrados à direita do gráfico. 42
- Figura 10 - Modelo de regulação da expressão de CsgD, mediada por dois módulos, um de ação global e outro de ação local, vinculados por c-di-GMP e pela enzima gatilho YciR. 46
- Figura 11 - Mapas dos vetores do *kit BacterioMatch II Two-Hybrid System*: (a) pBT e (b) pTRG..... 52
- Figura 12 - Esquematização da sequência de eventos de um ensaio de duplo-híbrido em nível molecular. Após a clonagem dos possíveis parceiros de interação nos vetores “isca” e “presa”, a cotransformação na linhagem repórter e a coexpressão das proteínas fusionadas, ocorre a interação entre as proteínas “isca” e “presa” e a consequente ativação da transcrição dos genes repórteres, resultando na seleção positiva. O vetor “isca” carrega o gene de resistência ao cloranfenicol (indicado em azul); o vetor “presa” carrega o gene de resistência à tetraciclina (indicado em laranja). 53
- Figura 13 - Esquema ilustrativo da técnica de LIC (*Ligase Independent Cloning*). Em verde, são mostradas as sequências correspondentes aos *primers* utilizados para a modificação dos vetores e para a amplificação dos insertos; em vermelho, as extensões que serão removidas pela T4 DNA polimerase para a criação de *overhangs* e, em negrito, o ponto em que ocorre o equilíbrio entre as atividades exonucleásica e polimerásica da T4 DNA polimerase. As extensões dos vetores são compostas somente por adeninas, citosinas e guaninas e são limitadas por uma timina (em negrito) na extremidade 5'; já as dos insertos, compostas por citosinas, guaninas e timinas e limitadas por uma adenina (em negrito) na extremidade 5'. Estão representados no esquema: (a) a modificação dos vetores pBT e pTRG; (b) o resultado da amplificação do vetor pBT e de um inserto a partir dos *primers* desenhados e o sentido da atividade exonucleásica da T4 DNA polimerase; (c) vetor e inserto após tratamento com a T4 DNA polimerase, na presença de dTTP e de dATP, respectivamente, evidenciando os *overhangs* gerados; (d) anelamento entre o vetor e o inserto tratados. 55
- Figura 14 - Sítios de clonagem de pBT e pTRG modificados para a técnica de LIC (*Ligase Independent Cloning*). Em verde, são mostradas os *primers* para a modificação dos vetores; em vermelho, as extensões que serão removidas pela T4 DNA polimerase para a criação de *overhangs* e, em negrito, o ponto em que ocorre o equilíbrio entre as atividades exonucleásica e polimerásica da T4 DNA polimerase. 60
- Figura 15 - Resultados da predição de hélices transmembranares obtidos pela ferramenta *online* TMHMM – em (a) e (c) – e da predição de domínios obtidos pela ferramenta do banco de dados Pfam – em (b) e (d). Em (a) e

	(b) são apresentados os resultados para a proteína PA3825 e, em (c) e (d), para a proteína PA1727.	71
Figura 16 -	Arquitetura de domínios das construções obtidas após o processamento das porções transmembranares. À esquerda é indicada a proteína de origem das construções, as quais foram identificadas especificamente por letras (de A a H) e números (de 1 a 9).	73
Figura 17 -	Imagens resultantes de eletroforese em gel de agarose para a análise de restrição de alguns dos clones de (a) “iscas” e (b) “presas” obtidos pela técnica de dupla-restrição enzimática. As bandas mais altas correspondem ao vetor (pBT ou pTRG) e as mais baixas, ao inserto clivado. Como sempre foram testados vetores extraídos de, ao menos, duas colônias por tentativa de clonagem, os índices (1) e (2) identificam a colônia de origem. Os tamanhos esperados para os insertos eram: 780 pb (PA0707_A1), 990 pb ((tnpM)_F1), 777 pb (PA1120_D5), 1419 pb (PA3311_D2), 2100 pb (PA0285_B4), 606 pb (PA1107_D7), 240 pb (PA3311_C2).	74
Figura 18 -	Imagens digitalizadas de placas mostrando o resultado final da etapa de padronização dos ensaios de duplo-híbrido do tipo um a um. Em (a) é mostrado o controle negativo; em (b) e em (c), os controles de eficiência de cotransformação dos controles negativo e positivo, respectivamente, e, em (d), o controle positivo.	77
Figura 19 -	Imagens digitalizadas de placas mostrando o resultado final da etapa de padronização dos ensaios para a estratégia de triagem por grupos. Em (a) é mostrado o controle negativo; em (b) e em (c), os controles de eficiência de cotransformação do controle negativo e da simulação de interação positiva, respectivamente, e, em (d), a simulação de interação positiva em um ensaio entre grupos.	77
Figura 20 -	Matriz ilustrativa dos resultados da estratégia de triagem por grupos. À direita, está indicada a legenda de cores. Os padrões de interação forte e fraca estão representados pelas imagens de digitalizadas de placas exibidas abaixo da matriz.	80
Figura 21 -	Gráfico apontando as parcelas de ensaios de interação da estratégia de triagem por grupos que resultaram em padrões de interação forte, padrão de interação fraca e em ausência de interação.	80
Figura 22 -	Primeira rede de interações resultante da etapa de triagem por pares. Optou-se pela utilização dos nomes das proteínas (quando já existente). Linhas tracejadas indicam interações fracas, linhas simples indicam interações de força média e linhas de maior espessura indicam interações fortes. Módulos de cor vermelha indicam proteínas que contêm domínio GGDEF; de cor vermelha e amarela, ambos os domínios GGDEF e EAL; de cor amarela, domínio EAL e de cor azul, domínio HD-GYP.	81
Figura 23 -	Imagens digitalizadas de placas mostrando o resultado final dos ensaios de interação complementares com a proteína FimX. Em (a), é mostrado o teste	

feito com a porção N-terminal truncada e, em (b), com o domínio EAL truncado. Acima das imagens, são apresentadas esquematizações das construções correspondentes. 83

Figura 24 - Imagens digitalizadas de placas mostrando o resultado final dos ensaios de interação complementares com PA0847_G5 e PA3343_F7. Em (a), é mostrado o teste feito com PA0847_HP; em (b), com PA0847_P; em (c), com PA0847_PG e, em (d), com PA0847_G. Acima das imagens, são apresentadas esquematizações das construções correspondentes. 84

Figura 25 - Ensaios de força de interação e testes de autoativação, feitos em meio duplamente seletivo. À esquerda de cada resultado, estão indicados os parceiros de interação testados. Abaixo das imagens dos testes de força estão indicadas as diluições correspondentes a cada coluna. À direita encontram-se os testes de autoativação de “iscas” e “presas” envolvidas em cada teste. Espaços em branco indicam testes que foram descartados devido à ausência de crescimento no controle de cotransformação correspondente (não mostrado). Os resultados estão agrupados de acordo com a data em que foram feitos – os controles correspondentes a cada uma das rodadas de testes são mostrados nas duas primeiras linhas. 85

Figura 26 - Rede de interações final encontrada. São apresentadas as novas interações evidenciadas durante a finalização da averiguação das interações positivas da etapa de triagem por grupos, conjuntamente com as interações já anteriormente encontradas. Linhas tracejadas indicam interações fracas, linhas simples indicam interações de força média e linhas de maior espessura indicam interações fortes. Módulos de cor vermelha indicam proteínas que contêm domínio GGDEF; de cor vermelha e amarela, ambos os domínios GGDEF e EAL; de cor amarela, domínio EAL; de cor azul, domínio HD-GYP e de cor roxa, domínio PilZ. 87

LISTA DE TABELAS

Tabela 1 -	Distribuição dos domínios GGDEF, EAL e HD-GYP em oito genomas bacterianos. A terceira coluna refere-se a proteínas que contém ambos os domínios GGDEF e EAL; a segunda e a terceira, somente domínio GGDEF ou EAL, respectivamente. Somente <i>C. crescentus</i> (Alphaproteobacteria), <i>B. cenocepacia</i> (Betaproteobacteria) e <i>C. difficile</i> (Clostridia; única Gram-positiva listada como exemplo) não pertencem à classe Gammaproteobacteria. Os organismos estão apresentados na ordem decrescente dos valores da última coluna, que apresenta a soma dos valores de todas as outras colunas para aquele organismo.	39
Tabela 2 -	Compilação de proteínas alvo. Algumas informações não foram encontradas para todas as proteínas, por essa razão alguns dos espaços não foram preenchidos na terceira, na sexta e sétima colunas.	69
Tabela 3 -	Lista de clones (“iscas” e “presas”) não obtidos. Cada “isca” ou “presa” foi designada pela vetor em que seria clonada, pela proteína de origem da construção e pelo código que designa a construção.	76
Tabela 4 -	Composição dos oito grupos de “iscas” formados. A denominação de cada componente identifica qual o vetor em que a construção foi clonada (no caso, pBT), a proteína de origem e o código da construção.	79
Tabela 5 -	Composição dos oito grupos de “presas” formados. A denominação de cada componente identifica qual o vetor em que a construção foi clonada (no caso, pTRG), a proteína de origem e o código da construção. ...	79
Tabela A1 -	Lista de primers para amplificação das sequências alvo para clonagem pelo método de dupla-restrição enzimática.	101
Tabela A2 -	Lista de primers para amplificação das sequências alvo para clonagem pela técnica LIC.	104
Tabela A3 -	Composição dos oito grupos de “iscas” formados. Na coluna “Descrição”, são dados os critérios de formação de cada grupo – as proteínas a que se faz referência nessa coluna são as proteínas que originaram as construções que compõe o grupo em questão.	106
Tabela A4 -	Composição dos oito grupos de “presas” formados. Na coluna “Descrição”, são dados os critérios de formação de cada grupo – as proteínas a que se faz referência nessa coluna são as proteínas que originaram as construções que compõe o grupo em questão.	107

LISTA DE ABREVIATURAS E SIGLAS

3-AT	3-amino-1,2,4-triazol
5'-pGpG	5'-fosfoguanilil-(3'-5')-guanosina
Ala	Alanina
Asp	Ácido aspártico
BSA	Albumina do soro bovino
c-di-GMP	Guanosina monofosfato (3'-5')-cíclica dimérica
cDNA	DNA complementar
CAP	Proteína ativadora do catabolismo
CRP	Proteína receptora de AMPc
DGC	Diguanilato Ciclase
DMSO	Dimetilsulfóxido
DNA	Ácido desoxirribonucleico
dATP	Desoxiadenosina trifosfato
dNTP	Desoxirribonucleotídeo trifosfato
DSF	<i>Diffusible Signal Factor</i>
DTT	DL-Ditiotreitol
dTTP	Desoxitimidina trifosfato
EPS	Substâncias poliméricas extracelulares
FRET	<i>Förster / Fluorescence Resonance Energy Transfer</i>
GFP	Proteína verde fluorescente
Gly	Glicina
Glu	Ácido glutâmico
GMP	Guanosina monofosfato
GTP	Guanosina trifosfato
HDO	<i>His-dropout</i>
His	Histidina
HK	Quinase de histidina
HTP	<i>High throughput</i>
IPTG	Isopropil- β -D-tiogalactopiranosídeo
KEGG	<i>Kyoto Encyclopedia of Genes and Genomes</i>
LB	Luria-Bertani

Leu	Leucina
LIC	<i>Ligase Independent Cloning</i>
MBEC	<i>Minimum biofilm-eliminating concentration</i>
MCS	Sítio múltiplo de clonagem
mRNA	RNA mensageiro
NEB	New England Biolabs
ORF	Quadro aberto de leitura
pb	Pares de bases
PCR	Reação em cadeia da polimerase
PDE	Fosfodiesterase
PEG	Polietilenoglicol
Phe	Fenilalanina
P _i	Fosfato inorgânico
PP _i	Pirofosfato
(p)ppGpp	Guanosina pentafofato ou guanosina tetrafofato
Pro	Prolina
RNA	Ácido ribonucleico
RNAP α	Subunidade α da RNA polimerase
SOC	<i>Super Optimal broth with Catabolite repression</i>
SR	Resposta estrigente
TErRA	Temático em Energias Renováveis e Meio Ambiente
T _m	Temperatura de fusão / desnaturação
TMHMM	<i>TransMembrane prediction using Hidden Markov Models</i>
Tyr	Tirosina
λ cI	Repressor cI do bacteriófago λ

SUMÁRIO

1	INTRODUÇÃO	23
1.1	Biofilmes bacterianos	23
1.2	Vias de sinalização mediadas por c-di-GMP	27
1.2.1	Síntese e degradação da molécula de c-di-GMP	28
1.2.1.1	Domínios catalíticos GGDEF, EAL e HD-GYP	28
1.2.1.2	Domínios sensores	31
1.2.2	Efetores, alvos e respostas fenotípicas	32
1.3	Multiplicidade das proteínas envolvidas no metabolismo de c-di-GMP	38
1.4	Sinalização global e sinalização específica	41
1.5	Interações proteicas nas vias de sinalização mediadas por c-di-GMP	43
2	OBJETIVOS	49
3	MATERIAIS E MÉTODOS	51
3.1	Principais técnicas utilizadas	53
3.1.1	Método de duplo-híbrido bacteriano	51
3.1.2	<i>Ligase Independent Cloning</i> (LIC)	54
3.2	Materiais e métodos	56
3.2.1	Compilação das sequências alvo e remoção de porções transmembranares	56
3.2.2	Construção das bibliotecas de DNA de “iscas” e “presas” por meio da técnica de dupla-restrição enzimática	57
3.2.2.1	Desenho de primers para clonagem por dupla-restrição enzimática	57
3.2.2.2	Clonagem das construções selecionadas nos vetores pBT e pTRG por meio de dupla-restrição enzimática	58
3.2.3	Construção das bibliotecas de DNA de “iscas” e “presas” por meio da técnica de <i>Ligase Independent Cloning</i> (LIC)	59
3.2.3.1	Modificação dos vetores pBT e pTRG e desenho de primers para LIC	59

3.2.3.2 Clonagem das construções selecionadas nos vetores pBT e pTRG por meio da técnica de LIC	61
3.2.4 Ensaios de duplo-híbrido bacteriano	63
3.2.4.1 Células competentes.....	63
3.2.4.2 Meios não seletivo, seletivo e duplamente seletivo e meio líquido HDO	63
3.2.4.3 Controles	64
3.2.4.4 Padronização e otimização do protocolo dos ensaios de interação.....	65
3.2.4.5 Estratégia de triagem por grupos, seguida de triagem por pares	66
3.2.4.6 Ensaios de força de interação.....	67
4 RESULTADOS E DISCUSSÃO.....	69
4.1 Compilação das sequências alvo e remoção de porções transmembranares.....	69
4.2 Construção das bibliotecas de DNA de “iscas” e “presas” por meio da técnica de dupla-restrição enzimática	74
4.3 Construção das bibliotecas de DNA de “iscas” e “presas” por meio da técnica de <i>Ligase Independent Cloning</i> (LIC)	75
4.4 Ensaios de duplo-híbrido bacteriano.....	76
4.4.1 Padronização dos ensaios de interação	76
4.4.2 Ensaios de interação: triagem por grupos	78
4.4.3 Ensaios de interação: triagem por pares.....	80
4.4.4 Ensaios de interação complementares.....	82
4.4.5 Ensaios de força de interação e testes de autoativação	84
4.4.6 Ensaios finais e mapa de interações proteicas	86
5 CONCLUSÕES.....	89
6 PERSPECTIVAS.....	91
REFERÊNCIAS	93
APÊNDICE A – Tabelas adicionais	101

1 INTRODUÇÃO

1.1 Biofilmes bacterianos

Um biofilme é definido, de forma ampla, como uma comunidade microbiana aderida a uma superfície – em geral, em uma interface sólido-líquido – e envolta por uma matriz altamente hidratada, composta, essencialmente, por uma variedade de polissacarídeos, proteínas, ácidos nucleicos e lipídios produzidos pelos próprios microrganismos que o constituem (Figura 1).¹⁻² O conjunto de biopolímeros da matriz, denominado EPS (do inglês, *extracellular polymeric substances*), dentre diversas outras funções, dá suporte à arquitetura tridimensional do biofilme, contribuindo para a adesão celular à superfície colonizada e para a coesão intercelular do agregado.² Em geral, menos de 10% da massa seca total de um biofilme corresponde ao agregado celular, sendo, portanto, os mais de 90% restantes correspondentes ao material extracelular que compõe a matriz.²

Figura 1 - Modelos da matriz extracelular em diferentes escalas. São ilustrados: a) um biofilme aderido à superfície; b) os principais componentes do EPS; c) interações entre polissacarídeos, que contribuem para a estabilidade da estrutura e d) a modelagem molecular da interação entre um polissacarídeo (alginato) e uma lipase extracelular na matriz.

Fonte: Adaptada de FLEMMING.²

O estado de agregação celular em biofilmes é também dito estado sésil (ou sedentário), em oposição ao estado móvel (ou planctônico), em que as células têm um estilo de vida livre.¹ A transição entre esses estados ocorre em estágios, os quais estão ilustrados, de forma simplificada, na Figura 2. A adesão inicial de células a uma superfície é seguida da produção do EPS que as permeia, formando uma estrutura primária, a qual se desenvolve até a formação do biofilme maduro. O estágio que completa o ciclo de transição é a dispersão de células do biofilme. As células planctônicas liberadas poderão então colonizar outras superfícies e formar novos biofilmes. Todo o processo é acompanhado pela expressão de genes distintos, regulada em resposta a alterações nas condições a que estão submetidas as células.²⁻⁵ Alguns exemplos de fenótipos tipicamente expressos quando a transição ocorre no sentido de formação do biofilme são a produção de estruturas de adesão, como *fimbriae*, e a síntese de polissacarídeos da matriz; já no sentido oposto, de dispersão do biofilme, são expressos genes associados à motilidade, como, por exemplo, aqueles envolvidos na síntese de flagelo, ou ainda enzimas que catalisam a degradação de exopolissacarídeos, facilitando a liberação das células.¹⁻²

Figura 2 - Etapas de formação de um biofilme e de sua dispersão em células planctônicas.

Fonte: Adaptada de KRASTEVA.⁶

A complexidade, heterogeneidade e flexibilidade de composição dos biofilmes e das funções de seus componentes são características que contribuem enormemente para o sucesso dessas estruturas nos ambientes em que se estabelecem e, na mesma proporção, para a dificuldade em estudá-los.^{2,5,7} Biofilmes podem ser compostos de uma ou mais espécies de microrganismos – neste caso, os consórcios microbianos formados podem render-lhes os mais diversos tipos de vantagens, como o compartilhamento de substâncias que por si só não produzem ou ainda a transferência horizontal de genes decorrente da presença de material

genético na matriz extracelular. Podem conter também vários tipos de biopolímeros, tanto secretados pelas células quanto resultantes da “reciclagem” dos produtos originados de lise celular.² A diversidade de composição não se restringe apenas à comparação entre biofilmes – ocorre também entre regiões de um mesmo biofilme, indicando a existência de microdomínios dentro da estrutura. Além disso, as características do meio em que se encontra um biofilme influenciam fortemente a sua composição. Alterações na composição de um biofilme surgem constantemente como respostas adaptativas a alterações nas condições do meio em que está inserido.⁵

A perda de informações relativas à heterogeneidade e à flexibilidade de composição de biofilmes é quase inevitável durante a execução de estudos, já que, em geral, muitos métodos de análise fornecem informações muito generalizadas, que ignoram variações espaciais dentro da estrutura e diferenças nas condições às quais as células estão submetidas em ambientes distintos.² Felizmente, nos últimos anos, avanços tecnológicos nas ferramentas de análise, especialmente nas áreas de microscopia e espectroscopia, têm permitido a elaboração de estudos mais abrangentes sobre a matriz extracelular de organismos já vastamente estudados.⁷

São diversos os benefícios atrelados à organização de microrganismos em biofilmes, dentre eles: a proximidade entre células, que favorece interações cooperativas e comunicação intercelular; a retenção de enzimas extracelulares na matriz, que permite a quebra de macromoléculas (inclusive do próprio EPS) e, conseqüentemente, a disponibilização de nutrientes para as células que compõem o biofilme; a retenção de água, que aumenta a tolerância do agregado celular ao ressecamento; a tolerância elevada a condições externas desfavoráveis, como cisalhamento e alterações no pH do meio e a resistência a diferentes tipos de agentes antimicrobianos, tais como os agentes de defesa de organismos hospedeiros e antibióticos usados para o tratamento de infecções.²⁻³

Biofilmes bacterianos estão amplamente associados a infecções crônicas – estima-se que até 80% das infecções estejam relacionadas à formação de biofilmes – e a infecções causadas por instrumentos e implantes médicos.⁸ Biofilmes instalam-se, por exemplo, em válvulas cardíacas, causando endocardite; nos pulmões de pacientes com fibrose cística, causando pneumonia crônica; no ouvido médio de pacientes com otite persistente e, muito frequentemente, em próteses, cateteres, *stents* e implantes médicos utilizados para os mais diversos fins. Estima-se que 50% das infecções nosocomiais estejam associadas à utilização desses implantes e, portanto, à presença de biofilmes.^{4,9}

As infecções caracterizadas pela presença de biofilme evidenciam-se, clinicamente, pela elevada resistência a tratamentos com antibióticos, os quais, embora auxiliem no controle dos

episódios intermitentes de manifestação aguda dos sintomas da infecção, mostram-se inefetivos para, de fato, erradicá-la. Tais episódios estão relacionados à dispersão eventual de células planctônicas do biofilme, evento que, de forma geral, está relacionado à maior produção de fatores de virulência e, por essa razão, é caracterizado pelo agravamento dos sintomas da doença.⁴ A eficácia da administração de antibióticos no controle desses episódios, mas não na eliminação definitiva da infecção, deve-se, em parte, ao fato de o desenho desses fármacos ter sido voltado à eliminação de células em seu estado planctônico.¹⁰ Experimentos realizados tanto *in vitro* quanto *in vivo* demonstraram que, em biofilmes, as bactérias são de dez a mil vezes mais resistentes à ação de antibióticos do que quando se encontram no estado planctônico. Logo, devido à toxicidade e a efeitos colaterais, é praticamente inviável que sejam alcançadas as concentrações de antibiótico necessárias para a erradicação do patógeno do organismo infectado.⁴

Há urgência em encontrar soluções eficientes para sanar o grave problema das infecções crônicas. Por essa razão, há crescente interesse em desvendar os fatores ou mecanismos relacionados a essa elevada resistência dos biofilmes. Dentre as causas já propostas, as mais relevantes atualmente indicam que a tolerância bacteriana é mediada por respostas de controle ao estresse oxidativo causado pelos antibióticos. Estudos identificaram a produção de radicais hidroxila (OH•) como um mecanismo comum de ação de antibióticos, que independe de seus alvos primários e que provoca a morte celular por danos oxidativos.¹¹ Segundo Nguyen e colaboradores, o controle do estresse oxidativo é mediado pela resposta estrigente (SR, do inglês, *stringent response*), um mecanismo regulatório que é acionado quando há limitação de nutrientes e que envolve a produção da molécula sinalizadora (p)ppGpp. Apesar da existência de estratégias de aumento da circulação de nutrientes pelo biofilme, a difusão desses nutrientes para as células menos próximas à superfície é reduzida, resultando no acionamento da SR e, conseqüentemente, no aumento da tolerância à ação de antibióticos. As células bacterianas que entram nesse estado fenotípico são ditas persistentes. Os experimentos feitos pelos autores indicaram ainda que, em oposição ao que se pensava anteriormente, o estado de baixo ou nulo crescimento de grande parte das células que compõem o biofilme não é um fator determinante para a ineficácia dos antibióticos.¹⁰⁻¹¹ Shatalin e colaboradores, por sua vez, evidenciaram a proteção contra o estresse oxidativo conferida pela atuação sinérgica resultante da produção de óxido nítrico (NO) e de sulfeto de hidrogênio (H₂S), que também explicaria a multirresistência bacteriana a antibióticos.¹²

O crescente interesse em esclarecer os mecanismos relacionados à formação e à manutenção dos biofilmes deve-se não somente à necessidade de combatê-los, mas também às

possíveis aplicações benéficas dessas estruturas em diferentes cenários. Devido à capacidade de acúmulo, imobilização ou degradação de substâncias em sua matriz, à sua maior resistência ao estresse mecânico e à sua capacidade de adaptação e sobrevivência em condições adversas, biofilmes têm grande utilidade em processos de biorremediação.¹³ Já Absalon e colaboradores propuseram a utilização de biofilmes como plataformas para o direcionamento de enzimas a superfícies para fins terapêuticos. Baseando-se no fato de que a matriz extracelular é rica em proteínas de adesão, os autores demonstraram a viabilidade de retenção de enzimas ativas na matriz, por meio de uma fusão entre RbmA, uma proteína de adesão de *Vibrio cholerae*, e uma quitinase, também de *V. cholerae*. Estratégias similares poderiam ser empregadas em muitas outras aplicações, envolvendo outros microrganismos, em que haja interesse em orientar enzimas a superfícies específicas. Como exemplos, os autores citam a possibilidade de usar *Escherichia coli* ou probióticos para a entrega de lactase ou enzimas pancreáticas ao epitélio do intestino de organismos deficientes nessas enzimas ou de usar bactérias não patogênicas modificadas de forma a permitir o acúmulo de alginases na matriz para a colonização dos pulmões de pacientes com fibrose cística, colaborando para a eliminação de infecções por *Pseudomonas aeruginosa*, normalmente associadas a essa condição.¹⁴

Uma importante limitação para o desenvolvimento de soluções e aplicações cada vez mais abrangentes e seguras para os biofilmes bacterianos é a compreensão ainda restrita sobre aspectos fundamentais de seu funcionamento. Embora muito já se tenha avançado nessa direção, ainda restam inúmeras questões em aberto, relacionadas, especialmente, aos mecanismos regulatórios que coordenam esses sistemas. Compreender esses mecanismos é, sem dúvida, central para que se estabeleçam as correlações entre processos, componentes, sinais, respostas, causas e consequências relacionadas à formação e à dispersão de biofilmes.

1.2 Vias de sinalização mediadas por c-di-GMP

Ao longo das quase três décadas transcorridas desde que foi descoberta, uma pequena molécula mensageira consolidou-se como um elemento central na regulação de processos celulares associados à formação e à dispersão de biofilmes bacterianos e à adaptação a alterações ambientais.^{8,15} Primeiramente descrita por Benziman e colaboradores como um ativador alostérico de uma sintase bacteriana de celulose, sabe-se hoje que a molécula guanosina monofosfato (3'-5')-cíclica dimérica (c-di-GMP) (Figura 3) é um sinalizador intracelular quase universal entre bactérias.^{1,16} Apesar de ter sido descrita também em bactérias Gram-positivas, a sinalização mediada por c-di-GMP foi muito mais amplamente

identificada entre espécies Gram-negativas, tendo sido mais extensivamente estudados exemplos pertencentes à classe Gammaproteobacteria, como por exemplo, os sistemas encontrados em representantes dos gêneros *Pseudomonas*, *Xanthomonas*, *Vibrio*, *Escherichia* e *Salmonella*.^{8,16}

Figura 3 - A molécula de c-di-GMP.

Fonte: Adaptada de HENGGE.¹

1.2.1 Síntese e degradação da molécula de c-di-GMP

1.2.1.1 Domínios catalíticos GGDEF, EAL e HD-GYP

A síntese da molécula de c-di-GMP é catalisada por enzimas diguanilato ciclases (DGCs – do inglês, *diguanylate cyclases*). As DGCs caracterizam-se pela presença do domínio GGDEF, o qual foi assim designado por conter, em sua sequência, o motivo conservado Gly-Gly-Asp-Glu-Phe. Este motivo corresponde ao sítio ativo, sendo que somente a mutação de Asp (ácido aspártico) para Glu (ácido glutâmico) não elimina a atividade enzimática. A síntese ocorre a partir de duas moléculas de guanosina trifosfato (GTP), em uma reação de duas etapas, com a liberação de uma molécula de pirofosfato (PP_i) em cada uma delas e tendo como intermediário o dinucleotídeo linear 5'-pppGpG (ou, de outra forma, GTP-GMP).^{1,8,17} A DGC ativa é um homodímero – de fato, o sítio ativo forma-se na interface dimérica (Figura 4a), sendo que cada uma das subunidades contribui com uma das moléculas de GTP para a reação, que ainda requer coordenação por dois íons Mg^{2+} ou Mn^{2+} . Há ainda um sítio inibitório (RxxD, sendo “x” qualquer resíduo), presente em aproximadamente metade das DGCs. Especula-se que aquelas que não contêm sítio inibitório possam ser inibidas competitivamente pela ligação de c-di-GMP aos seus sítios ativos.⁸

Já a degradação do c-di-GMP é catalisada por enzimas fosfodiesterases (PDEs – do inglês, *phosphodiesterases*). Há dois tipos de fosfodiesterases envolvidas na hidrólise desse sinal: aquelas que contêm domínio EAL, assim denominado devido à existência do motivo conservado Glu-Ala-Leu, e aquelas que contêm domínio HD-GYP (trata-se de um

subconjunto da família de domínios hidrolíticos HD), assim designado também em virtude de um motivo conservado – no caso, His-Asp-(...)-Gly-Tyr-Pro.⁸

A catálise pelo domínio EAL ocorre por meio da linearização do c-di-GMP a 5'-pGpG (ou, de outra forma, di-GMP). Embora o domínio EAL exiba alguma atividade catalítica para a quebra deste produto, ela não é consistente com a taxa de hidrólise de c-di-GMP observada *in vivo*, o que indica que a quebra do 5'-pGpG a duas moléculas de GMP é feita por enzimas alternativas.⁸ Recentemente, a proteína oligoribonuclease (Orn), uma exonuclease 3'→5' que hidrolisa nanoRNAs (oligômeros compostos por dois a cinco nucleotídeos), foi identificada como uma enzima fundamental para a degradação de 5'-pGpG em *Pseudomonas aeruginosa*.¹⁸⁻¹⁹

Figura 4 - Estrutura dos sítios ativos de domínios envolvidos no metabolismo da molécula c-di-GMP em complexo com seus respectivos substratos. Em (a), é mostrado um análogo não hidrolisável de GTP (GTPαS) no sítio ativo do domínio GGDEF; em (b), a molécula de c-di-GMP no sítio ativo do domínio EAL. Os átomos de carbono dos substratos estão coloridos em cinza, assim como as glicinas do motivo GGDEF. Átomos catalíticos importantes estão coloridos em amarelo e indicados por setas. As esferas cor-de-rosa representam os íons metálicos Mg²⁺.

Fonte: Adaptada de RÖMLING.⁸

Assim como ocorre com o domínio GGDEF, as evidências indicam que um dímero é a forma ativa da enzima, muito embora, neste caso, o monômero do domínio EAL também retenha alguma atividade catalítica.⁸ A reação ocorre por meio da coordenação de duas moléculas de água por dois íons metálicos – uma das moléculas de água promove um ataque hidrolítico em uma ligação fosfoéster do c-di-GMP. O resíduo Glu do motivo EAL está diretamente envolvido na coordenação de um desses íons e, por essa razão, é imprescindível para a atividade catalítica do domínio (Figura 4b). Por fim, é interessante apontar que, embora a reação ocorra também por meio da coordenação por íons Mg²⁺, a coordenação por Mn²⁺

resulta em distâncias ótimas de interação entre esses íons e as moléculas de água, resultando em uma configuração mais favorável para a catálise. Já íons Ca^{2+} inibem a atividade fosfodiesterase por provocarem uma grande distorção nessas distâncias de interação.⁸

Devido à dificuldade de cristalização de domínios HD-GYP, apenas recentemente obteve-se acesso à estrutura desse domínio ativo.^{8,20} O êxito de Bellini e colaboradores na obtenção da estrutura cristalina da PDE ativa *PmGH* de *Persephonella marina*, tanto isolada quanto em complexo com íons metálicos, com c-di-GMP e com GTP, forneceu novos indícios a respeito do funcionamento desse domínio. Um dos aspectos mais interessantes identificados pelos autores foi a presença de um centro catalítico formado por três íons Fe arranjados em uma geometria triangular, suportados por interações com oito resíduos da estrutura proteica. A substituição de qualquer um desses resíduos por alanina resultou na inibição ou redução drástica da atividade catalítica do domínio, indicando a relevância desses três íons para a catálise ou, ao menos, para a coordenação do sítio catalítico. A estrutura complexada com c-di-GMP revelou que esta molécula adota uma conformação em “U” e que fica próxima ao centro de íon Fe (Figura 5). Com base em outras análises, os autores também identificaram a possibilidade de que estes íons formem o sítio de ligação ao c-di-GMP.^{20,21}

Figura 5 - Configuração do centro catalítico formado por três íon Fe e modelagem da molécula de c-di-GMP no sítio ativo do domínio HD-GYP da PDE *PmGH*. Em (a), é dado enfoque à coordenação dos íons Fe G e M e, em (b), à coordenação dos íons Fe M e H. Os íons Fe estão representados como esferas cor de laranja e foram rotulados G, H e M devido à sua posição no sítio: G e H devido à proximidade aos motivos GYP e HD respectivamente e M por estar no meio. As moléculas cujos átomos de carbono estão coloridos em amarelo são oriundas do tampão de cristalização e as esferas em vermelho representam moléculas do solvente. Átomos de carbono coloridos em cinza pertencem à cadeia lateral dos resíduos envolvidos nas interações. A imagem (c), que representa a ligação do c-di-GMP ao sítio ativo, foi gerada a partir de uma superposição de estruturas. Nela é possível ver a conformação do c-di-GMP no sítio e a provável interação com um dos íons Fe.

Fonte: Adaptada de BELLINI.²⁰

Apesar das novas evidências, estudos complementares serão necessários para que o mecanismo pelo qual os domínios HD-GYP funcionam possa ser devidamente descrito. Até o momento, é bem estabelecido que a hidrólise por PDEs que contêm esse domínio tem, como produto final, duas moléculas de GMP, diferentemente do que ocorre na catálise feita pelo

domínio EAL, em que o produto final é o dinucleotídeo linear 5'-pGpG. Segundo análises bioquímicas feitas com proteínas de *P. aeruginosa*, a hidrólise pelo domínio HD-GYP acontece em duas etapas, tendo 5'-pGpG como produto intermediário.²²

1.2.1.2 Domínios sensores

Uma grande parte das enzimas diguanilato ciclases e fosfodiesterases envolvidas no metabolismo do c-di-GMP são modulares. Além dos típicos domínios catalíticos, GGDEF, EAL e HD-GYP, geralmente localizados em proximidade à extremidade C-terminal da proteína, muitas dessas enzimas contêm também um ou mais domínios sensores localizados na porção N-terminal da cadeia proteica. A atividade dos domínios enzimáticos é controlada por esses domínios sensores em resposta ao recebimento de diferentes sinais de entrada, estímulos ambientais ou sinais celulares, resultando no ajuste dos níveis do segundo mensageiro c-di-GMP e de respostas fenotípicas por ele reguladas. É interessante observar que, como parte dessas enzimas contém mais de um domínio sensor, é possível que o controle de sua atividade catalítica dê-se por meio da integração entre diferentes sinais recebidos.^{1,15,23}

Dentre os domínios sensores mais comumente encontrados associados aos domínios de síntese e de degradação de c-di-GMP, estão, por exemplo, os domínios CHASE, GAF, HAMP, MASE, PAS/PAC e REC.^{8,24} Muitas vezes, esses domínios ficam localizados no periplasma ou no espaço extracelular, conectados aos domínios catalíticos (citoplasmáticos) por meio de hélices transmembranares. Além dos domínios sensores, algumas dessas proteínas transmembranares contêm grandes loops periplasmáticos que, acredita-se, também podem estar envolvidos na percepção e transmissão de estímulos externos.²⁴ Apesar de um número reduzido dos ligantes específicos desses domínios sensores ter sido identificado, sabe-se que alguns dos sinais que podem ser percebidos e convertidos em respostas catalíticas incluem, por exemplo, antibióticos, fosforilação, íons, luz, moléculas gasosas (CO, NO, O₂), nucleotídeos, peptídeos e sinais de *quorum sensing* (autoindutores).^{1,8,23}

Um caso particularmente interessante é o da associação dos domínios catalíticos de c-di-GMP com domínios REC. As enzimas formadas por essa combinação são conhecidas como reguladores de resposta e participam de sistemas de dois-componentes com sensores do tipo quinase de histidina (HK, do inglês, *histidine kinase*) para a transmissão de sinais. O sinal recebido pelo sensor HK, o qual está frequentemente associado à membrana (e tipicamente contém domínio sensor periplasmático ou extracelular) é transmitido por meio de etapas sequenciais que envolvem a sua autofosforilação, seguida da fosforilação do domínio REC do

regulador de resposta cognato, desencadeando, por sua vez, a resposta do domínio catalítico (Figura 6).^{8,25} Até o início do ano de 2011, dentre o total de reguladores de resposta bacterianos descritos, contabilizou-se que aqueles que continham domínios do metabolismo de c-di-GMP representavam a expressiva parcela de 5,4%.²⁶

Figura 6 - Representação simplificada de um sistema de dois componentes.

Fonte: Adaptada de BRETL.²⁷

1.2.2 Efetores, alvos e respostas fenotípicas

Sequencialmente à síntese de moléculas de c-di-GMP, a cadeia de sinalização avança com a ligação deste segundo mensageiro a receptores, causando-lhes alterações conformacionais. Estes receptores ou efetores, sob o controle alostérico do c-di-GMP, passam a interagir de forma diferente com outras moléculas (ou, possivelmente, no caso de proteínas, com outro domínio dentro de sua própria estrutura), os alvos da sinalização, os quais estão diretamente associados à resposta fenotípica celular.^{1,28-29} As moléculas que atuam na síntese, degradação e reconhecimento desse sinal e as moléculas alvo, em conjunto, formam o módulo de sinalização básico desse segundo mensageiro (Figura 7).¹

Um aspecto notável da sinalização por c-di-GMP é a excepcional variedade de receptores que reconhecem essa molécula, o que revela não somente o quão amplo é o leque de possíveis respostas celulares que podem estar sob seu controle, mas também a sua versatilidade e flexibilidade de ligação.⁸ Embora as formas monomérica e dimérica tenham sido muito mais frequentemente encontradas complexadas com receptores, foi relatada, até o momento, ao menos uma ocorrência da ligação de um tetrâmero de c-di-GMP a um receptor – no caso, um fator de transcrição de *Streptomyces spp.*²⁹⁻³⁰ Na realidade, demonstrou-se que,

em condições adequadas, o c-di-GMP pode formar até mesmo octâmeros, ao passo que a sua forma linear, o 5'-pGpG, não forma nem mesmo tetrâmeros em concentrações similares.³¹⁻³² Essa capacidade formar oligômeros gera novas possibilidades de interação com outras moléculas. Além de seu potencial de oligomerização, as várias configurações adotadas pela molécula também contribuem para essa versatilidade (Figura 8).^{29,33} É, inclusive, interessante atentar para o fato de que o c-di-GMP liga-se diferentemente mesmo aos domínios que o degradam: uma conformação mais estendida foi verificada na ligação ao sítio catalítico do domínio EAL; já, ligado ao domínio HD-GYP, adota uma conformação em “U” (configuração *cis*).^{20,33-34} Adicionalmente, características inerentes à molécula, como, por exemplo, a superfície hidrofóbica provida pelas guaninas, aumentam ainda mais a diversidade de interações possíveis.³³

Figura 7 - Metabolismo e módulo de controle fenotípico básico da molécula sinalizadora c-di-GMP, cuja estrutura é mostrada no centro da imagem. As formas em verde representam os sinais extracelulares percebidos pelas enzimas que sintetizam e degradam o sinalizador.

Fonte: Adaptada de HENGGE¹; SONDERMANN.³⁵

Ao contrário do que acontece com os domínios que sintetizam e degradam o c-di-GMP, poucos são os efetores imediatamente identificados por sua sequência ou estrutura, o que decorre da grande variedade de moléculas que se ligam a esse mensageiro e da baixa

similaridade entre elas.⁸ De fato, somente algumas classes de receptores já foram identificadas e estudadas mais extensivamente – é o caso de proteínas que contêm domínio PilZ, as quais, ligadas ao mensageiro, regulam outros domínios ou proteínas por meio de interações proteicas; de proteínas com domínios GGDEF e EAL degenerados, os quais não têm função catalítica, mas ligam-se ao c-di-GMP e atuam como efetores nas vias de sinalização; de fatores de transcrição que se ligam ao sinal para regular a transcrição de genes e, por fim, de *riboswitches* que, da mesma forma, ligam-se ao c-di-GMP para regular as etapas de transcrição ou tradução (Figura 7).²⁹

Figura 8 - Superposição de conformações adotadas pela molécula de c-di-GMP em complexo com diferentes proteínas, mostrada de ângulos distintos. Os códigos 3PJT, 3GFX, 3HV8, 4F3H, 3QYY, 2RDE e 4F5D correspondem às estruturas depositadas no PDB (*Protein Data Bank*) das quais foram resgatadas as imagens.

Fonte: Adaptada de CHOU.³³

A primeira classe de receptores identificada com sucesso foi a das proteínas que contêm domínio PilZ. Um vínculo geral entre esse domínio e a molécula de c-di-GMP estabeleceu-se, inicialmente, pela observação de que havia ampla correlação entre a presença de domínios similares à proteína PilZ de *P. aeruginosa* – a qual está envolvida na regulação da formação de *pili* e da motilidade do tipo *twitching* e deu origem ao nome do domínio – e a presença de domínios envolvidos no metabolismo do c-di-GMP em genomas bacterianos.^{29,36} Estudos posteriores demonstraram, no entanto, a existência de três classes distintas de domínios PilZ, sendo que somente os domínios do tipo I podem ligar-se diretamente ao c-di-GMP. A ligação a esse sinal depende de resíduos contidos nos motivos RxxxR e (D/N)xSxxG desses domínios.³³

Em alguns casos, o domínio PilZ localiza-se junto ao terminal carboxila de domínios GGDEF, EAL e HD-GYP; é também frequente a associação do domínio PilZ com os domínios que influenciam diretamente a resposta celular. É o caso, por exemplo, da proteína Alg44 de *P. aeruginosa*, que participa da síntese do alginato, um dos polissacarídeos

encontrados na matriz extracelular de bactérias dessa espécie.^{1,37-38} É o caso também da proteína YcgR de *E. coli*, que, ligada ao c-di-GMP, interage com as proteínas FliG e FliM, componentes do rotor flagelar, reduzindo o torque gerado e alterando o padrão de rotação, o que, consequentemente, prejudica a motilidade celular.^{1,29,37,39} De forma geral, pode-se dizer que os domínios efetores PilZ, estruturalmente alterados pela ligação ao c-di-GMP, influenciam a resposta fenotípica via interações com outros domínios ou outras proteínas que estão, diretamente ou indiretamente, vinculados a essa resposta. Não obstante, os aspectos específicos dos eventos relacionados à participação dos efetores PilZ nas vias de sinalização variam entre si, podendo envolver, por exemplo, a formação ou dissociação de dímeros desses efetores.²⁹

Por algum tempo, imaginou-se que, analogamente a outros sistemas de sinalização celular, deveria haver um receptor universal único para a molécula de c-di-GMP. No entanto, as mesmas análises de bioinformática que indicaram o domínio PilZ como receptor apontaram também para a improbabilidade de que ele fosse universal, já que, dentre os genomas bacterianos examinados que continham domínios de metabolismo de c-di-GMP, nem todos continham também o recém-revelado receptor. Essa incoerência sugeria a existência de outros receptores atuantes nesse sistema de sinalização.^{33,36-37}

Dentre as outras classes de efetores encontradas, está a das proteínas com domínios degenerados GGDEF e EAL, os quais não desempenham função catalítica, mas, assim como os domínios PilZ, mediados pela ligação ao c-di-GMP, modulam a função ou a atividade enzimática de outros componentes via interações proteicas.^{1,29,33} No caso do domínio GGDEF, o controle alostérico exercido por meio da ligação de c-di-GMP ao seu sítio inibitório foi, desde o início, um indicativo de seu potencial como molécula efetora.³³ São exemplos desse tipo de receptor, as proteínas PelD e FimX de *P. aeruginosa* e a proteína LapD de *P. fluorescens*.^{29,37} PelD é codificada por um dos genes do operon *pel* e contém um domínio GGDEF degenerado. A ligação de c-di-GMP ao sítio inibitório desse domínio ocasiona o aumento da produção do exopolissacarídeo Pel e a formação de biofilme.⁴⁰ No caso de FimX, que contém os domínios GGDEF e EAL, ambos degenerados, a molécula de c-di-GMP liga-se com alta afinidade ao domínio EAL, provocando mudanças conformacionais na estrutura e, consequentemente, regulação da formação de *pili* do tipo IV e da motilidade do tipo *twitching*.^{29,34,41} Assim como FimX, a proteína transmembranar LapD também contém domínios GGDEF e EAL degenerados. A ligação de c-di-GMP ao domínio EAL citoplasmático da LapD provoca uma extensa mudança conformacional que se transfere por toda a estrutura proteica, alterando, por fim, a conformação do domínio periplasmático dessa

proteína e aumentando assim a sua afinidade pela protease LapG. Esta proteína, quando não interage com o domínio LapD, cliva a proteína LapA, envolvida na adesão a superfícies e, portanto, essencial para a formação e manutenção de biofilmes. Logo, a ligação de c-di-GMP ao efector LapD previne a dissociação do biofilme por possibilitar o sequestro da protease LapG, impedindo a clivagem da adesina LapA. É interessante ainda acrescentar que esse mecanismo é regulado de acordo com a disponibilidade de fosfato inorgânico P_i . Quando este se torna escasso, a expressão da PDE RapA e, conseqüentemente, a degradação do c-di-GMP são aumentadas, resultando na liberação da LapG e clivagem da LapA, o que, por sua vez, leva à dispersão do biofilme.⁴²

Fatores de transcrição compõem outro grupo de efetores da rede de sinalização por c-di-GMP. FleQ, o primeiro exemplo descrito dessa classe, é o regulador da transcrição dos genes do operon *pel*, envolvidos na síntese do polissacarídeo Pel (componente da matriz extracelular) e é também o regulador central de genes relacionados à motilidade flagelar em *P. aeruginosa*. Livre de c-di-GMP, FleQ atua como repressor do operon *pel*; curiosamente, a ligação do sinalizador inverte a sua regulação e FleQ passa a atuar como ativador da transcrição dos genes desse operon. Estudos recentes evidenciaram que esse processo é mediado por uma alteração significativa na estrutura quaternária proteica, decorrente da ligação ao c-di-GMP. Enquanto repressor, esse efector é um hexâmero que adota o formato de um anel e liga-se aos boxes 1 e 2 do promotor *pelA*, o que provoca uma distorção no DNA e impede o início da transcrição. Quando ligado ao c-di-GMP, o hexâmero adota uma conformação equivalente a um dímero de trímeros – nesse formato, a afinidade de interação por alguns sítios do promotor é possivelmente reduzida, o que ameniza a distorção no DNA e ocasiona a ativação da transcrição.^{29,43-44}

O fator de transcrição VpsT de *V. cholerae*, quando ligado ao c-di-GMP, regula positivamente a expressão de genes *vps* para a produção do polissacarídeo Vps, componente da matriz extracelular, ao passo que regula negativamente a motilidade. A ligação ao sinal promove a formação de um dímero de VpsT, imprescindível para o reconhecimento da sequência-alvo de DNA e para a regulação transcricional.^{37,45} É interessante observar ainda que o fator de transcrição VpsR, um regulador central de formação de biofilme em *V. cholerae*, também é um receptor de c-di-GMP e liga-se a ele para regular positivamente a transcrição de *vpsT*. Acredita-se que os sistemas de sinalização por c-di-GMP e por *quorum sensing* atuem de forma integrada para regular a transcrição de *vpsT*, já que os sítios de ligação de VpsR e do fator HapR se sobrepõem no promotor de *vpsT*. É curioso observar também que HapR, cuja expressão resulta da presença de altos níveis de indutores em

situações de alta densidade celular (*quorum sensing*), participa também da regulação da expressão de diversas enzimas DGCs e PDEs e, portanto, influencia a produção de c-di-GMP.⁸ Outro exemplo, reportado recentemente, é o do fator de transcrição BldD de *Streptomyces spp.*, envolvido na regulação da diferenciação multicelular durante a esporulação, cuja dimerização é mediada pela ligação a um tetrâmero de c-di-GMP e é necessária para a sua ligação ao DNA.^{30,33} Inversamente aos dois casos anteriores, o fator CLP de *X. campestris*, cuja sequência é altamente similar à do fator CAP (do inglês, *catabolite activator protein*), também designado CRP (do inglês, *cAMP receptor protein*), ao ligar-se ao c-di-GMP, perde a capacidade de ligar-se com alta afinidade ao DNA. O c-di-GMP, nesse caso, controla negativamente a expressão de genes de virulência bacterianos.⁴⁶

O último grupo de receptores encontrado é formado por *riboswitches* – segmentos não codificantes de RNA mensageiro (mRNA) com estruturas secundárias específicas, que se ligam a moléculas pequenas, como o c-di-GMP, para regular a expressão de genes.^{8,29,47} Esses efetores estão, geralmente, posicionados próximos a ORFs (do inglês, *open reading frames*) de enzimas do metabolismo do c-di-GMP ou então de genes regulados por esse mensageiro. Esses *riboswitches* podem, portanto, regular também a produção e degradação do seu ligante, c-di-GMP.^{29,47} A ligação do sinalizador aos *riboswitches* pode afetar a transcrição ou a tradução de genes ou ainda a estabilidade do mRNA, sendo que o controle ocorre, geralmente, em virtude do rearranjo estrutural do mRNA que é provocado pela ligação do c-di-GMP.^{8,47} Esse rearranjo pode levar, por exemplo, à formação de uma estrutura de *hairpin* que ocasione o término precoce da transcrição ou também de uma estrutura que impeça a ligação do ribossomo ao RBS (do inglês, *ribosome binding site*).⁴⁷

Há duas classes de *riboswitches* receptores de c-di-GMP: os que pertencem à primeira caracterizam-se por conter o motivo GEMM (do inglês, *Genes for the Environment, for Membranes and for Motility*); já aqueles que pertencem à segunda contêm motivo distinto e são associados a eventos de *splicing*.^{8,48-49} Um representante desta última classe é o *riboswitch* receptor de c-di-GMP do patógeno *Clostridium difficile*, o qual, associado a uma ribozima, regula a expressão de um gene de virulência. A ligação do sinalizador ao *riboswitch* estabiliza um sítio específico de *splicing* para uma ribozima próxima, o que resulta na formação de um RBS, permitindo a ativação da tradução do gene. Nessa situação, após o *splicing*, o *riboswitch* fica localizado próximo ao RBS e mantém-se como regulador da tradução do transcrito, controlando a acessibilidade desse RBS de acordo com a ligação ou não de c-di-GMP. Alternativamente, quando, na primeira etapa, não há c-di-GMP ligado ao *riboswitch*, o *splicing* ocorre em um sítio distinto, resultando em um transcrito truncado.⁵⁰

Embora múltiplos exemplos de receptores pertencentes a essas classes, juntamente com seus alvos e processos, tenham sido revelados em um curto (e recente) espaço de tempo, a distribuição dos domínios metabólicos do c-di-GMP nos genomas bacterianos ainda indica a probabilidade de que existam outros tipos de efetores.⁵¹

Uma tendência geral observada nas vias de sinalização já descritas (e refletida pela maioria dos exemplos revistos nesta seção) é a de que níveis aumentados de c-di-GMP, resultantes de alta atividade de DGCs, estimulam a expressão de fenótipos relacionados à formação de biofilmes, tais como a síntese de polissacarídeos da matriz e a expressão de adesinas extracelulares, e inibem, por exemplo, a mobilidade mediada por flagelos. Em contrapartida, baixos níveis desse dinucleotídeo, associados à atividade PDE, suprimem a síntese de componentes da matriz e estimulam o desenvolvimento de fenótipos relacionados à motilidade, promovendo a dispersão do biofilme e o aumento da virulência bacteriana (Figura 7).^{1,8}

Curiosamente, nem todos os exemplos já descritos podem ser preditos diretamente por esse padrão.^{8,52} Um caso particularmente interessante é o de duas DGCs ativas de *X. campestris*, as quais, por meio de interações com o regulador de resposta RpfG (uma PDE ativa que contém domínio HD-GYP), contribuem para a regulação positiva da motilidade mediada por *pili* do tipo IV.⁵³ Por sua vez, em um estudo em escala genômica da atividade e de fenótipos relacionados a DGCs e PDEs de *P. aeruginosa*, Kulasekara e colaboradores contestaram a robustez do modelo geral mencionado, apontando como exemplo duas DGCs cuja expressão não afetava o fenótipo de biofilme da bactéria, mas que, segundo suas análises, eram capazes de produzir altos níveis de c-di-GMP.⁵⁴ Em conjunto, evidências como essas expõem pontos de fragilidade da tendência de correlação mencionada, explicitando que somente a análise dos mecanismos subjacentes à esse sistema de sinalização poderá explicar plenamente os resultados observados.

1.3 Multiplicidade das proteínas envolvidas no metabolismo de c-di-GMP

Embora a primeira menção direta ao domínio GGDEF tenha sido feita no ano de 1995 em um artigo sobre o regulador de resposta PleD de *Caulobacter crescentus* – nele era reportada a relação dessa proteína com a transição entre fenótipos associados à motilidade e ao estado sésil bacteriano –, somente no ano de 1998 foi conjecturada a possibilidade de que esse domínio, juntamente com o domínio EAL, pudesse estar associado ao metabolismo da molécula de c-di-GMP.⁵⁵⁻⁵⁶ A relação entre esse dinucleotídeo, os domínios GGDEF e EAL e

a transição entre os estados sésil e móvel só seria mais explicitamente descrita seis anos mais tarde. Contudo, ainda entre os anos de 1998 e 2001, surgiu um crescente interesse pelos recém-revelados domínios, graças ao início e rápida progressão do sequenciamento de genomas inteiros, que permitiram a constatação de que os genes que codificam tais domínios têm ampla distribuição entre genomas bacterianos e, ainda mais, estão presentes múltiplas vezes em muitos deles.⁵⁷⁻⁵⁹

De fato, ainda que muito mais acentuadamente nos genomas de bactérias Gram-negativas do que de Gram-positivas, observa-se que os domínios GGDEF, EAL (e em menor proporção, HD-GYP) estão codificados diversas vezes no genoma de diversas espécies, integrando proteínas com arquiteturas variadas, estando frequentemente associados a um domínio sensor amino-terminal (veja subseção 1.2.1.2).¹ A Tabela 1 apresenta a quantidade identificada desses domínios nos genomas de alguns organismos escolhidos como exemplo.

Tabela 1 - Distribuição dos domínios GGDEF, EAL e HD-GYP em oito genomas bacterianos. A terceira coluna refere-se a proteínas que contém ambos os domínios GGDEF e EAL; a segunda e a quarta, somente domínio GGDEF ou EAL, respectivamente. Somente *C. crescentus* (Alphaproteobacteria), *B. cenocepacia* (Betaproteobacteria) e *C. difficile* (Clostridia; única Gram-positiva listada como exemplo) não pertencem à classe Gammaproteobacteria. Os organismos estão apresentados na ordem decrescente dos valores da última coluna, que apresenta a soma dos valores de todas as outras colunas para aquele organismo.

Organismo	GGDEF	GGDEF + EAL	EAL	HD-GYP	Total
<i>Vibrio cholerae</i> El Tor N16961	31	10	12	9	62
<i>Pseudomonas aeruginosa</i> PAO1	17	16	5	3	41
<i>Xanthomonas axonopodis</i> pv. <i>citri</i> str. 306	20	10	4	3	37
<i>Clostridium difficile</i> str. 630	18	18	1	0	37
<i>Escherichia coli</i> str. K-12 substr. MG1655	12	7	10	0	29
<i>Burkholderia cenocepacia</i> AU 1054	13	5	6	2	26
<i>Salmonella enterica</i> serovar Typhimurium LT2	5	7	8	0	20
<i>Caulobacter crescentus</i> CB15	4	7	3	0	14

Fonte: DISTRIBUTION...⁵¹

Essa multiplicidade de domínios cuja função na sinalização celular, em primeira análise, poderia ser suposta redundante é, ao que tudo indica, uma das premissas mais importantes para explicar a abrangência e a complexidade da sinalização por c-di-GMP. Em conjunto, essa elevada quantidade de proteínas com domínios GGDEF, EAL e HD-GYP e a grande variedade de receptores desse segundo mensageiro alicerçam, certamente, a existência dessa intrincada rede de sinalização, capaz de responder a diferentes tipos de estímulo, regulando de forma eficiente e dinâmica inúmeros fenótipos celulares importantes.^{1,6,8,28,33}

Nos últimos anos, estudos como o de Kulasekara e colaboradores, nos quais é avaliada a expressão de diversos fenótipos em um organismo em resposta a algum tipo de intervenção

na expressão ou atividade de domínios associados ao metabolismo de c-di-GMP, têm indicado que domínios que supostamente exercem a mesma função catalítica podem induzir respostas celulares bastante distintas.^{54,60,61} De acordo com esses resultados, é possível supor que o c-di-GMP metabolizado por um determinado par de domínios GGDEF e EAL, por exemplo, poderia ter um alvo celular específico, o que é bastante surpreendente, considerando a ampla quantidade de alvos controlados por essas vias de sinalização e que o c-di-GMP, por ser uma molécula pequena e solúvel, provavelmente difunde-se livremente pelo citoplasma.⁵⁴

Por fim, como se verifica na Tabela 1, há uma quantidade elevada de proteínas que contém ambos os domínios GGDEF e EAL (terceira coluna). Com algumas exceções, esses domínios estão geralmente arranjados de forma consecutiva na sequência proteica, próximos ao terminal carboxila, com o domínio GGDEF antecedendo o domínio EAL. Esse tipo de associação também é vista entre o domínio GGDEF e o domínio HD-GYP.⁸ Considerando o cenário anteriormente descrito, em que as concentrações de um sinalizador altamente difusível devem, de alguma forma, atuar em alvos bastante específicos, a existência de proteínas que potencialmente acumulam as funções DGC e PDE – e, desta forma, poderiam evitar o “vazamento” do sinal – é bastante sugestiva. Nesse contexto, seria coerente presumir ainda que a proximidade dessas proteínas com efetores e com seus alvos diretos poderia manter a sinalização restrita a uma via específica.

As proteínas com domínios GGDEF e EAL ou HD-GYP, na realidade, apesar de conservarem, em alguns casos, ambas as funções, podem também ter ambos os domínios cataliticamente inativos ou ainda, e mais comumente, podem ter somente um deles ativo.^{1,8,29} No caso da proteína PdeA de *Caulobacter crescentus*, por exemplo, o domínio GGDEF, apesar de cataliticamente inativo, liga-se ao GTP para ativar alostericamente o domínio EAL adjacente, aumentando a sua atividade PDE.^{1,29} Há casos interessantes, como o da proteína scrC de *Vibrio parahaemolyticus*, em que a modulação por duas proteínas acessórias, scrA e scrB parece definir qual das duas atividades prevalece, e o da proteína MSDGC1 de *Mycobacterium smegmatis*, que apresenta ambas as atividades, mas cujos domínios são inativos quando isolados.^{1,8} Alguns desses domínios degenerados podem também atuar como efetores (seção 1.2.2), e participar da regulação de respostas fenotípicas por meio de interações com proteínas ou RNAs alvo ou ainda com outros membros da mesma via de sinalização.⁸

1.4 Sinalização global e sinalização específica

A abundância de domínios GGDEF e EAL em espécies únicas e a grande diversidade de comportamentos celulares que respondem a essa sinalização fazem emergir uma questão intrigante: como um sistema de controle aparentemente mediado pelas concentrações intracelulares de uma pequena molécula sinalizadora pode gerir simultaneamente tantos alvos celulares diferentes? Embora, muito provavelmente, não haja uma resposta única e simples para essa questão, uma possível abordagem consiste em, inicialmente, tentar esclarecer uma outra questão correlata – a sinalização ocorre de forma global ou específica?

No modelo de sinalização global ou de baixa especificidade, as respostas fenotípicas celulares são reguladas por um *pool* comum de c-di-GMP, para o qual contribuem todas as DGCs e PDEs celulares. Nesse caso, supõe-se que todo o c-di-GMP produzido difunde-se livremente pelo citoplasma e, desta forma, a sua concentração seria uniforme em toda a célula. Neste modelo, portanto, qualquer nível de especificidade de resposta advém exclusivamente da variação de afinidade pela molécula de c-di-GMP entre os diferentes efetores.^{1,60}

Já no modelo de sinalização específica ou de alta especificidade, o c-di-GMP metabolizado por DGCs e PDEs individuais contribui para um *pool* local, ao qual se atribui a regulação de uma resposta fenotípica ou de um conjunto pequeno delas. Nesse modelo, ocorre o isolamento do sinal e da resposta a ele associada, por meio de estratégias que, acredita-se, envolvam, por exemplo, o sequestro desse sinal em complexos proteicos formados por potenciais interações entre enzimas, efetores e/ou alvos – dessa forma, devido à proximidade entre os elementos de uma via específica, o sinal ficaria espacialmente restrito –; ou ainda o sequestro temporal do sinal, mediado pela expressão diferencial de enzimas DGCs e PDEs em resposta a condições celulares e ambientais variantes no tempo. Neste último caso, devido ao fato de que essas enzimas podem conter domínios sensores que reconhecem sinais distintos, a expressão de diferentes DGCs e PDEs no tempo significaria ainda que somente alguns estímulos específicos poderiam ser reconhecidos por essas enzimas em determinado momento, levando à regulação da sua atividade e, portanto, ao controle do nível de c-di-GMP.^{1,60}

Alguns estudos recentes parecem lançar alguma luz sobre essa questão de especificidade. É o caso de um estudo publicado por Massie e colaboradores, em que foi analisada a correlação entre o aumento da concentração de c-di-GMP gerado por sete DGCs e o aumento da formação de biofilme em *V. cholerae*.^{37,60}

Todas as DGCs desse organismo foram clonadas em vetores de expressão, sob o controle do promotor Ptac, o qual é induzido por c-di-GMP. As DGCs foram selecionadas, primeiramente, por meio de ensaios MBEC (do inglês, *minimum biofilm-eliminating concentration*), que determinaram quais das DGCs de *V. cholerae* eram ativas para a formação de biofilme nas condições do ensaio. Em seguida, a formação de biofilme pelas DGCs selecionadas no primeiro teste foi examinada por citometria de fluxo. Esta técnica foi escolhida pelos autores para que, fosse possível acessar, a partir do mesmo experimento, a quantidade de c-di-GMP produzida (o que foi feito por meio de cromatografia líquida acoplada a espectrometria de massas), garantindo assim que essas medidas e as de formação de biofilme fossem feitas nas mesmas condições.⁶⁰ Desta segunda seleção, restaram nove DGCs – as outras foram eliminadas ou devido à baixa formação de biofilme observada mediante indução ou por produzirem biofilmes robustos mesmo na ausência de indução. Duas DGCs foram eliminadas por produzirem c-di-GMP em taxas excessivas quando comparadas às concentrações típicas encontradas nas células bacterianas e porque altos níveis dessa molécula podem inibir o crescimento bacteriano. Para confirmar que a formação de biofilme por essas DGCs é dependente, de fato, da produção do c-di-GMP, foram feitas mutações no sítio ativo dessas enzimas, que levaram, como esperado, à eliminação da capacidade de formação de biofilme.⁶⁰

Figura 9 - Gráficos de correlação entre formação de biofilme e concentração de c-di-GMP produzido por sete DGCs de *V. cholerae*. Em (a) é mostrada a reta obtida para a correlação global e, em (b), as retas obtidas para as correlações específicas para cada DGC. Em (a), o coeficiente de correlação r^2 está indicado no gráfico; em (b), os índices são mostrados à direita do gráfico.

Fonte: Adaptada de MASSIE.⁶⁰

A formação de biofilme e a produção de c-di-GMP foram mensuradas para oito diferentes estados de indução, para cada uma das sete DGCs eleitas. Foi evidenciada uma forte correlação entre o c-di-GMP produzido individualmente por cada uma das DGCs e o aumento na formação de biofilme; já a correlação global, entre o total de c-di-GMP produzido e o aumento total de formação de biofilme, foi baixa (Figura 9).⁶⁰ Tais resultados constituem forte evidência de que, em *Vibrio cholerae*, o sistema de sinalização mediado por c-di-GMP

funciona via alta especificidade. Além disso, as inclinações das retas de correlação entre concentração de c-di-GMP e índice de formação de biofilme apresentadas no gráfico (b) da Figura 9 indicam que algumas DGCs estão relacionadas a respostas rápidas de formação de biofilme face ao aumento da concentração de c-di-GMP, enquanto que outras, a respostas mais lentas.⁶⁰

Outro caso muito interessante em que se evidenciou a alta especificidade foi descrito por Hobley e colaboradores. *Bdellovibrio bacteriovorus* é uma bactéria que alterna entre os estilos de vida axênico e predatório de outras bactérias. Os autores demonstraram, por meio de experimentos de *knockout* em DGCs individuais desse organismo, que vias específicas de sinalização por c-di-GMP desempenham funções únicas e totalmente distintas na transição entre esses estilos de vida de *B. bacteriovorus*. A deleção do gene da DGC DgcB desse organismo resulta em uma bactéria incapaz de entrar no modo de vida predatório; inversamente, o mutante da DGC DgcC é um predador obrigatório; já a deleção do gene da DGC DgcA inviabiliza a motilidade do tipo *gliding* desse organismo.⁶¹

Juntamente a esses, outro exemplos, como o do já mencionado estudo de Kulasekara e colaboradores, em que a formação de biofilme por algumas das DGCs examinadas não se correlaciona com as alterações nas concentrações de c-di-GMP por elas produzido, atestam pela alta especificidade de sinalização.^{37,54} Na realidade, a possibilidade de que *pools* discretos e um ou mais *pools* comuns coexistam na célula e de que ambos os tipos integrem de alguma forma e em diferentes escalas esse sistema de regulação é também bastante plausível – de fato, a integração entre ações globais e locais de sinalização já foi evidenciada (veja subseção 1.5).^{8,62}

1.5 Interações proteicas nas vias de sinalização mediadas por c-di-GMP

Conforme discutido na seção anterior, a existência de complexos proteicos ou, de outra forma, microcompartimentos que segreguem as moléculas de c-di-GMP formadas por DGCs e PDEs individuais em uma via específica, mantendo-as próximas tanto do efetor quanto de seu alvo (ou restrito conjunto de alvos) e evitando o seu “escape” para outras vias, é um dos mecanismos moleculares propostos para a manutenção de um sistema de alta especificidade.¹ Criar proximidade entre os elementos envolvidos em uma determinada via de sinalização é, portanto, o ponto-chave implícito no conceito de microcompartimentos e de sequestro espacial ou funcional de c-di-GMP. Um tipo de mecanismo molecular efetivo para a

colocalização desses elementos é a interação proteica direta entre membros da via de sinalização.⁸

Com efeito, vários exemplos de interações proteicas entre membros de vias de sinalização já foram revelados. É interessante, no entanto, observar que esse tipo de mecanismo parece ter ascendido de uma função talvez pressuposta secundária na sinalização para assumir papéis regulatórios centrais em algumas vias de sinalização. Em alguns casos, a regulação da via pode ter tornado-se inteiramente dependente de interações proteicas entre esses componentes, dispensando até mesmo a mediação do segundo mensageiro c-di-GMP.^{1,63} Há, de fato, muitas proteínas com domínios degenerados envolvidas em interações proteicas e, embora possam não exercer função catalítica e até mesmo tenham perdido sua capacidade de ligação ao substrato, por vezes ainda desempenham papéis associados à sinalização por c-di-GMP, mediando, por exemplo, a regulação da formação de biofilme.⁸ É o caso da proteína YcgF de *E. coli*, cujo domínio EAL não se liga ao c-di-GMP e, conseqüentemente, também não o degrada. YcgF contém ainda o domínio sensor de luz azul BLUF. Quando exposta à luz azul, esta proteína interage com o regulador transcricional YcgE; este repressor, por sua vez, libera um sítio operador na molécula de DNA, desimpedindo a expressão de um conjunto de genes que codificam, dentre outras, as proteínas YmgA and YmgB, as quais estão relacionadas à produção de ácido colânico, um polissacarídeo da matriz extracelular, e à repressão da síntese de fímbrias *curli*.⁶³

Um exemplo interessante, descrito por Ryan e colaboradores, é o de duas proteínas com domínios GGDEF ativos que formam um complexo com o regulador de resposta RpfG – por meio de interações com o seu domínio HD-GYP – e com uma proteína com domínio PilZ. Este complexo, por sua vez, interage com as proteínas PilU e PilT para regular positivamente a motilidade dependente de *pilus* em *Xanthomonas campestris*.^{53,64} RpfG contém, além do domínio HD-GYP, um domínio sensor REC e participa de um sistema de dois-componentes (veja subseção 1.2.1.2) com a quinase de histidina (HK) RpfC, a qual tem um domínio sensor associado à membrana, que reconhece o sinal intercelular de *quorum sensing* DSF (do inglês, *Diffusible Signal Factor*). Por meio de fusão com GFP (do inglês, *green fluorescent protein*), os autores verificaram que o regulador de resposta RpfG é direcionado aos pólos da célula durante a sinalização por DSF. O reconhecimento desse sinal pela HK RpfC desencadeia a sua autofosforilação, seguida da fosforilação do domínio REC da RpfG. Segundo proposição dos autores, isso levaria a uma alteração conformacional no domínio HD-GYP que seria favorável à sua interação com os dois domínios GGDEF, verificada por experimentos de

FRET (do inglês, *Förster / Fluorescence Resonance Energy Transfer*), e, paralelamente, ao aumento da sua atividade PDE.⁵³

Experimentos de duplo-híbrido em levedura, usando como iscas os domínios GGDEF, identificaram que uma proteína com domínio PilZ interagia com ambos os domínios, o que foi confirmado ainda por experimentos de *Far-Western blotting* e análises por FRET. Os autores demonstraram ainda que as interações entre o domínio HD-GYP e os domínios GGDEF e o recrutamento de PilZ são induzidos pela sinalização por DSF.⁶⁴ Por fim, novos experimentos de duplo-híbrido em levedura – usando a proteína com domínio PilZ como isca –, *Far-Western blotting* e análises por FRET indicaram que as proteínas PilU e PilT, que são requeridas para a motilidade mediada por *pilus* em *X. campestris*, interagem com a proteína com domínio PilZ. Concluiu-se, portanto, que o complexo proteico induzido pela sinalização por DSF, formado por RpfG e pelas proteínas com domínios GGDEF, recrutam a proteína com domínio PilZ, a qual por sua vez interage, independentemente de c-di-GMP, com PilU e PilT para regular positivamente a motilidade mediada por *pilus* em *X. campestris*.⁶⁴

Complementarmente, as análises feitas demonstraram também que tanto a formação desse complexo quanto a regulação da motilidade não requerem a ligação de c-di-GMP a PilZ e que a atuação dos domínios GGDEF depende somente do motivo de interação ao domínio HD-GYP e não de suas atividades catalíticas.⁶⁴ É interessante ainda apontar que a proteína RpfG, apesar de, segundo estes estudos, não desempenhar função catalítica na regulação da motilidade, participa da regulação de outros fenótipos (notadamente, síntese de fatores de virulência, além da formação de biofilme) exercendo sua atividade PDE.⁶⁴

Já Lindenberg e colaboradores propuseram um modelo envolvendo o conceito de enzimas gatilho para tentar explicar o porquê de algumas das proteínas DGCs e PDEs envolvidas na expressão de CsgD, um regulador de transcrição que ativa genes para a produção de celulose e fímbrias *curli* em *E. coli*, terem uma atuação mais específica, participando unicamente da regulação desse processo, enquanto que outras tem uma atuação mais ampla, participando da regulação de outros processos – no caso, da regulação de motilidade. Nomeadamente, foi evidenciado que as DGCs YegE e YdaM e as PDEs YhjH e YciR têm efeitos na regulação da expressão de CsgD, mas que somente YegE e YhjH também têm efeitos na regulação da motilidade.^{62,65}

Os resultados obtidos pelos autores os levaram à concepção de um modelo de regulação no qual YegE, YdaM, YhjH e YciR são agrupados em dois módulos. Nesse modelo, o módulo I, constituído por YegE e YhjH, controla o módulo II, formado por YdaM

e YciR, o qual é um “interruptor” (controlado pelo módulo I) para a ativação da transcrição do gene *csgD*, mediada pelo fator de transcrição MlrA (Figura 10).⁶⁵ A DGC YegE e a PDE YhjH controlam os níveis de um *pool* de c-di-GMP: o estilo de vida celular planctônico corresponderia à prevalência da ação PDE de YhjH; já no estilo de vida celular sésil, predominaria a ação DGC de YegE.^{62,65} Quando a ação de YegE predomina, a produção de c-di-GMP é elevada e esta molécula promove o acionamento do “interruptor” da expressão de *csgD*, ou seja, do módulo I; além disso, as moléculas de c-di-GMP originárias desse *pool* ativam também outra proteína, a YcgR, que interfere negativamente na motilidade (veja subseção 1.2.2).^{39,62,65}

Figura 10 - Modelo de regulação da expressão de CsgD, mediada por dois módulos, um de ação global e outro de ação local, vinculados por c-di-GMP e pela enzima gatilho YciR.

Fonte: Adaptada de JENAL.⁶⁵

O módulo II é na verdade um complexo mediado pela interação entre as proteínas YdaM e YciR. As duas proteínas interagem ainda com o fator MlrA, formando um complexo dinâmico entre essas três proteínas. As análises ainda revelaram que YciR é essencial para que o sinal “enviado” pelo módulo I seja reconhecido e ative o módulo II, pois essa proteína tem um domínio EAL que não somente é cataliticamente ativo como também atua como sensor de c-di-GMP nesse sistema. A ligação de c-di-GMP a YciR é o que determina a ativação do complexo, embora o mecanismo exato pelo qual isso ocorre não tenha sido especificamente determinado.^{62,65} Foi demonstrado também que, apesar de YciR inibir a atividade DGC de YdaM, isso não é determinante para a repressão da expressão de *csgD* e

que essa repressão deve ser ocasionada pelas interações de YciR com MlrA e com YdaM. Pressupõe-se que, quando YciR liga-se ao c-di-GMP, estas interações se modificam, YdaM ativa MlrA e, conseqüentemente, a transcrição do gene *csgD* e a expressão de fenótipos controlados pelo regulador CsgD. Os autores ainda supõem que o c-di-GMP que é sintetizado por YdaM possa promover um *feedback* positivo, contribuindo para o mesmo *pool* mantido por YegE e YhjH e estimulando ainda mais a ativação do módulo II e expressão de CsgD. Por fim, é interessante destacar a ação como enzima gatilho de YciR, vinculando as ações globais do módulo I com as ações locais do módulo II, por intermédio do c-di-GMP.^{62,65}

Guzzo e colaboradores, por sua vez, determinaram a estrutura cristalina de um complexo formado pelas proteínas PilZ e FimX (pelo domínio EAL, especificamente) de *X. citri* subspécies citri e por c-di-GMP. Os autores ainda propuseram um modelo para a atuação do complexo na regulação da atividade motora mediada por *pilus*, no qual, em paralelo, foram ainda inclusos os resultados anteriormente obtidos por Ryan e colaboradores, já descritos nesta seção.⁶⁶

Recentemente, foi descrito um sistema relacionado à resposta de dispersão do biofilme induzida por glutamato em *P. aeruginosa*, envolvendo a formação de um complexo entre as proteínas NicD (uma DGC), DipA (uma PDE) e BdlA. Roy e Sauer propuseram um modelo em que NicD, ao reconhecer um sinal indutor de dispersão (como o glutamato) por meio de seu domínio sensor periplasmático, seria desfosforilada, o que ocasionaria um aumento momentâneo da sua atividade DGC e, portanto, do nível local de c-di-GMP.⁶⁷ Nesse modelo, a ativação de NicD, promoveria a ativação por fosforilação da proteína sensora BdlA, com a qual forma um complexo, juntamente com a PDE DipA. Tanto o aumento momentâneo no nível de c-di-GMP quanto a fosforilação provocariam também a clivagem de BdlA por um processo que requereria ainda a chaperona ClpD e a protease ClpP. Uma vez clivada, BdlA promoveria a ativação de DipA, resultando na diminuição dos níveis de c-di-GMP e favorecendo a dispersão do biofilme. Os autores ainda propuseram que outra PDE, a RbdA, que também interage com BdlA, possivelmente poderia ser também recrutada para o complexo, contribuindo ainda mais para a diminuição nos níveis de c-di-GMP e para a dispersão do biofilme.⁶⁷

2 OBJETIVOS

Em vista do surgimento de evidências isoladas de que interações proteicas entre membros constantes das vias de sinalização por c-di-GMP podem desempenhar papel fundamental na regulação e transmissão de sinal nessas vias, é possível pressupor que esse tipo de mecanismo (regulação mediada por interações proteicas) não esteja restrito a alguns casos particulares, mas que seja possivelmente frequente nesses sistemas de sinalização.

Nesse sentido, o presente trabalho teve como objetivo averiguar a relevância desse tipo de interação para o sistema de sinalização como um todo, investigando, para isso, a rede global de interações entre proteínas que contêm domínios GGDEF, EAL, HD-GYP e PilZ de um organismo modelo para estudos de formação de biofilme – *Pseudomonas aeruginosa* (linhagem PA14), uma bactéria patogênica oportunista Gram-negativa, capaz de infectar uma variedade de hospedeiros, incluindo humanos.⁶⁸

Para tanto, foram metas específicas deste trabalho:

- (i) A construção de duas bibliotecas de DNA direcionadas, por meio da clonagem dos genes de *P. aeruginosa* que codificam domínios GGDEF, EAL, HD-GYP e PilZ nos vetores “isca”, pBT, e “presa”, pTRG, a partir das quarenta e nove proteínas que contêm tais domínios identificadas nesse organismo.
- (ii) A execução de ensaios de duplo-híbrido bacteriano entre todas as proteínas de *P. aeruginosa* que contêm os domínios supracitados para a identificação de possíveis parceiros de interação.
- (iii) A obtenção de um mapa representativo das interações proteicas envolvidas nas vias de sinalização mediadas por c-di-GMP de *P. aeruginosa*.

3 MATERIAIS E MÉTODOS

3.1 Principais técnicas utilizadas

3.1.1 Método de duplo-híbrido bacteriano

O sistema de duplo-híbrido bacteriano é empregado para detecção de interações proteína-proteína *in vivo*, as quais são evidenciadas pela ativação de genes repórteres. Esse sistema utiliza células de uma linhagem de *Escherichia coli* como hospedeiras para os ensaios de interação, trazendo, assim, algumas vantagens em relação ao sistema que utiliza leveduras, devido ao crescimento muito mais rápido e à maior eficiência de transformação de *E. coli*, características que permitem que um grande número de análises seja feito em menor tempo.⁶⁹⁻

⁷⁰ A descrição do sistema feita nesta subseção é baseada em um sistema comercializado pela Agilent Technologies, o kit *BacterioMatch II Two-Hybrid System*.

O sistema dispõe de dois vetores para a clonagem das sequências codificantes das proteínas de que se quer verificar a interação: o vetor “isca”, pBT, em que é clonado o gene de uma proteína de interesse (a qual será a “isca” da interação), e o vetor “presa”, pTRG, com o qual se constrói uma biblioteca para encontrar possíveis parceiros de interação da proteína de interesse (essas proteínas serão as “presas” da interação) (Figura 11).⁶⁹⁻⁷⁰

O vetor “isca” contém uma sequência que codifica para o repressor λ (λ cI), do bacteriófago λ , o qual, por sua vez, contém um domínio de ligação ao DNA. Essa sequência está arranjada de tal forma no vetor que, quando induzida a expressão, o repressor λ fica fusionado à proteína de interesse. Da mesma forma, o vetor “presa” contém uma sequência que codifica para a subunidade α da RNA polimerase (RNAP α), a qual fica fusionada à proteína “presa” quando induzida a expressão.⁶⁹⁻⁷⁰

A linhagem repórter de *E. coli* utilizada neste sistema carrega um plasmídeo que contém a sequência do operador λ , à qual o repressor λ se liga. É importante ressaltar que esse repressor não tem função repressora neste sistema – é somente usado por exibir o domínio de ligação ao DNA. Tal plasmídeo contém ainda as sequências dos genes repórteres, sendo que, entre elas e o operador λ , está a sequência promotora desses genes, à qual se liga a subunidade α da RNA polimerase.⁶⁹⁻⁷⁰

Quando a linhagem repórter é cotransformada com os vetores “isca” e “presa” e a expressão é induzida, o repressor λ , que está fusionado à “isca”, liga-se ao operador λ presente no plasmídeo da linhagem, ancorando a “isca” para essa região. Caso “isca” e “presa”

interajam, a subunidade α da RNA polimerase, que está fusionada à “presa”, é recrutada para a região promotora, ativando a transcrição dos genes repórteres.⁶⁹⁻⁷⁰ Esses eventos estão esquematizados na Figura 12.

Figura 11 - Mapas dos vetores do *BacterioMatch II Two-Hybrid System Vector Kit*: (a) pBT e (b) pTRG.

Fonte: Adaptada de BACTERIOMATCH....⁶⁹

Figura 12 - Esquemática da sequência de eventos de um ensaio de duplo-híbrido em nível molecular. Após a clonagem dos possíveis parceiros de interação nos vetores “isca” e “presa”, a cotransformação na linhagem repórter e a coexpressão das proteínas fusionadas, ocorre a interação entre as proteínas “isca” e “presa” e a consequente ativação da transcrição dos genes repórteres, resultando na seleção positiva. O vetor “isca” carrega o gene de resistência ao cloranfenicol (indicado em azul); o vetor “presa” carrega o gene de resistência à tetraciclina (indicado em laranja).

Fonte: Adaptada de BACTERIOMATCH...⁷¹

São dois os genes repórteres: o primeiro é o gene *HIS3*, que complementa uma mutação (dessa linhagem repórter) no gene que codifica para uma enzima da via biossintética da histidina; o segundo é o gene *aadA*, que confere resistência à estreptomicina. Quando não há interação entre “isca” e “presa”, a linhagem não cresce em meio deficiente em histidina, contendo um inibidor competitivo da enzima (3-amino-1,2,4-triazol ou 3-AT).⁶⁹ A adição desse inibidor é necessária, pois, sem a ativação transcricional, o produto do gene *HIS3* é ainda produzido, mesmo que a baixos níveis, o que permite o crescimento da linhagem em meio deficiente em histidina. Logo, somente quando há interação, há ativação transcricional e a enzima é produzida a níveis suficientes para superar a inibição competitiva do 3-AT, permitindo o crescimento bacteriano nesse meio seletivo. Com a adição de estreptomicina à cultura, eleva-se ainda mais a pressão seletiva, devido à presença do gene repórter secundário.

3.1.2 *Ligase Independent Cloning* (LIC)

A técnica de *Ligase Independent Cloning* (LIC) baseia-se na atividade exonucleásica 3'-5' da T4 DNA polimerase para gerar pontas coesivas em vetores e insertos, os quais são, posteriormente, anelados na ausência da enzima ligase, baseando a ligação dos fragmentos de DNA somente na complementaridade e especificidade das pontas criadas.⁷²⁻⁷³ A clonagem independente, portanto, não só da execução de uma reação de ligação anteriormente à transformação da célula, mas também de enzimas de restrição, o que torna o processo de clonagem muito mais objetivo.

Na ausência de dNTPs que possam ser adicionados à fita de DNA, a T4 DNA polimerase passa a remover nucleotídeos da cadeia no sentido 3'-5', ao invés de estendê-la no sentido 5'-3'. Fundamentada nisso, a técnica de LIC consiste na adição de extensões contendo somente três dos quatro tipos de nucleotídeos nas fitas de DNA do vetor e do inserto. Desta forma, quando os fragmentos de DNA são tratados com a T4 DNA polimerase unicamente na presença do tipo de dNTP complementar ao que não foi usado para gerar a extensão, os nucleotídeos começam a ser removidos a partir da extremidade 3' da fita.⁷²⁻⁷³ Isso acontece até que seja encontrado o primeiro nucleotídeo do tipo que não foi incluído na extensão – colocado intencionalmente após o final da extensão, antes do início da sequência original do fragmento –, o que resulta no equilíbrio entre as atividades polimerase e exonuclease da T4 DNA polimerase, criando um *overhang* em ambas as extremidades tanto do vetor quanto do inserto. Como as extensões são desenhadas de forma que haja complementaridade específica entre os *overhangs* gerados no vetor e no inserto, formam-se pontas coesivas que permitem a ligação corretamente orientada entre os fragmentos. A ligação é baseada unicamente na complementaridade entre essas pontas. A Figura 13 ilustra o método, utilizando como exemplo o vetor pBT modificado para LIC e a representação de um inserto amplificado.

Uma vez amplificados, vetores e insertos são tratados com a enzima T4 DNA polimerase e são, então, incubados juntos (na ausência da enzima ligase) para que ocorra o anelamento. O produto do anelamento é usado para a transformação de células competentes de *E. coli* e os vetores recombinantes são extraídos e analisados para a identificação de clones positivos.

a)

Vetor pBT modificado

Fim da sequência da
proteína de fusão (λ cI) —————

5'...GAA GAG ACG TTT GGC GCG GCC GCT **CAG GGC GCC** ATG...inserto...TGA **CCG CGT CGG GTC AAT TAA GAT**...3'
3'...CTT CTC TGC AAA CCG **CGC CGG CGA GTC CCG CGG** TAC...inserto...ACT **GGC GCA GCC CAG** TTA ATT CTA...5'

Vetor pTRG modificado

Fim da sequência da
proteína de fusão (RNAP α) —————

5'...AAA CCA GAG GCG GCC GGA TCC GCG GCC GCT **CAG GGC GCC** ATG...inserto...
3'...TTT GGT CTC CGC CGG **CCT AGG CGC CGG CGA GTC CCG CGG** TAC...inserto...
...TGA **CCG CGT CGG GTC AAT TAA TGA** ACT AGT GAG ATC CTC TAG AGT CGA C...3'
...ACT **GGC GCA GCC CAG** TTA ATT ACT TGA TCA CTC TAG GAG ATC TCA GCT G...5'

b)

Vetor pBT modificado

T4 DNA polimerase, na presença de dTTP
(ausência de dATP, dCTP e dGTP)

5'...GAA GAG ACG TTT GGC GCG GCC GCT **CAG GGC GCC A** ← atividade
exonucleásica **CCG CGT CGG GTC AAT TAA GAT**...3'
3'...CTT CTC TGC AAA CCG **CGC CGG CGA GTC CCG CGG T** → atividade
exonucleásica **GGC GCA GCC CAG** TTA ATT CTA...5'

Inserto amplificado

T4 DNA polimerase, na presença de dATP
(ausência de dTTP, dCTP e dGTP)

5' **CAG GGC GCC** ATG...inserto...TGA **CCG CGT CGG GTC** 3' ← atividade
exonucleásica
3' **GTC CCG CGG** TAC...inserto...ACT **GGC GCA GCC CAG** 5' → atividade
exonucleásica

(continua)

(continuação)

Figura 13 - Esquema ilustrativo da técnica de LIC (*Ligase Independent Cloning*). Em verde, são mostradas as sequências correspondentes aos *primers* utilizados para a modificação dos vetores e para a amplificação dos insertos; em vermelho, as extensões que serão removidas pela T4 DNA polimerase para a criação de *overhangs* e, em negrito, o ponto em que ocorre o equilíbrio entre as atividades exonucleásica e polimerásica da T4 DNA polimerase. As extensões dos vetores são compostas somente por adeninas, citosinas e guaninas e são limitadas por uma timina (em negrito) na extremidade 5'; já as dos insertos, compostas por citosinas, guaninas e timinas e limitadas por uma adenina (em negrito) na extremidade 5'. Estão representados no esquema: (a) a modificação dos vetores pBT e pTRG; (b) o resultado da amplificação do vetor pBT e de um inserto a partir dos *primers* desenhados e o sentido da atividade exonucleásica da T4 DNA polimerase; (c) vetor e inserto após tratamento com a T4 DNA polimerase, na presença de dTTP e de dATP, respectivamente, evidenciando os *overhangs* gerados; (d) anelamento entre o vetor e o inserto tratados.

Fonte: Adaptada de CAMILO.⁷³

3.2 Materiais e métodos

3.2.1 Compilação das sequências alvo e remoção de porções transmembranares

A compilação de todas as sequências nucleotídicas e proteicas codificantes de proteínas de *P. aeruginosa* que contêm domínios GGDEF, EAL, HD-GYP e PilZ foi feita por meio de consultas a bancos de dados disponíveis *online*. Especificamente, as sequências foram recuperadas do banco KEEG (*Kyoto Encyclopedia of Genes and Genomes*) e, para a identificação da arquitetura de domínios de cada proteína, foi utilizada a ferramenta de análise

de sequências proteicas do banco Pfam. Esta etapa havia sido cumprida previamente pela Dr. Ana Isabel de Camargo.

Uma vez compiladas todas as proteínas alvo, suas sequências foram submetidas à análise pelo programa de predição de hélices transmembranares TMHMM (*TransMembrane prediction using Hidden Markov Models*) – versão 2.0 – para a distinção entre porções transmembranares, porções citoplasmáticas e porções periplasmáticas. Os dados obtidos por essa análise permitiram que as sequências de porções transmembranares fossem removidas, o que foi feito com a intuito de evitar problemas de solubilidade na execução das etapas seguintes do projeto, assegurando o não comprometimento dos experimentos de duplo-híbrido. Após a remoção das porções transmembranares, restaram somente construções com porções proteicas preditas solúveis.

3.2.2 Construção das bibliotecas de DNA de “iscas” e “presas” por meio da técnica de dupla-restrição enzimática

3.2.2.1 Desenho de *primers* para clonagem por dupla-restrição enzimática

Uma vez definidas as construções proteicas que seriam os insertos para clonagem em ambos os vetores pBT e PTRG – e, portanto, para a construção das bibliotecas de “iscas” e “presas” respectivamente –, a próxima etapa foi a definição dos pares de enzimas de restrição adequados para permitir a posterior clivagem de insertos e vetores e a ligação específica e orientada entre eles. Para tanto, foram analisados os sítios múltiplos de clonagem (MCS – do inglês, *Multiple Cloning Site*) nos mapas dos vetores pBT e pTRG (Figura 11), os quais estão ilustrados no manual de instruções (revisão C) do *BacterioMatch II Two-Hybrid System Vector Kit* (Agilent Technologies), adquirido para o projeto, o qual fornecia além dos vetores, as linhagens de propagação e repórter (ambas originárias de *E. coli* XL1-Blue). É importante ressaltar que os MCSs de pBT e pTRG não são idênticos e, portanto, foram necessárias análises separadas voltadas para cada um dos vetores. Além disso, no manual, é indicado que a endonuclease Not I, presente no MSC de ambos pBT e pTRG, por codificar um *linker* de alanina, facilitaria a orientação e o enovelamento das proteínas fusionadas – por essa razão, quando possível, Not I foi utilizada preferencialmente. Além da análise dos mapas dos vetores, foi utilizado a ferramenta *NEBcutter* (NEB, New England Biolabs), disponível *online*, para determinar quais, dentre as endonucleases que tivessem sítios nos vetores, não clivariam internamente as sequências nucleotídicas das construções. As sequências foram

examinadas uma a uma e foram definidos os pares de enzimas de restrição que seriam utilizados para cada caso.

Em seguida, foi utilizada a ferramenta *Oligo PerfectTM Designer* (Invitrogen / Thermo Fisher Scientific), também disponibilizada *online* para gerar os *primers* que seriam utilizados para a amplificação por PCR (*Polymerase Chain Reaction*) dos insertos. Foram utilizados parâmetros relativamente amplos de entrada, os quais, sempre que possível, foram: tamanho entre 15 pb e 30 pb (tamanho ótimo de 20 pb), temperatura de desnaturação (T_m) entre 50 °C e 70 °C (T_m ótima de 60 °C) e conteúdo GC entre 20% e 80% (conteúdo GC ótimo de 50%). Foram inseridas também como dados de entrada na ferramenta as enzimas de restrição previamente eleitas.

O desenho dos pares de *primers* foi concluído com a inclusão de códons de parada nos *primers* reversos daquelas sequências que, por terem o terminal carboxila interno à sequência proteica original, não os continham. Para garantir que as sequências dos insertos, correspondentes às “iscas” e às “presas”, estariam no mesmo quadro de leitura das sequências correspondentes às proteínas a que estariam fusionadas, foi examinado e ajustado, quando necessário, o quadro de leitura, pela adição de uma ou duas bases alternativamente, após a sequência do sítio de reconhecimento da enzima de restrição nos *primers* diretos. Por fim, foi adicionada uma pequena extensão (CATG) no início das sequências dos *primers* para fornecer suporte estrutural para a clivagem dos produtos de amplificação pelas enzimas de restrição nas etapas seguintes.

3.2.2.2 Clonagem das construções selecionadas nos vetores pBT e pTRG por meio de dupla-restrição enzimática

Para a clonagem das sequências nos vetores pBT e pTRG fornecidos pelo *BacterioMatch II Two-Hybrid System Vector Kit* (Agilent Technologies), as sequências foram amplificadas a partir dos *primers* desenhados e do DNA genômico de *Pseudomonas aeruginosa* PA14 previamente purificado. Foram também utilizadas, alternativamente, as enzimas *High Fidelity PCR Enzyme Mix* (Fermentas / Thermo Fisher Scientific) e *Phusion® High-Fidelity DNA Polymerase* (NEB), com soluções tampão respectivas, e solução de dNTPs 10 mM cada (*Thermo Scientific*). Alternativamente, DMSO (dimetilsulfóxido) e soluções de BSA (do inglês, *Bovine Serum Albumin*) foram também utilizados para melhorar a especificidade e quantidade de produto obtido. Como os tamanhos e conteúdos de cada sequência eram bastante variados, as quantidades usadas nas reações e parâmetros de tempo e

temperatura aplicados nos ciclos de reação foram também bastante variados. Após amplificadas, as sequências cujos produtos foram inespecíficos foram purificadas em gel de agarose 0,8%. Todos os produtos de amplificação foram purificados com o *kit* de purificação *Wizard® SV Gel and PCR Clean-Up System* (Promega).

Os vetores pBT e pTRG foram propagados na linhagem de *E. coli* fornecida pelo *kit* *BacterioMatch II Two-Hybrid System* (Agilent Technologies) e foram extraídos em quantidades estoque para a execução das clonagens, utilizando o *kit* *Wizard® Plus SV Minipreps DNA Purification System* (Promega). Assim como os produtos de amplificação dos insertos, os vetores foram duplamente clivados com as enzimas de restrição adequadas (escolhidas de acordo com os passos descritos anteriormente) – em geral, por praticidade devido ao volume de amostras, foram utilizadas enzimas *FastDigest* (Fermentas / Thermo Fisher Scientific) – e foi feita a purificação em gel de agarose 0,8% dos produtos obtidos (no caso dos vetores), utilizando também o *kit* *Wizard® SV Gel and PCR Clean-Up System* (Promega) (para vetores e insertos). As reações de clivagem foram montadas de acordo com indicações do fabricante das enzimas, com exceção do tempo de reação que foi, em geral, estendido para até 1 hora (em estufa a 37 °C).

Com insertos e vetores já purificados, foram montadas reações de ligação, utilizando a enzima *T4 DNA Ligase* (Invitrogen / Thermo Scientific), de acordo com as indicações do fabricante – novamente, como os tamanhos e conteúdos das sequências eram bastante variados, as quantidades utilizadas nas reações foram também variadas. As reações foram utilizadas para a transformação de células quimiocompetentes (previamente preparadas) da linhagem de *E. coli* para propagação fornecida pelo *kit* de duplo-híbrido. Os vetores recombinantes foram então propagados e extraídos, utilizando o *kit* *Wizard® Plus SV Minipreps DNA Purification System* (Promega). Para a validação dos clones obtidos, foram cumpridas etapas de análise de restrição com eletroforese em gel de agarose 0,8% e de sequenciamento. Os sequenciamentos foram executados pelas técnicas Andressa Patrícia Alves Pinto e Patrícia Fernandes Matheus após fornecimento de soluções contendo *primer* e amostra para cada reação.

3.2.3 Construção das bibliotecas de DNA de “iscas” e “presas” por meio da técnica de *Ligase Independent Cloning* (LIC)

3.2.3.1 Modificação dos vetores pBT e pTRG e desenho de *primers* para LIC

Em parceria com o pesquisador Dr. César Moisés Camilo, do Grupo de Biotecnologia Molecular (IFSC-USP), coordenado pelo Prof. Dr. Igor Polikarpov, foram executadas as clonagens das construções por meio da técnica de *Ligase Independent Cloning* (descrita na subseção 3.1.2), já bem estabelecida pelo pesquisador no grupo mencionado e mais adequada à clonagem em larga escala. Além disso, algumas etapas de clonagem foram executadas na plataforma automatizada *Freedom EVO 200* (Tecan) alocada no Polo TErRA (Polo Temático em Energias Renováveis e Meio Ambiente) do mesmo grupo e pela qual o pesquisador parceiro era responsável.

Conforme indicado na subseção 3.1.2, para que a técnica pudesse ser aplicada à clonagem em pBT e pTRG, foi necessária a modificação destes vetores pela adição de extensões LIC nas pontas desses vetores linearizados, idênticas para pBT e pTRG, para a formação dos *overhangs*. No caso, as extensões, desenhadas pelo pesquisador parceiro para a modificação dos vetores, foram 5'-TGGCGCCCTGA-3' e 5'-CCGCGTCGGGTCA-3', as quais, a menos das últimas bases 3', constituem os *overhangs*. Experimentalmente, essa modificação foi feita por meio da amplificação dos vetores por PCR, utilizando a enzima de alta fidelidade e processividade *Phusion® High-Fidelity DNA Polymerase* (NEB), a partir de *primers* contendo essas sequências de extensão e tendo como molde os vetores originais. Para facilitar a amplificação, os vetores foram previamente linearizados por digestão com EcoRI e XhoI. Na Figura 14, (idêntica à imagem (a) da Figura 13, subseção 3.1.2), os *primers* estão indicados em verde e, em vermelho, as sequências complementares aos *overhangs*, ou seja, as bases que serão removidas pela T4 DNA polimerase. É essencial observar ainda que a clonagem de qualquer inserto que contenha *overhangs* correspondentes às sequências em vermelho é compatível com ambos os vetores, já que, por sua vez, as sequências dos *overhangs* dos vetores são idênticas para pBT e pTRG.

Uma vez amplificados, os vetores foram tratados com a enzima DpnI para garantir a eliminação de pBT e pTRG originais utilizados como molde, de forma que não interferissem nas etapas de clonagem subsequentes. Foram também purificados em gel de agarose e extraídos utilizando o kit *Wizard® SV Gel and PCR Clean-Up System* (Promega).

Os *primers* para a amplificação dos insertos foram desenhados da mesma forma, mas continham sequências de extensão LIC complementares às dos *primers* para modificação dos vetores, de forma que, após o tratamento com a T4 DNA polimerase, vetores e insertos tivessem *overhangs* complementares, permitindo o anelamento. Adicionalmente, foi utilizada a ferramenta *online* HTP-OligoDesigner (*High Throughput OligoDesigner*), desenvolvida pelo próprio Dr. César Camilo, para o desenho dos *primers*.⁷⁴

Vetor pBT modificado

Fim da sequência da
proteína de fusão (λ cI) —————
5'...GAA GAG ACG TTT GGC GCG GCC GCT **CAG GGC GCC** ATG...inserto...TGA **CCG CGT CGG GTC AAT TAA GAT**...3'
3'...CTT CTC TGC AAA CCG **CGC CGG CGA GTC CCG CGG** TAC...inserto...ACT **GGC GCA GCC CAG** TTA ATT CTA...5'

Vetor pTRG modificado

Fim da sequência da
proteína de fusão (RNAP α) —————
5'...AAA CCA GAG GCG GCC GGA TCC GCG GCC GCT **CAG GGC GCC** ATG...inserto...
3'...TTT GGT CTC CGC CGG **CCT AGG CGC CGG CGA GTC CCG CGG** TAC...inserto...
...TGA **CCG CGT CGG GTC AAT TAA TGA ACT AGT GAG ATC CTC TAG AGT CGA C**...3'
...ACT **GGC GCA GCC CAG** TTA ATT ACT TGA TCA CTC TAG GAG ATC TCA GCT G...5'

Figura 14 - Sítios de clonagem de pBT e pTRG modificados para a técnica de LIC (*Ligase Independent Cloning*). Em verde, são mostradas os *primers* para a modificação dos vetores; em vermelho, as extensões que serão removidas pela T4 DNA polimerase para a criação de *overhangs* e, em negrito, o ponto em que ocorre o equilíbrio entre as atividades exonucleásica e polimerásica da T4 DNA polimerase.

Fonte: Adaptada de CAMILO.⁷³

Ainda é preciso observar que, como uma das extensões LIC necessariamente fica posicionada entre o final da sequência correspondente à proteína de fusão (λ cI ou RNAP α) e o início do sítio de clonagem, a sequência adicionada estaria contida no *linker* pelo qual as duas proteínas expressas estariam fusionadas. Para minimizar o efeito que a alteração nesse *linker* poderia supostamente causar nas interações testadas adiante, buscou-se manter o seu comprimento dentro do intervalo de comprimentos originalmente previstos caso fossem utilizados os sítios múltiplos de clonagem dos vetores e, novamente, optou-se também por conservar uma sequência de três alaninas, a qual facilita a orientação e o enovelamento das proteínas fusionadas.

3.2.3.2 Clonagem das construções selecionadas nos vetores pBT e pTRG por meio da técnica de LIC

Uma vez amplificados e purificados, vetores e insertos foram submetidos, separadamente, ao tratamento com a enzima T4 DNA polimerase. Para tanto, para cada 500 ng de vetor purificado, foram adicionadas 3 unidades de T4 DNA Polimerase (NEB), 1X NEBuffer 2 (NEB), 4 mM de DTT (Sigma), 2,5 mM de dTTP (Thermo Scientific) e BSA 1X (NEB), para um volume final de 20 µL. Similarmente, no caso dos insertos, para cada 200 ng de produto purificado, foram adicionadas 3 unidades de T4 DNA Polimerase (NEB), 1X NEBuffer 2 (NEB), 4 mM de DTT (Sigma), 2,5 mM de dATP (Thermo Scientific) e BSA 1X (NEB), para um volume final de 20 µL. As reações, tanto com vetores quanto com insertos, foram incubadas por 30 minutos a 22 °C e, em seguida, por 20 min a 75 °C para inativação.

Para o anelamento entre cada inserto e cada um dos dois tipos de vetor, todos tratados com T4 DNA polimerase, foi feita a seguinte mistura padrão (a proporção entre vetor e inserto foi alterada, conforme necessário, de acordo com o tamanho dos insertos), que foi incubada a 25 °C por 30 minutos: 1 uL de vetor tratado foi misturado a 3 uL de inserto tratado.

Para a transformação, o produto da etapa de anelamento foi diluído para um volume total de 10 uL e foram adicionados 30 uL de um tampão de transformação, previamente mantido em banho de gelo, composto por 100 mM de KCl, 30 mM de CaCl₂, 50 mM de MgCl₂ e 1,5% (m/v) de PEG 4000. Células quimiocompetentes de *E. coli* XL1-Blue, a linhagem de propagação fornecida pelo conjunto *BacterioMatch II Two-Hybrid System Vector Kit*, haviam sido previamente preparadas pelo método Inoue de preparo de células ultracompetentes.⁷⁵ Para cada transformação, 50 uL dessas células competentes foram adicionados à mistura anterior. As reações de transformação ficaram em banho de gelo por 30 minutos e, em seguida, foram incubadas por mais 10 min à temperatura ambiente (25 °C). Foram adicionados então 200 µL de meio LB (Luria-Bertani) em cada transformação, as quais foram então incubadas por 1 hora a 37 °C. Após esse intervalo, as transformações foram centrifugadas a 3500 g por 5 min e grande parte do volume sobrenadante foi descartado – as células coletadas foram ressuspensas no volume restante, que foi então espalhado sobre meio sólido LB-ágar com 50 µg/mL de canamicina, contendo, adicionalmente, 25 µg/mL de cloranfenicol no caso das transformações envolvendo pBT e 12,5 µg/mL de tetraciclina no caso das transformações envolvendo pTRG. Após, em média, 24 horas em incubação a 37 °C,

as colônias obtidas foram transferidas para meio líquido (contendo os mesmos antibióticos) e cultivadas *overnight*, a 37 °C, sem agitação. Os vetores recombinantes foram então extraídos. Para a validação dos clones obtidos, as amostras foram testadas por PCR seguida de eletroforese em gel de agarose 0,8% para verificação por tamanho e, em seguida, submetidas a sequenciamento. As reações de sequenciamento foram executadas pela técnica Andressa Patrícia Alves Pinto após fornecimento de soluções contendo *primer* e amostra para cada reação.

É importante ressaltar que a construção destas bibliotecas de “iscas” e “presas” foi feita de forma direcionada; ou seja, em oposição ao que ocorre na produção de bibliotecas de cDNA ou de bibliotecas obtidas a partir da fragmentação aleatória de DNA genômico, estas bibliotecas, embora comparativamente muito menores em tamanho, foram feitas a partir da clonagem específica e planejada de cada um de seus componentes com a intenção de que se pudesse assegurar a obtenção de produtos cuja qualidade não interferisse negativamente nos experimentos posteriores.

Além disso, é válido lembrar que a escolha de que todas as sequências alvo fossem clonadas em ambos os vetores pBT e pTRG adveio da intenção de que fossem cobertas todas as possibilidades de interação entre as proteínas alvo.

3.2.4 Ensaios de duplo-híbrido bacteriano

Conforme já indicado nas subseções anteriores, tanto para a construção das bibliotecas de “iscas” e “presas” quanto para a execução dos testes de interação proteica, foi adquirido o kit *BacterioMatch II Two-Hybrid System* (Agilent Technologies), que contém todos os vetores (pBT, pTRG, pBT-LGF2 e pTRG-Gal11) e ambas as linhagens de *E. coli* XL1-Blue (tanto a linhagem para a propagação de pBT e pTRG recombinantes quanto a linhagem repórter para ensaios de interação) utilizados. Todas as soluções e meios de cultura utilizados constantemente durante a execução dos experimentos foram sempre preparados conforme orientações e protocolos fornecidos no manual do kit.⁶⁹ A solução de sais M9 foi sempre preparada conforme protocolo fornecido no manual (não foi adquirida comercialmente) e, também segundo indicado, o suplemento de aminoácidos sem histidina (*His dropout amino acid supplement*) foi adquirido da Clontech Laboratories (Takara).⁶⁹

3.2.4.1 Células competentes

As células competentes da linhagem repórter foram preparadas pelo método Inoue para obtenção de células ultracompetentes. Este protocolo foi utilizado para garantir a alta eficiência de transformação das células, imprescindível para o sucesso dos ensaios de interação. Todos os estoques foram congelados em nitrogênio líquido e conservados a - 80 °C por até quarenta dias.

3.2.4.2 Meios não seletivo, seletivo e duplamente seletivo e meio líquido HDO

Os testes de interação proteica foram avaliados por meio do cultivo das células cotransformadas com os vetores “isca” e “presa” nos meios de seleção descritos nesta subseção. Todos os meios têm como base comum, o meio mínimo sem histidina, cujo protocolo é indicado no manual de instruções do *kit* utilizado.⁶⁹ Os antibióticos utilizados, assim como o inibidor 3-AT, também foram preparados de acordo com as indicações do manual. Após prontas, as placas com meio sólido eram conservadas a 4 °C, protegidas da luz, por até uma semana.

O meio não seletivo é composto de meio mínimo sem histidina com 1,5% (m/v) ágar e 50 µM de IPTG (isopropil-β-D-tiogalactopiranosídeo) para indução da expressão das proteínas de fusão, contendo, além de 50 µL/mL de canamicina, 25 µL/mL de cloranfenicol, e 12,5 µL/mL de tetraciclina, para seleção dos plasmídeos “isca” e “presa”, respectivamente.

O meio seletivo é composto de meio mínimo sem histidina com 1,5% (m/v) ágar e 50 µM de IPTG, contendo 50 µL/mL de canamicina, 25 µL/mL de cloranfenicol, e 12,5 µL/mL de tetraciclina. Contém também 5 mM do inibidor 3-AT. Esta concentração de inibidor permite a seleção de células cotransformadas com pares de interação positiva.

O meio duplamente seletivo é composto de meio mínimo sem histidina com 1,5% (m/v) ágar e 50 µM de IPTG, contendo 50 µL/mL de canamicina, 25 µL/mL de cloranfenicol, e 12,5 µL/mL de tetraciclina. Contém também, além de 5 mM do inibidor 3-AT, 12,5 µL/mL de estreptomicina. A adição deste antibiótico aumenta a pressão seletiva para pares de interação positiva.

O meio líquido HDO (*His-dropout*) é composto de meio mínimo sem histidina com 50 µM de IPTG. É utilizado para a adaptação das células cotransformadas a esse tipo de meio

(após terem sido recuperadas em meio SOC) e para o início da expressão das proteínas de fusão, anteriormente ao plaqueamento nos meios seletivos.

3.2.4.3 Controles

O sistema contém, além dos vetores pBT e pTRG para a clonagem de “iscas” e de “presas”, os vetores para controle positivo, pBT-LGF2 e pTRG-Gal11^P, um par de interação robusto. A sequência de LGF2 codifica um domínio de dimerização do ativador transcricional de levedura, Gal4, enquanto a sequência indicada como Gal11^P codifica um domínio desse mutante da proteína Gal11, também de levedura.⁷⁶ No sistema de duplo-híbrido, quando a linhagem repórter é cotransformada com ambos os plasmídeos, verifica-se um padrão forte de interação, revelado pelo crescimento abundante nas placas de meio seletivo. Este controle foi incluído em todas as rodadas de testes de interação para assegurar a efetividade do sistema, além de ter sido usado como referência de força de interação.^{69,76}

Como controle negativo, foram cotransformados o vetor pBT vazio (sem fragmento clonado) e o vetor pTRG-Gal11^P na linhagem repórter. Conforme esperado, não se observa crescimento no meio seletivo para este par de cotransformação. Este controle também foi incluído em todas as rodadas de testes para atestar a qualidade seletiva do meio, que é inteiramente dependente do inibidor 3-AT.

Como controles de eficiência de cotransformação, todas as cotransformações feitas (testes e controles positivo e negativo) sempre foram também plaqueadas em meio não seletivo. Esses controles foram usados para determinar a capacidade de cotransformar as células competentes da linhagem repórter com “iscas” e “presas”, ou, de outra forma, para permitir a detecção de falsos-negativos.

Como controles de autoativação, para “iscas” e “presas” que originaram padrões positivos de interação (por praticidade, serão designadas “iscas positivas” e “presas positivas”), foram feitas, juntamente aos ensaios de força de interação (veja subseção 3.2.4.6), cotransformações com pBT vazio e “presa positiva” e com “isca positiva” e pTRG-Gal11^P.

3.2.4.4 Padronização e otimização do protocolo dos ensaios de interação

Anteriormente ao início efetivo dos testes de interação, foi feita a padronização do ensaio de duplo-híbrido bacteriano para a determinação de um protocolo único, viável e reproduzível ao longo de todos os testes de interação, baseado nos protocolos de

cotransformação e plaqueamento indicados no manual de instruções do kit adquirido.⁶⁹ Para tanto, foram utilizados os controles descritos na subseção 3.2.4.3 (notadamente, controle positivo, controle negativo e controles de cotransformação – no caso, dos controles positivo e negativo). Nesta etapa, foi definido ainda o protocolo que seria utilizado para o preparo de células competentes (subseção 3.2.4.1) e foram testados todos os protocolos de preparo de soluções e meios de cultura (fornecidos no manual do *kit* adquirido) para, da mesma forma, garantir a sua reprodutibilidade e a manutenção da uniformidade dos ensaios ao longo de todo o período subsequente de testes de interação.

Ao final dos testes de padronização, obteve-se o protocolo descrito a seguir.

Para os testes de interação acompanhados dos respectivos controles de cotransformação, microtubos de 2 mL contendo 100 µL de células competentes da linhagem repórter foram retirados do estoque, conservado a -80 °C, e deixados sobre gelo para descongelamento. Às células, em cada microtubo, 80 ng de cada um dos vetores contendo os possíveis parceiros de interação (pBT-“isca” e pTRG-“presa”) foram adicionados. Os microtubos foram então submetidos à agitação leve e incubados em gelo por 30 minutos, sendo que, passados os primeiros 15 minutos e ao fim dos 30 minutos, foram novamente agitados. Após o período total de incubação, os tubos foram imersos por 35 segundos exatos em banho de água a 42 °C e imediatamente retirados e devolvidos ao banho de gelo, no qual foram deixados por mais 2 minutos. Ao final desse tempo, os tubos foram retirados do gelo e, a cada um, foram adicionados 900 µL de meio SOC previamente aquecido em banho de água a 42 °C. Os microtubos foram então incubados em *shaker* a 37°C, 235 rpm por 90 minutos. Ao final desse período, as células foram coletadas por centrifugação a 2000 g, por 10 min, à temperatura ambiente. O sobrenadante (meio SOC) foi descartado e, às células coletadas, foi adicionado 1 mL de meio HDO. As células foram ressuspensas por inversão até a obtenção de uma mistura homogênea e, subsequentemente, incubadas em *shaker* a 37 °C e 235 rpm por mais 2 horas para a adaptação ao crescimento em meio mínimo. Após esse período, as células foram mais uma vez coletadas por centrifugação a 2000 g, por 10 min, à temperatura ambiente. Aproximadamente 70 % do sobrenadante foi descartado e as células foram ressuspensas, por inversão, nos 30 % restantes. Metade do volume final restante em cada tubo (aproximadamente 150 uL) foi plaqueado em placa seletiva; a outra metade, em placa não seletiva (controle de cotransformação). Todas as placas foram previamente deixadas destampadas em ambiente estéril por, em média, 20 minutos para evitar umidade excessiva.

Por fim, as placas foram incubadas em estufa a 37 °C por 24 horas e, à temperatura ambiente, ao abrigo da luz, por mais 16 horas.

3.2.4.5 Estratégia de triagem por grupos, seguida de triagem por pares

Em virtude da vasta quantidade de interações potencialmente existentes entre “iscas” e “presas”, determinada pelo tamanho das bibliotecas obtidas, verificou-se a inviabilidade de averiguá-las em ensaios independentes. Em busca de um método mais eficiente para a detecção das interações, foi desenhada uma estratégia alternativa de triagem das bibliotecas – a de cotransformação com grupos de “iscas” e “presas” e busca, dentre os elementos dos grupos que revelassem interações positivas, de parceiros específicos de interação. Desta forma, todas as interações hipotetizadas entre os elementos que compusessem os grupos cujos testes de interação gerassem resultado negativo poderiam ser descartadas e aquelas cujos testes tivessem resultado positivo poderiam ser averiguadas por meio do desmembramento dos grupos e realização de novos testes de interação.

Para a execução dessa estratégia, foram determinados critérios para a distribuição de “iscas” e “presas” em grupos. Primeiramente, o número de grupos formados deveria ser tal que a quantidade de testes iniciais fosse sensivelmente reduzida; em contrapartida, como, quanto maior fosse o número de componentes de cada grupo, maior seria a probabilidade de que, entre eles, houvesse alguma interação positiva, não seria interessante que os grupos fossem compostos por muitos componentes. Nesse cenário, seriam reduzidas as chances de eliminar um grande volume de testes a partir de resultados negativos de interação entre grupos e, portanto, perderia-se a vantagem pressuposta pelo método desenhado. Secundariamente, a composição dos grupos foi definida pela formação de *clusters* de proteínas contendo mesmo tipo de domínio alvo (GGDEF, EAL, GGDEF-EAL, HD-GYP e PilZ). Como terceiro critério, foram mantidas juntas proteínas contendo mesmos tipos de domínios (além dos domínios alvo). Os dois últimos critérios foram estabelecidos para facilitar a análise de resultados posteriormente.

O protocolo para realização dos ensaios de interação entre grupos foi idêntico ao descrito na subseção 3.2.4.4 e foi também validado por testes de padronização. Conforme indicado, 80 ng de cada um dos vetores de “iscas” e “presas” foi usado em cada ensaio.

Uma vez finalizada a primeira triagem e encontrados os pares de grupos cujo resultado para interação foi positivo, iniciou-se a triagem pelos pares interagentes específicos. Nesta segunda etapa, os pares de grupos revelados foram desmembrados e foram feitos ensaios do

tipo um a um entre “iscas” e “presas”, com a finalidade de que fossem determinados quais, dentre os componentes dos grupos, eram, de fato, os parceiros de interação que haviam originado os resultados positivos observados. Os ensaios desta etapa também foram feitos conforme indicado no protocolo descrito na subseção 3.2.4.4.

3.2.4.6 Ensaios de força de interação

Para a confirmação dos testes entre parceiros que apresentaram resultados positivos para interação e como uma tentativa de mensurar a força de interação, foi elaborado um teste adicional, baseado na pressão seletiva provida pela adição do antibiótico estreptomicina, o segundo agente seletivo do sistema de duplo-híbrido utilizado, ao preparo do meio mínimo utilizado para os ensaios de interação, tornando-o duplamente seletivo (subseção 3.2.4.2).

Inicialmente, foram feitas culturas de 2 mL de meio LB, contendo os antibióticos para seleção dos vetores (25 µg/mL de cloranfenicol e 12,5 µg/ mL de tetraciclina, além de 50 µg/mL de canamicina) originadas de colônias frescas de células repórteres cotransformadas com os parceiros de interação identificados nos testes feitos anteriormente. Após 6 horas de crescimento em *shaker* a 37 °C e 235 rpm, as culturas foram centrifugadas a 2000 g, por 10 min, à temperatura ambiente. O meio sobrenadante foi descartado e, às células coletadas, foi adicionado 1 mL de meio HDO (contendo também 25 µg/mL de cloranfenicol, 12,5 µg/ mL de tetraciclina e 50 µg/mL de canamicina). Após 2 horas de incubação (em shaker a 37 °C e 235 rpm), foram feitas as seguintes diluições seriadas da cultura final obtida em meio HDO: 1/2, 1/4, 1/8, 1/16, 1/32. As diluições feitas foram então gotejadas em placa duplamente seletiva, as quais foram então cultivadas em estufa a 37 °C por 24 horas e, em seguida, foram deixadas à temperatura ambiente, protegidas da luz, por 16 horas adicionais.

O mesmo procedimento foi feito, paralelamente, para a cotransformação das células repórteres com controles de autoativação de “iscas” e “presas” (subseção 3.2.4.3) envolvidos em interações positivas. Além disso, como nos ensaios anteriores, foram também incluídos os controles positivo e negativo.

4 RESULTADOS E DISCUSSÃO

4.1 Compilação das sequências alvo e remoção de porções transmembranares

Foram identificadas, ao todo, quarenta e nove proteínas de *P. aeruginosa* (PA14) que contêm domínios GGDEF, EAL, HD-GYP ou PilZ. A Tabela 2 traz a lista com todas as proteínas compiladas. Ao todo, foram contabilizadas dezesseis proteínas com domínio GGDEF, dezesseis proteínas com ambos os domínios GGDEF e EAL (excluem-se desta conta as proteínas que contêm somente um desses domínios), sete proteínas com domínio EAL, três proteínas com domínio HD-GYP e sete proteínas com domínio PilZ.

Tabela 2 - Compilação de proteínas alvo. Algumas informações não foram encontradas para todas as proteínas, por essa razão alguns dos espaços não foram preenchidos na terceira, na sexta e sétima colunas.

Código PA	PA14	Nomes	# Resíduos	Domínio alvo contido	Funções associadas	Referência
PA0707	PA14_55160	ToxR, RegA	259	EAL	Regulador de transcrição do gene ToxA.	WALKER. ⁷⁷
PA3825	PA14_14530	-	523	EAL	-	-
PA2133	PA14_36990	-	285	EAL	-	-
-	PA14_15435	TnpM	329	EAL	-	-
PA2200	PA14_36260		531	EAL	-	-
PA3947	PA14_12810	RocR	392	EAL	-	-
-	PA14_59790	PvrR	399	EAL	Regulação negativa do <i>cluster</i> de genes CupD	MIKKELSEN. ⁷⁸
PA3311	PA14_21190	-	785	GGDEF e EAL	-	-
PA5442	PA14_71850	-	950	GGDEF e EAL	-	-
PA4601	PA14_60870	MorA	1415	GGDEF e EAL	Regulador global de fenótipos associados à formação de biofilme.	RAVICHANDRAN. ⁷⁹
PA0861	PA14_53140	RbdA	823	GGDEF e EAL	Fosfodiesterase envolvida na dispersão de biofilme e motilidade.	AN. ⁸⁰
PA1727	PA14_42220	MucR	685	GGDEF e EAL	Regulador da síntese de alginato.	HAY. ⁸¹
PA5295	PA14_69900	-	558	GGDEF e EAL	-	-
PA2072	PA14_37690	-	864	GGDEF e EAL	-	-
PA0575	PA14_07500	-	1245	GGDEF e EAL	-	-
PA1433	PA14_45930	-	650	GGDEF e EAL	-	-
PA0285	PA14_03720	-	760	GGDEF e EAL	-	-
PA2567	PA14_31330	-	499	GGDEF e EAL	-	-
PA3258	PA14_21870	-	601	GGDEF e EAL	-	-
PA4959	PA14_65540	FimX	691	GGDEF e EAL	Regulação de motilidade do tipo <i>twitching</i> .	NAVARRO. ³⁴
PA4367	PA14_56790	BifA	687	GGDEF e EAL	Envolvida na regulação negativa de formação de biofilme e positiva da motilidade do tipo <i>swarming</i> .	KUCHMA. ⁸²
PA5017	PA14_66320	DipA	899	GGDEF e EAL	Fosfodiesterase envolvida na dispersão de biofilme induzida por nutriente.	ROY. ⁸³
PA1181	PA14_49160	-	1120	GGDEF e EAL	-	-
PA5487	PA14_72420	-	671	GGDEF	-	-

(continua)

(continuação)

Tabela 2 - Compilação de proteínas alvo. Algumas informações não foram encontradas para todas as proteínas, por essa razão alguns dos espaços não foram preenchidos na terceira, na sexta e sétima colunas.

Código PA	PA14	Nomes	# Resíduos	Domínio alvo contido	Funções associadas	Referência
PA1120	PA14_49890	TpbB, YfiN	435	GGDEF	Envolvida na regulação positiva da síntese do exopolissacarídeo Pel e expressão de genes de adesão e negativa da motilidade <i>swarming</i> .	UEDA. ⁸⁴ GIARDINA. ⁸⁵
PA0169	PA14_02110	SiaD	235	GGDEF	Envolvida na indução de agregação celular na presença de agente tóxico.	KLEBENSBERGER. ⁸⁶
PA0847	PA14_53310	-	681	GGDEF	-	-
PA4396	PA14_57140	-	366	GGDEF	-	-
PA3177	PA14_23130	-	307	GGDEF	-	-
PA0290	PA14_03790	-	323	GGDEF	-	-
PA3702	PA14_16500	WspR	347	GGDEF	Envolvida na regulação positiva da formação de biofilme.	DE. ⁸⁷
PA2870	PA14_26970	-	525	GGDEF	-	-
PA4929	PA14_65090	NicD	673	GGDEF	Ciclase induzida por nutriente.	ROY. ⁶⁷
PA1851	PA14_40570	-	401	GGDEF	-	-
PA0338	PA14_04420	-	376	GGDEF	-	-
PA4332	PA14_56280	SadC	375	GGDEF	Envolvida na regulação negativa da motilidade do tipo <i>swarming</i> e positiva da formação de biofilme.	MERRIT. ⁸⁸
PA1107	PA14_50060	RoeA	398	GGDEF	Envolvida na regulação da síntese do polissacarídeo Pel e negativa de motilidade <i>swarming</i> .	MERRIT. ⁸⁹
PA4843	PA14_64050	GcbA	542	GGDEF	Envolvida na modulação da motilidade, atuando somente nos estágios iniciais de adesão a superfícies.	PETROVA. ⁹⁰
PA3343	PA14_20820	-	389	GGDEF	-	-
PA2960	PA14_25770	PilZ	118	PilZ	Envolvida na síntese de pili do tipo IV e regulação da motilidade do tipo <i>twitching</i> .	ALM. ⁹¹
PA4608	PA14_60970	-	125	PilZ	-	-
PA2989	PA14_25420	-	254	PilZ	-	-
PA0012	PA14_00130	-	88	PilZ	-	-
PA3542	PA14_18550	Alg44	389	PilZ	Envolvida na síntese de alginato.	MERIGUI. ⁹²
PA2799	PA14_27930	-	99	PilZ	-	-
PA3353	PA14_20700	-	263	PilZ	-	-
PA4108	PA14_10820	-	414	HD-GYP	-	-
PA2572	PA14_30830	-	447	HD-GYP	-	-
PA4781	PA14_63210	-	393	HD-GYP	-	-

Fonte: Elaborada pela autora.

Em seguida, foram removidas as porções transmembranares de cada sequência – ponderando a existência de hélices possivelmente preditas pela ferramenta TMHMM, mas contidas em domínios não transmembranares – e determinadas as porções solúveis para o desenho de *primers*. Na Figura 15, são apontados dois exemplos (de PA3825 e de PA1727) de resultados utilizados no processamento feito para a obtenção de sequências correspondentes a

porções solúveis – as imagens ilustram os resultados obtidos pela ferramenta TMHMM e a predição de domínios pela ferramenta do banco de dados Pfam.

Figura 15 - Resultados da predição de hélices transmembranares obtidos pela ferramenta *online* TMHMM – em (a) e (c) – e da predição de domínios obtidos pela ferramenta do banco de dados Pfam – em (b) e (d). Em (a) e (b) são apresentados os resultados para a proteína PA3825 e, em (c) e (d), para a proteína PA1727.

Fonte: Adaptada de TMHMM...⁹³⁻⁹⁴; PFAM...⁹⁵

No caso da proteína PA3825, optou-se por fazer três construções distintas da sequências, devido à predição de menor probabilidade de uma hélice transmembranar entre os resíduos 235 e 254. Foi excluída também toda a porção anterior ao resíduo 36 (de 1 a 35), devido à predição de alta probabilidade de hélice transmembrana entre os resíduos 13 e 35 – a região anterior (resíduos 1 a 12) também foi excluída, pois havia sido determinado um limiar mínimo de quarenta resíduos para que se mantivesse uma porção solúvel curta dentre as

construções. As construções de PA3825 foram então delimitadas da seguinte forma: a primeira construção estendeu-se do resíduo 36 ao resíduo 234 (identificada como B1 na Figura 16); a segunda, do resíduo 255 ao 523 (identificada como C1 na Figura 16) e a terceira – nesta, a hélice de baixa predição foi desconsiderada –, do resíduo 36 ao 523 (identificada como D1 na Figura 16). No caso da proteína PA1727, toda a região transmembranar inicial foi excluída (do resíduo 1 ao resíduo 237), resultando na construção que se estendeu do resíduo 238 ao resíduo 685 (identificada como B3 na Figura 16).

Desta etapa, resultaram sessenta e seis construções de sequências de *P. aeruginosa*. A Figura 16 ilustra um painel com as construções obtidas para cada uma das proteínas selecionadas. As construções foram identificadas, por praticidade, por letras (de A a H) e números (de 1 a 9), conforme indicado no painel, ao lado da identificação das proteínas das quais foram originadas. Considerando que cada construção é clonada em ambos os vetores pBT e pTRG para possibilitar o teste de todas as interações possíveis, juntas, as bibliotecas completas deveriam contabilizar cento e trinta e dois clones de *P. aeruginosa*.

Figura 16 - Arquitetura de domínios das construções obtidas após o processamento das porções transmembranares. À esquerda é indicada a proteína de origem das construções, as quais foram identificadas especificamente por letras (de A a H) e números (de 1 a 9).

Fonte: Elaborada pela autora.

4.2 Construção das bibliotecas de DNA de “iscas” e “presas” por meio da técnica de dupla-restrição enzimática

Uma vez obtidas as construções de sequências alvo, foram projetados *primers* adequados para a sua amplificação e posterior clonagem nos vetores pBT e pTRG, conforme descrito na subseção 3.2.2.1. Devido à sua extensão, a tabela em que são listados todos os *primers* desenhados está apresentada no Apêndice A (Tabela A1).

Utilizando o método de clonagem por dupla-restrição enzimática (assim designado para diferenciá-lo do método de clonagem por LIC), embora aproximadamente 81 % de todas as construções tenham sido amplificadas com sucesso – contabilizando insertos para clonagem tanto em pBT quanto em pTRG – e clivadas com os pares adequados de enzimas, não mais que 35 % e 39 % de cobertura das bibliotecas de “iscas” e “presas” (clonadas e validadas) foram obtidos em um prolongado espaço de tempo. A Figura 17 ilustra alguns exemplos de análises de restrição para primeira confirmação de algumas das clonagens feitas nesta etapa.

Figura 17 - Imagens resultantes de eletroforese em gel de agarose para a análise de restrição de alguns dos clones de (a) “iscas” e (b) “presas” obtidos pela técnica de dupla-restrição enzimática. As bandas mais altas correspondem ao vetor (pBT ou pTRG) e as mais baixas, ao inserto clivado. Como sempre foram testados vetores extraídos de, ao menos, duas colônias por tentativa de clonagem, os índices (1) e (2) identificam a colônia de origem. Os tamanhos esperados para os insertos eram: 780 pb (PA0707_A1), 990 pb ((tnpM)_F1), 777 pb (PA1120_D5), 1419 pb (PA3311_D2), 2100 pb (PA0285_B4), 606 pb (PA1107_D7), 240 pb (PA3311_C2).

Fonte: Elaborada pela autora.

De fato, em vista do tempo transcorrido desde que haviam sido iniciados os experimentos, muito maior do que o previsto originalmente, verificou-se a inadequação desta estratégia tradicional para a clonagem em maior escala. A clonagem por dupla-restrição enzimática requer que o planejamento e os ajustes experimentais sejam feitos, praticamente, de forma específica para cada uma das sequências, além de envolver a execução de um número razoável de etapas. Neste caso, somou-se a isso, a necessidade de que todas as

sequências fossem clonadas em dois tipos de vetores, cujos MCSs não são diretamente compatíveis (Figura 11, subseção 3.1.1). Foi necessário, então, buscar novos meios para a conclusão desta etapa e foi por esta razão que se firmou, no final do ano de 2013, a já referida colaboração com o pesquisador Dr. César Moisés Camilo e foi iniciada a etapa de clonagens pela técnica de LIC (*Ligase Independent Cloning*).

4.3 Construção das bibliotecas de DNA de “iscas” e “presas” por meio da técnica de *Ligase Independent Cloning* (LIC)

Conforme já indicado na metodologia (seção 3), uma vez feitas as modificações nos vetores para que contivessem extensões LIC, bastou a inclusão de extensões LIC complementares a essas nos *primers* para amplificação de insertos para que pudessem ser iniciadas as etapas de clonagem. Neste ponto, destacam-se duas vantagens da técnica para a construção das bibliotecas almejadas: a modificação dos vetores tornou os sítios de clonagem de pBT e pTRG compatíveis, eliminando a necessidade de que um mesmo inserto fosse amplificado separadamente para clonagem em cada um dos vetores; ao mesmo tempo, como a técnica não envolve clivagem por enzimas de restrição, o desenho dos *primers* dos insertos para LIC foi simplificado, já que, em todos os casos, puderam ser utilizadas as mesmas extensões LIC (uma para o *primer* direto e outra para o *primer* reverso). A Tabela A2, no Apêndice A, traz a lista de todos os *primers* desenhados para a amplificação dos insertos para LIC. Após a amplificação das sequências, os passos subsequentes para clonagem, descritos na subseção 3.2.3.2, eram diretamente aplicáveis a todas as amostras, já que a técnica baseia-se, quase que exclusivamente, na complementaridade gerada pela criação dos *overhangs*, idênticos para todas elas, o que constitui a principal vantagem da técnica e a razão pela qual foi escolhida em detrimento da continuação do uso da técnica até então empregada.

Embora a utilização desta técnica tenha sido, sem dúvida, muito prolífica e oportuna, não foi possível, em tempo hábil, obter total cobertura das bibliotecas (optou-se pelo início dos ensaios de triagem por grupos). Ao final, foram obtidos e validados, por sequenciamento, aproximadamente 70% dos vetores “isca” e 67% dos vetores “presa” almejados. As sequências cujas tentativas de clonagem nos vetores pBT ou pTRG não foram bem-sucedidas estão indicadas na Tabela 3.

Tabela 3 - Lista de clones (“iscas” e “presas”) não obtidos. Cada “isca” ou “presa” foi designada pela vetor em que seria clonada, pela proteína de origem da construção e pelo código que designa a construção.

“Iscas”	“Presas”
pBT-PA5442_E2	pTRG-PA3825_C1
pBT- PA460_G2	pTRG-PA2200_H1
pBT- PA0861_A3	pTRG-PA4601_G2
pBT- PA5295_C3	pTRG-PA2072_E3
pBT- PA2072_E3	pTRG-PA0575_F3
pBT- PA0575_F3	pTRG-PA1433_H3
pBT- PA1433_A4	pTRG-PA1433_A4
pBT- PA2567_C4	pTRG-PA2567_C4
pBT- PA3258_D4	pTRG-PA3258_D4
pBT- PA0847_F5	pTRG-PA1120_C5
pBT- PA0847_G5	pTRG-PA1120_D5
pBT- PA4396_H5	pTRG-PA3177_A6
pBT- PA3177_A6	pTRG-PA0290_B6
pBT- PA0290_B6	pTRG-PA3702_C6
pBT- PA2870_E6	pTRG-PA2870_D6
pBT- PA1851_A7	pTRG-PA2870_E6
pBT- PA4332_C7	pTRG-PA1851_H6
pBT- PA4843_E7	pTRG-PA1851_A7
pBT- PA2960_H7	pTRG-PA4332_C7
pBT- PA4781_B9	pTRG-PA4843_E7
	pTRG-PA4608_A8
	pTRG-PA4108_H8

Fonte: Elaborada pela autora.

4.4 Ensaios de duplo-híbrido bacteriano

4.4.1 Padronização dos ensaios de interação

Os padrões obtidos ao final dos testes feitos nesta etapa e, por conseguinte, resultantes da aplicação do protocolo estabelecido para a realização dos testes de interação (exposto na subseção 3.2.4.4) estão ilustrados nas figuras a seguir. A Figura 18 mostra a padronização para os ensaios do tipo um a um, em que uma única “isca” e uma única “presa” são cotransformadas; já a Figura 19, a padronização para os ensaios de triagem por grupos. Neste último caso, para o controle positivo, os testes foram feitos cotransformando a linhagem repórter com 80 ng de pBT-LGF2, 400 ng de pBT-PA3825_B1, 80 ng de pTRG-Gal11^P e 400 ng de pTRG-PA3343_G7 para simular a existência de um único par “isca” e “presa” interagente em um experimento em que “iscas” e “presas” não interagentes estivessem presentes cinco vezes em excesso no processo de cotransformação – ou seja, para simular a

cotransformação com grupos de seis componentes de “iscas” e seis componentes de “presas” e averiguar a possibilidade de que, existindo somente um par interagente, fosse observado um padrão positivo bem definido. A escolha dos vetores pBT-PA3825_B1 e pTRG-PA3343_G7 como par não interagente para estes testes, foi previamente certificada por ensaio de duplo-híbrido entre eles (do tipo um a um) em que se observou padrão negativo para interação, idêntico ao observado no controle negativo (não houve crescimento em placa seletiva).

Figura 18 - Imagens digitalizadas de placas mostrando o resultado final da etapa de padronização dos ensaios de duplo-híbrido do tipo um a um. Em (a), é mostrado o controle negativo; em (b) e em (c), os controles de eficiência de cotransformação dos controles negativo e positivo, respectivamente, e, em (d), o controle positivo.

Fonte: Elaborada pela autora.

Figura 19 - Imagens digitalizadas de placas mostrando o resultado final da etapa de padronização dos ensaios para a estratégia de triagem por grupos. Em (a), é mostrado o controle negativo; em (b) e em (c), os controles de eficiência de cotransformação do controle negativo e da simulação de interação positiva, respectivamente, e, em (d), a simulação de interação positiva em um ensaio entre grupos.

Fonte: Elaborada pela autora.

A comparação entre as imagens exibidas na Figura 18 e na Figura 19 revela que, embora a eficiência de cotransformação não tenha se alterado, o padrão positivo para a cotransformação de grupos foi perceptivelmente menos forte, o que é coerente com o fato de que uma fração necessariamente menor de células foi efetivamente cotransformada com o par interagente nesse caso. Não obstante, ficou comprovada a viabilidade da estratégia de triagem por grupos.

4.4.2 Ensaios de interação: triagem por grupos

A primeira etapa da execução da estratégia de triagem por grupos foi a divisão das bibliotecas de “iscas” e de “presas” em conjuntos formados segundo os critérios expostos na metodologia. Da adoção desses critérios, resultaram oito grupos de “iscas” e oito grupos de “presas”, os quais estão resumidos, respectivamente, na Tabela 4 e na Tabela 5. A Tabela A3 e a Tabela A4, encontradas no Apêndice A trazem mais informações. É importante observar que, embora similar, a composição dos grupos de “iscas” e “presas” indicados pelo mesmo índice (de 1 a 8) não é correspondente entre si, o que se deve ao fato de que algumas das construções foram clonadas com sucesso somente em um dos dois tipos de vetor.

Definida a composição de grupos, foram executados sessenta e quatro ensaios de interação – cada um dos oito grupos de “iscas” foi testado contra cada um dos oito grupos de “presas” –, todos acompanhados dos controles indicados na subseção 3.2.4.3. Os resultados estão assinalados na matriz mostrada na Figura 20. Os resultados classificados como “padrão de interação forte” e “padrão de interação fraca” estão representados, de forma generalizada, pelas imagens de placas de ensaios de duplo-híbrido efetivos mostradas também na Figura 20. A designação “ausência de interação” foi utilizada para indicar testes que geraram placas em que não se observou crescimento algum. A correlação entre o crescimento obtido nas placas de teste (analisado sempre em conjunto com o respectivo controle de eficiência de cotransformação) e a força de interação está indicada no manual do *kit* adquirido para os ensaios. O gráfico da Figura 21 resume o resultado obtido em termos das frações do total de testes realizados. É, todavia, válido ressaltar que esse gráfico foi elaborado para a extração e melhor visualização dos dados contidos na Figura 20, mas que as frações indicadas não representam uma tendência final, já que em cada teste entre grupos havia entre vinte e trinta e seis possibilidades implícitas de interação.

Apesar de preliminar, este resultado foi valioso, pois indicou não somente a existência de interações, determinando e orientando a continuidade dos testes, mas também o sucesso da estratégia empregada, já que permitiu uma drástica redução na quantidade de testes que deveriam ser executados para o esgotamento das possibilidades de interação. Consideradas as coberturas das bibliotecas de “iscas” e de “presas”, o número de possíveis interações a serem testadas foi reduzido de duas mil e vinte e quatro para setecentas e dezesseis – ou seja, foram descartadas, pela triagem por grupos, mil trezentas e oito interações possíveis e, conseqüentemente, o número de testes que deveriam ser realizados foi reduzido na mesma proporção.

Tabela 4 - Composição dos oito grupos de “iscas” formados. A denominação de cada componente identifica qual o vetor em que a construção foi clonada (no caso, pBT), a proteína de origem e o código da construção.

Grupo	Componentes
I1	pBT-PA0707_A1 ; pBT-PA2133_E1 ; pBT- (tnpM) _F1 ; pBT-PA3947_A2 ; pBT- (pvrR) _B2
I2	pBT-PA3825_B1 ; pBT-PA3825_C1 ; pBT-PA3825_D1 ; pBT-PA2200_G1 ; pBT-PA2200_H1 ; pBT-PA3353_G8
I3	pBT-PA3311_C2 ; pBT-PA3311_D2 ; pBT-PA1727_B3 ; pBT-PA4601_F2 ; pBT-PA4367_F4 ; pBT-PA4367_G4
I4	pBT-PA0575_G3 ; pBT-PA0285_B4 ; pBT-PA4959_E4 ; pBT-PA5017_H4 ; pBT-PA1181_A5 ; pBT-PA2072_D3
I5	pBT-PA5487_B5 ; pBT-PA0169_E5 ; pBT-PA1107_D7 ; pBT-PA3343_F7 ; pBT-PA3343_G7
I6	pBT-PA1120_C5 ; pBT-PA1120_D5 ; pBT-PA3702_C6 ; pBT-PA4929_F6 ; pBT-PA4929_G6 ; pBT-PA0338_B7
I7	pBT-PA4608_A8 ; pBT-PA2989_B8 ; pBT-PA0012_C8 ; pBT-PA3542_D8 ; pBT-PA3542_E8 ; pBT-PA2799_F8
I8	pBT-PA0861_H2 ; pBT-PA1433_H3 ; pBT-PA2870_D6 ; pBT-PA1851_H6 ; pBT-PA4108_H8 ; pBT-PA2572_A9

Fonte: Elaborada pela autora.

Tabela 5 - Composição dos oito grupos de “presas” formados. A denominação de cada componente identifica qual o vetor em que a construção foi clonada (no caso, pTRG), a proteína de origem e o código da construção.

Grupo	Componentes
P1	pTRG-PA0707_A1 ; pTRG-PA2133_E1 ; pTRG- (tnpM) _F1 ; pTRG-PA3825_B1 ; pTRG-PA3825_D1 ; pTRG-PA2200_G1
P2	pTRG-PA3311_C2 ; pTRG-PA3311_D2 ; pTRG-PA1727_B3 ; pTRG-PA5295_C3 ; pTRG-PA4367_F4 ; pTRG-PA4367_G4
P3	pTRG-PA5442_E2 ; pTRG-PA0861_H2 ; pTRG-PA0861_A3 ; pTRG-PA0575_G3 ; pTRG-PA0285_B4 ; pTRG-PA4959_E4
P4	pTRG-PA4601_F2 ; pTRG-PA2072_D3 ; pTRG-PA5017_H4 ; pTRG-PA1181_A5
P5	pTRG-PA5487_B5 ; pTRG-PA0169_E5 ; pTRG-PA1107_D7 ; pTRG-PA3343_F7 ; pTRG-PA3343_G7
P6	pTRG-PA0847_F5 ; pTRG-PA0847_G5 ; pTRG-PA4396_H5 ; pTRG-PA4929_F6 ; pTRG-PA4929_G6 ; pTRG-PA0338_B7
P7	pTRG-PA2960_H7 ; pTRG-PA2989_B8 ; pTRG-PA0012_C8 ; pTRG-PA3542_D8 ; pTRG-PA3542_E8
P8	pTRG-PA3947_A2 ; pTRG- (pvrR) _B2 ; pTRG-PA2572_A9 ; pTRG-PA4781_B9 ; pTRG-PA2799_F8 ; pTRG-PA3353_G8

Fonte: Elaborada pela autora.

Figura 20 - Matriz ilustrativa dos resultados da estratégia de triagem por grupos. À direita, está indicada a legenda de cores. Os padrões de interação forte e fraca estão representados pelas imagens de digitalizadas de placas exibidas abaixo da matriz.

Fonte: Elaborada pela autora.

Figura 21 - Gráfico apontando as parcelas de ensaios de interação da estratégia de triagem por grupos que resultaram em padrões de interação forte, padrão de interação fraca e em ausência de interação.

Fonte: Elaborada pela autora.

4.4.3 Ensaios de interação: triagem por pares

Nesta etapa, os pares de grupos que produziram resultados positivos para interação foram desmembrados e foram feitos ensaios de duplo-híbrido em que cada um dos componentes pertencentes ao grupo “isca” envolvido em uma determinada interação positiva

(revelada na etapa anterior) foram testados contra cada um dos componentes pertencentes ao grupo “presa” envolvido nessa mesma interação.

Embora reduzido pela estratégia de triagem por grupos, o número de testes do tipo um a um que deveria ser executado nesta etapa era ainda vultoso, o que motivou a determinação de uma ordem de prioridade para a apuração das interações encontradas. Previsivelmente, os grupos que geraram resultados classificados como “padrão de interação forte” foram examinados prioritariamente. O resultado da triagem por pares de interação é apresentado na Figura 22 e foi o primeiro mapeamento de interações obtido. Nele, as cores indicam qual é tipo de domínio alvo (GGDEF, EAL, HD-GYP e PilZ) contido na proteína que originou aquela construção e a espessura das linhas que conectam os parceiros de interação indicam o padrão de força observado na placa resultante do teste que os revelou. Neste caso, foi necessário reconfigurar a definição dos padrões de força, em virtude do maior crescimento de colônias observado nas placas de ensaios do tipo um a um (veja seção 4.4.1) – optou-se então por classificar as interações encontradas em três padrões de força (fraca, média e forte).

Figura 22 - Primeira rede de interações resultante da etapa de triagem por pares. Optou-se pela utilização dos nomes das proteínas (quando já existente). Linhas tracejadas indicam interações fracas, linhas simples indicam interações de força média e linhas de maior espessura indicam interações fortes. Módulos de cor vermelha indicam proteínas que contêm domínio GGDEF; de cor vermelha e amarela, ambos os domínios GGDEF e EAL; de cor amarela, domínio EAL e de cor azul, domínio HD-GYP.

Fonte: Elaborada pela autora.

O mapa da Figura 22, embora esteja fundamentado em experimentos iniciais e exploratórios – tal qual é o caráter e a intenção deste estudo – evidencia a possibilidade de que exista, de fato, uma rede diversificada de interações proteicas nas vias de sinalização estudadas, em acordo com as expectativas iniciais. Surpreendentemente, embora interações envolvendo proteínas com domínio PilZ já tenham sido descritas em outros organismos, estas

não foram encontradas dentre os parceiros de interação deste primeiro conjunto de testes em busca de pares específicos.^{64,66} Por outro lado, fica clara a predominância de proteínas contendo domínios GGDEF ou ambos os domínios GGDEF e EAL na rede, o que não era inesperado devido à maior distribuição dessas proteínas no genoma de *P. aeruginosa* em relação à distribuição de proteínas contendo os domínios EAL, HD-GYP e PilZ.

Dentre os parceiros de interação encontrados, destacam-se, pela força de interação, SiaD (PA0169) e PA3825, PA1851 e PA0847, PA3343 e PA0847, PA0575 e PA0285, além das dimerizações de FimX (PA4959) e MucR (PA1727). Algumas destas interações foram submetidas a novos ensaios de interação para a obtenção de informações mais específicas, conforme explicitado na subseção seguinte.

4.4.4 Ensaios de interação complementares

Para avaliar mais extensivamente algumas das interações reveladas, foram executados ensaios de duplo-híbrido adicionais com alguns dos parceiros de interação. No caso de FimX contra FimX (PA4959), por exemplo, para averiguar se o resultado positivo era fruto da já relatada interação dimérica que ocorre por meio dos domínios N-terminais de FimX ou se estava relacionado a uma cogitada interface de dimerização no domínio EAL, foram feitas duas novas construções dessa proteína: uma com sequência truncada na região N-terminal, eliminando toda a porção proteica envolvida na interação já conhecida, e outra com sequência truncada na extremidade C-terminal, eliminando o domínio EAL. A relevância deste resultado decorre da observação feita por Navarro e colaboradores da existência de uma extensa interface de dimerização no domínio EAL da proteína, a qual não poderia contribuir para a interação dimérica de acordo com o que se evidenciou para FimX em solução, mas, especulou-se, poderia ser, neste contexto, a base para a formação de oligômeros de ordem superior em locais de alta concentração proteica na célula.³⁴

Os ensaios de duplo-híbrido com as novas construções clonadas nos vetores pBT e pTRG indicaram que a interação observada era, de fato, ocasionada pela interface de dimerização já anteriormente reportada e não envolvia o domínio EAL C-terminal. Foi observado crescimento, com padrão de interação forte, quando foram testadas as construções com a porção C-terminal truncada (ou seja, sem domínio EAL); em contrapartida, quando foram testadas as construções com N-terminal truncado não houve crescimento algum na placa seletiva (Figura 23).

Figura 23 - Imagens digitalizadas de placas mostrando o resultado final dos ensaios de interação complementares com a proteína FimX. Em (a), é mostrado o teste feito com a porção N-terminal truncada e, em (b), com o domínio EAL truncado. Acima das imagens, são apresentadas esquematizações das construções correspondentes.

Fonte: Elaborada pela autora.

Uma das proposições de Navarro e colaboradores indica a possibilidade de que o domínio EAL degenerado de FimX interaja com domínios EAL não degenerados de outras proteínas, o que impediria sua atividade catalítica. Segundo os autores, isso ocorreria na ausência de c-di-GMP, em concordância com a constatação por eles feita de que a dimerização do domínio EAL não seria compatível com a ligação a essa molécula. Na presença de c-di-GMP, então, a proteína com domínio EAL não degenerado seria liberada da interação e sua atividade catalítica seria restaurada. A verificação de que, em solução, o domínio EAL de uma proteína FimX não forma dímero com o domínio EAL de outra proteína FimX possivelmente corrobora esta hipótese de regulação da atividade de proteínas com domínio EAL não degenerado.

Uma interação tão interessante quanto surpreendente foi a de PA0847_G5 com PA3343_F7 – posteriormente, revelou-se também a interação de PA0847 com PA1851, de arquitetura muito similar a PA3343. Conforme se observa na Figura 16, PA3343_F7 tem uma sequência peptídica de apenas cinquenta e um resíduos e separa-se do restante da porção solúvel de PA3343 por uma extensa região transmembranar – ambas as porções solúveis são citoplasmáticas, o que foi confirmado pelo resultado de predição de hélices transmembranares obtido pelo método MEMSAT. PA0847_G5 é a porção citoplasmática de uma proteína também transmembranar e contém, além de um domínio GGDEF C-terminal, um domínio HAMP e um domínio PAS. Esta proteína é também alvo dos estudos para obtenção de título de doutorado do aluno Éverton Silva (também sob orientação do Prof. Dr. Marcos Navarro).

Com o intuito de investigá-la de forma mais específica, foram planejados novos ensaios de interação, em colaboração com o aluno Éverton Silva – agora com construções de PA0847_G5 contendo apenas alguns dos domínios. Para tanto, foram desenhados *primers*

para a amplificação de construções contendo, alternativamente, somente as porções correspondentes aos domínios HAMP e PAS (PA0847_HP), somente ao domínio PAS (PA0847_P), somente PAS e GGDEF (PA0847_PG) e somente GGDEF (PA0847_G). Cada uma delas foi clonada com sucesso no vetor pTRG e foram feitos então os ensaios para averiguar a interação com PA3343_F7. Os resultados são mostrados na Figura 24.

Figura 24 - Imagens digitalizadas de placas mostrando o resultado final dos ensaios de interação complementares com PA0847_G5 e PA3343_F7. Em (a), é mostrado o teste feito com PA0847_HP; em (b), com PA0847_P; em (c), com PA0847_PG e, em (d), com PA0847_G. Acima das imagens, são apresentadas esquematizações das construções correspondentes.

Fonte: Elaborada pela autora.

Os testes revelaram que a interação independe do domínio GGDEF e não ocorre na ausência do domínio HAMP, já que o padrão forte de crescimento só foi recuperado na placa correspondente ao teste feito com a construção PA0847_HP e não foi observado crescimento algum nas demais placas. Considerando a curta extensão da sequência peptídica de PA3343_F7, é coerente que a interação ocorra com a porção de PA0847 mais próxima à membrana. Adicionalmente, o aluno Éverton Silva executou experimentos de *pull-down* com PA3343_F7 e PA0847_G5, os quais indicaram a ocorrência de interação quando as proteínas foram pré-incubadas com c-di-GMP ou GTP (substrato para a síntese de c-di-GMP), o que indica que o c-di-GMP seja um intermediário para a interação.

4.4.5 Ensaios de força de interação e testes de autoativação

Como uma tentativa de acessar de forma mais direta a força das interações reveladas e como uma segunda confirmação, foram feitos ensaios utilizando o meio duplamente seletivo, conforme descrito na subseção 3.2.4.6. Conjuntamente, foram feitos também ensaios de autoativação de “iscas” e de “presas”. Os resultados estão apresentados na Figura 25.

Figura 25 - Ensaios de força de interação e testes de autoativação, feitos em meio duplamente seletivo. À esquerda de cada resultado, estão indicados os parceiros de interação testados. Abaixo das imagens dos testes de força estão indicadas as diluições correspondentes a cada coluna. À direita encontram-se os testes de autoativação de “iscas” e “presas” envolvidas em cada teste. Espaços em branco indicam testes que foram descartados devido à ausência de crescimento no controle de cotransformação correspondente (não mostrado). Os resultados estão agrupados de acordo com a data em que foram feitos – os controles correspondentes a cada uma das rodadas de testes são mostrados nas duas primeiras linhas.

Fonte: Elaborada pela autora.

De forma geral, os resultados de força de interação em placa duplamente seletiva correlacionam-se bem com as avaliações que haviam sido feitas previamente com base nas placas dos ensaios em meio seletivo. Observou-se um crescimento mais consistente nos casos de interações já classificadas como fortes, o qual, em geral, não se evanesce com o aumento da diluição da cultura, enquanto que, nos casos de interações classificadas como médias, o crescimento parece persistir até a segunda ou terceira diluições e, nos casos de interação fraca, até a primeira ou segunda diluições. Algumas interações, no entanto, não foram confirmadas, como é o caso de PA0338_B7 (“isca”) com PA2133_E1 (“presa”) e PA1120_D5 e TnpM_F1.

Os resultados dos testes de autoativação – outro parâmetro importante para a avaliação das interações encontradas – lançam dúvida sobre a validade de algumas das interações mais fracas encontradas, já que, nesses casos, os crescimentos foram comparáveis aos dos de força de interação, mesmo sem diluição da cultura. É preciso observar, no entanto, que, para a execução dos ensaios de autoativação de “iscas”, por razões secundárias, foi usado o vetor pTRG-Gal11^P (e não pTRG vazio) e é possível que a presença de Gal11^P tenha influenciado os resultados obtidos nesses testes. Nota-se, inclusive, a ocorrência de crescimento muito maior e mais frequente nos ensaios de autoativação de “iscas” comparativamente aos ensaios de autoativação de “presas” (mesmo quando “iscas” e “presas” são originadas da mesma construção). Além disso, como uma parcela desses testes foi feita mais de uma vez para uma mesma “isca” ou “presa”, pode-se observar que, entre algumas dessas repetições, há discrepâncias entre os resultados obtidos paralelamente (é o caso da “presa” PA0285_B4). Por essa razão, uma repetição futura destes ensaios (após algumas adequações no protocolo utilizado), seria conveniente para melhor orientar o julgamento destes resultados.

4.4.6 Ensaios finais e mapa de interações proteicas

Para dar continuidade aos ensaios do tipo um a um planejados para a apuração das interações positivas restantes encontradas na etapa de triagem por grupos, a autora contou com a colaboração do também aluno de mestrado Raphael Meneghello, cuja assistência foi indispensável para a finalização desta etapa em tempo hábil. Estes testes finais revelaram um novo conjunto de pares interagentes (algumas das proteínas já haviam sido reveladas como participantes de outras interações descobertas nos testes anteriores), os quais foram dispostos, em um novo mapa de interações proteicas (Figura 26), ilustrando a totalidade de interações encontradas. Embora, dentre os novos pares interagentes evidenciados, a vasta maioria tenha apresentado padrão fraco de interação, estas novas ramificações da rede ampliam ainda mais o

espaço de possíveis interações entre as proteínas alvo deste estudo e indicam que este tipo de cenário pode ser predominante e relevante no sistema de sinalização do qual participam. É interessante destacar que algumas proteínas, como a RegA e a PA2133 (ambas constituídas de apenas um domínio EAL), emergiram como participantes de um número surpreendentemente alto de interações, o que poderia indicar, por exemplo, sua atuação como reguladores com atuação global nesse sistema.

Figura 26 - Rede de interações final encontrada. São apresentadas as novas interações evidenciadas durante a finalização da averiguação das interações positivas da etapa de triagem por grupos, conjuntamente com as interações já anteriormente encontradas. Linhas tracejadas indicam interações fracas, linhas simples indicam interações de força média e linhas de maior espessura indicam interações fortes. Módulos de cor vermelha indicam proteínas que contêm domínio GGDEF; de cor vermelha e amarela, ambos os domínios GGDEF e EAL; de cor amarela, domínio EAL; de cor azul, domínio HD-GYP e de cor roxa, domínio PilZ.

Fonte: Elaborada pela autora.*

Curiosamente, a interação direta, relatada por Roy e Sauer, envolvendo duas proteínas que estavam inclusas nos ensaios realizados neste estudo – NicD, uma DGC, e DipA, uma PDE – não foi evidenciada por estes testes.⁶⁷ Apesar disso, NicD aparece como parceira de interação fraca de PA4108 e uma proteína relacionada, RbdA, como parceira de interação de outras três proteínas, PA0285, PA1181 e PA0575. Como esses autores identificaram essa interação por coimunoprecipitação, especula-se que fatores fisiológicos adicionais

* Coautoria de Raphael Meneghello.

influenciem a ocorrência dessa interação – de fato, de acordo com o estudo publicado, essas proteínas estariam implicadas em um enlace maior de interações e processos que subsequentemente levariam à dispersão de biofilme – e que, por essa razão, o mesmo resultado não tenha sido reproduzido nestes testes de duplo-híbrido bacteriano.

5 CONCLUSÕES

Este estudo dividiu-se, basicamente, em duas etapas principais: a obtenção de duas bibliotecas de DNA direcionadas, construídas a partir das sequências codificantes de proteínas contendo domínios tipicamente envolvidos nas vias de sinalização por c-di-GMP de *Pseudomonas aeruginosa*, e a utilização destas bibliotecas – de “iscas” e de “presas” – em ensaios de duplo-híbrido bacteriano para a determinação da rede de interações entre essas proteínas.

Durante a execução da primeira etapa, optou-se pela troca do método utilizado para a construção das bibliotecas – notadamente, passou-se a utilizar a técnica de *Ligase Independent Cloning*, no lugar da técnica de clonagem por meio de dupla-restrição enzimática. A adoção deste outro método, mais adequado para clonagens em maior escala, foi fundamental para a obtenção de uma cobertura maior das bibliotecas de “iscas” e de “presas” em menor espaço de tempo para que pudessem ser executados os experimentos subsequentes, ainda que não tenham sido conseguidos todos os clones previstos.

A segunda etapa mostrou-se tão desafiadora em termos de volume de experimentos quanto a primeira, o que, da mesma forma, estimulou a busca por estratégias que pudessem tornar mais eficiente a investigação por parceiros de interação proteica entre os alvos deste estudo. Projetou-se então a estratégia de triagem por grupos para que fossem acessados mais rapidamente os pares interagentes, por meio da eliminação conjunta de um número maior de possibilidades de interação a cada ensaio de duplo-híbrido. Esta estratégia revelou-se eficiente e útil para esse propósito, tornando mais ágil o processo de exploração das possibilidades de interação.

Para o esgotamento destas possibilidades e identificação específica de potenciais interações entre as proteínas alvo, foram feitos os ensaios de duplo-híbrido de forma individual entre os componentes dos grupos de “iscas” e “presas” que haviam produzido padrões positivos de interação nos testes anteriormente executados. Estes experimentos revelaram uma ampla e diversificada rede de interações proteicas, tanto em relação ao número e distinção entre forças de interação, quanto em relação à variedade de proteínas envolvidas.

Há muito o que se discutir e investigar a respeito das interações encontradas; há um grande espaço em aberto para a sua análise e para a revelação dos possíveis mecanismos moleculares nos quais, acredita-se, estariam envolvidas. Logicamente, cada uma dessas interações deverá ser examinada com maior minúcia, por experimentos que possam caracterizá-las e também validá-las adicionalmente, uma vez que foram retiradas do seu real

contexto fisiológico para a execução destes experimentos. Fornecer um ponto de partida para análises futuras mais específicas era, de fato, uma das perspectivas iniciais deste estudo. Não obstante, como a construção das bibliotecas foi feita de forma direcionada à obtenção de produtos solúveis e bem delimitados, acredita-se que a confiança destes experimentos seja maior do que o de experimentos em que são utilizadas bibliotecas de cDNA ou feitas a partir de DNA genômico por exemplo.

No contexto geral deste estudo, considerando-se o seu caráter exploratório, a rede de interações encontrada constitui uma evidência favorável às hipóteses levantadas a respeito dos mecanismos que medeiam a especificidade de sinalização nas vias que regulam a expressão de fenótipos relacionados, em especial, à formação e dispersão de biofilmes bacterianos. O amplo conjunto de interações revelado é um indício de que, ao menos em *P. aeruginosa*, interações entre proteínas tipicamente associadas às vias de sinalização por c-di-GMP poderiam ser um mecanismo comum pelo qual o sinal é transmitido ou segregado na célula bacteriana, permitindo o controle paralelo da expressão de uma grande variedade de fenótipos.

Por fim, duas das interações encontradas, a de PA0847 com PA3343 e também de PA0847 com PA1851, foram especialmente surpreendentes e de maior interesse devido ao fato de que PA0847 já era alvo de outros estudos no mesmo grupo de pesquisas do qual a autora faz parte. A interação entre as curtas porções solúveis de PA3343 e de PA1851 (nas quais não há também domínios preditos) e a extensa porção solúvel de PA0847 revela a possibilidade de que haja um mecanismo de controle que ocorra de uma forma original e inesperada entre essas proteínas.

6 PERSPECTIVAS

São perspectivas imediatas deste estudo a finalização das bibliotecas de DNA e a realização dos ensaios de duplo-híbrido bacteriano complementares com as novas “iscas” e “presas” obtidas, concluindo assim a investigação das interações entre proteínas tipicamente relacionadas às vias de sinalização por c-di-GMP de *Pseudomonas aeruginosa*. Além disso, seria pertinente a repetição dos ensaios de força de interação (ou de outros ensaios utilizando o agente seletivo secundário, a estreptomicina) e de autoativação para uma avaliação mais minuciosa das interações encontradas. Experimentos adicionais, como *pull-down*, coimunoprecipitação, RT-PCR, fusões com proteínas fluorescentes para localização das proteínas ou ainda experimentos envolvendo a análise da influência de outras moléculas nas interações reveladas neste estudo exploratório inicial seriam interessantes para a sua contextualização e validação e permitiriam que fossem feitas inferências mais específicas a respeito dos mecanismos moleculares nos quais possivelmente estão envolvidas.

Assim como foi feito para *P. aeruginosa*, planeja-se também investigar a rede de interações entre proteínas tipicamente relacionadas às vias de sinalização por c-di-GMP de *Xanthomonas citri* subsp. *citri* (anteriormente, *Xanthomonas axonopodis* pv. *citri*). Os *primers* (tanto para a clonagem pela técnica de dupla-restrição enzimática quanto para a clonagem pela técnica LIC) para a construção de bibliotecas de “iscas” e “presas” foram desenhados e adquiridos conjuntamente com os que foram usados para a construção das bibliotecas deste estudo.

REFERÊNCIAS

- 1 HENGGE, R. Principles of c-di-GMP signalling in bacteria. **Nature Reviews Microbiology**, v. 7, n. 4, p. 263-273, Apr. 2009.
- 2 FLEMMING, H.; WINGENDER, J. The biofilm matrix. **Nature Reviews Microbiology**, v. 8, n. 9, p. 623-633, Sept. 2010.
- 3 KIM, J.; PARK, H.-D.; CHUNG, S. Microfluidic approaches to bacterial biofilm formation. **Molecules**, v. 17, n. 8, p. 9818-9834, Aug. 2012.
- 4 WU, H. et al. Strategies for combating bacterial biofilm infections. **International Journal of Oral Science**, v. 7, n. 1, p. 1-7, Mar. 2015. I
- 5 AN, D.; PARSEK, M. The promise and peril of transcriptional profiling in biofilm communities. **Current Opinion in Microbiology**, v. 10, n. 3, p. 292-296, June 2007.
- 6 KRASTEVA, P. V.; GIGLIO, K. M.; SONDERMANN, H. Sensing the messenger: the diverse ways that bacteria signal through c-di-GMP. **Protein Science**, v. 21, n. 7, p. 929-948, July 2012.
- 7 HOBLEY, L. et al. Giving structure to the biofilm matrix: an overview of individual strategies and emerging common themes. **Fems Microbiology Reviews**, v. 39, n. 5, p. 649-669, Sept. 2015.
- 8 ROMLING, U.; GALPERIN, M.; 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.
- 9 HOIBY, N. et al. Antibiotic resistance of bacterial biofilms. **International Journal of Antimicrobial Agents**, v. 35, n. 4, p. 322-332, Apr. 2010.
- 10 BELOIN, C. et al. Novel approaches to combat bacterial biofilms. **Current Opinion in Pharmacology**, v. 18, p. 61-68, Oct. 2014. doi: 10.1016/j.coph.2014.09.005.
- 11 NGUYEN, D. et al. Active starvation responses mediate antibiotic tolerance in biofilms and nutrient-limited bacteria. **Science**, v. 334, n. 6058, p. 982-986, Nov. 2011.
- 12 SHATALIN, K. et al. H₂S: a universal defense against antibiotics in bacteria. **Science**, v. 334, n. 6058, p. 986-990, Nov. 2011.
- 13 SINGH, R.; PAUL, D.; JAIN, R. Biofilms: implications in bioremediation. **Trends in Microbiology**, v. 14, n. 9, p. 389-397, Sept. 2006.
- 14 ABSALON, C.; YMELE-LEKI, P.; WATNICK, P. The bacterial biofilm matrix as a platform for protein delivery. **Mbio**, v. 3, n. 4, July 2012. doi: 10.1128/mBio.00127-12.

15 KARATAN, E.; WATNICK, P. Signals, regulatory networks, and materials that build and break bacterial biofilms. **Microbiology and Molecular Biology Reviews**, v. 73, n. 2, p. 310-+, June 2009.

16 BORDELEAU, E. et al. Cyclic Di-GMP riboswitch-regulated type IV Pili contribute to aggregation of *Clostridium difficile*. **Journal of Bacteriology**, v. 197, n. 5, p. 819-832, Mar. 2015.

17 DE, N. et al. Phosphorylation-independent regulation of the diguanylate cyclase WspR. **Plos Biology**, v. 6, n. 3, p. 601-617, Mar. 2008.

18 ORR, M. et al. Oligoribonuclease is the primary degradative enzyme for pGpG in *Pseudomonas aeruginosa* that is required for cyclic-di-GMP turnover. **Proceedings of the National Academy of Sciences of the United States of America**, v. 112, n. 36, p. E5048-E5057, Sept. 2015.

19 COHEN, D. et al. Oligoribonuclease is a central feature of cyclic diguanylate signaling in *Pseudomonas aeruginosa*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 112, n. 36, p. 11359-11364, Sept. 2015.

20 BELLINI, D. et al. Crystal structure of an HD-GYP domain cyclic-di-GMP phosphodiesterase reveals an enzyme with a novel trinuclear catalytic iron centre. **Molecular Microbiology**, v. 91, n. 1, p. 26-38, Jan. 2014.

21 WIGREN, E.; LIANG, Z. X.; RÖMLING, U. Finally! the structural secrets of a HD-GYP phosphodiesterase revealed. **Molecular Microbiology**, v. 91, n. 1, p. 1-5, Jan. 2014.

22 STELITANO, V. et al. C-di-GMP hydrolysis by *Pseudomonas aeruginosa* HD-GYP phosphodiesterases: analysis of the reaction mechanism and novel roles for pGpG. **Plos One**, v. 8, n. 9, Sept. 2013.

23 RÖMLING, U.; GOMELSKY, M.; GALPERIN, M. C-di-GMP: the dawning of a novel bacterial signalling system. **Molecular Microbiology**, v. 57, n. 3, p. 629-639, Aug. 2005.

24 JONAS, K.; MELEFORS, O.; RÖMLING, U. Regulation of c-di-GMP metabolism in biofilms. **Future Microbiology**, v. 4, n. 3, p. 341-358, Apr. 2009.

25 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.

26 RESPONSE regulators encoded in bacterial and archaeal genomes. 2011. Disponível em: <http://www.ncbi.nlm.nih.gov/Complete_Genomes/RRcensus.html#Footnote1>. Acesso em: 24 out. 2015.

27 BRETL, D. J.; DEMETRIADOU, C.; ZAHRT, T. C. Adaptation to environmental stimuli within the host: two-component signal transduction systems of *Mycobacterium tuberculosis*. **Microbiology and Molecular Biology Reviews**, v. 75, n. 4, p. 566-582, Dec. 2011.

- 28 MILLS, E. et al. The bacterial second messenger c-di-GMP: mechanisms of signalling. **Cellular Microbiology**, v. 13, n. 8, p. 1122-1129, Aug. 2011.
- 29 KALIA, D. et al. Nucleotide, c-di-GMP, c-di-AMP, cGMP, cAMP, (p)ppGpp signaling in bacteria and implications in pathogenesis. **Chemical Society Reviews**, v. 42, n. 1, p. 305-341, 2013.
- 30 TSCHOWRI, N. et al. Tetrameric c-di-GMP mediates effective transcription factor dimerization to control *Streptomyces* development. **Cell**, v. 158, n. 5, p. 1136-1147, Aug. 2014.
- 31 ZHANG, Z. et al. Polymorphism of the signaling molecule c-di-GMP. **Journal of the American Chemical Society**, v. 128, n. 21, p. 7015-7024, May 2006.
- 32 NAKAYAMA, S. et al. A C-di-GMP-proflavine-hemin supramolecular complex has peroxidase activity--implication for a simple colorimetric detection. **Molecular BioSystems**, v. 8, n. 3, p. 726-729, Mar. 2012.
- 33 CHOU, S. H.; GALPERIN, M. Y. Diversity of c-di-GMP-binding proteins and mechanisms. **Journal of Bacteriology**, June 2015. doi: 10.1128/JB.00333-15. Disponível em: < <http://jb.asm.org/content/early/2015/06/03/JB.00333-15.full.pdf+html>>. Acesso em 10 nov. 2015.
- 34 NAVARRO, M. 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.
- 35 SONDERMANN, H.; SHIKUMA, N.; YILDIZ, F. You've come a long way: c-di-GMP signaling. **Current Opinion in Microbiology**, v. 15, n. 2, p. 140-146, Apr. 2012.
- 36 AMIKAM, D.; GALPERIN, M. Y. PilZ domain is part of the bacterial c-di-GMP binding protein. **Bioinformatics**, v. 22, n. 1, p. 3-6, Jan. 2006.
- 37 BOYD, C.; O'TOOLE, G.; SCHEKMAN, R. Second messenger regulation of biofilm formation: breakthroughs in understanding c-di-GMP effector systems. **Annual Review of Cell and Developmental Biology**, v. 28, p. 439-462, 2012. doi: 10.1146/annurev-cellbio-101011-155705.
- 38 FRANKLIN, M. et al. Biosynthesis of the *Pseudomonas aeruginosa* extracellular polysaccharides, alginate, Pel, and Psl. **Frontiers in Microbiology**, v. 2, 2011. doi: 10.3389/fmicb.2011.00167.
- 39 PAUL, K. et al. The c-di-GMP binding protein YcgR controls flagellar motor direction and speed to affect chemotaxis by a "backstop brake" mechanism. **Molecular Cell**, v. 38, n. 1, p. 128-139, Apr. 2010.
- 40 LEE, V. et al. A cyclic-di-GMP receptor required for bacterial exopolysaccharide production. **Molecular Microbiology**, v. 65, n. 6, p. 1474-1484, Sept. 2007.

- 41 QI, Y. et al. Binding of cyclic diguanylate in the non-catalytic EAL domain of FimX induces a long-range conformational change. **Journal of Biological Chemistry**, v. 286, n. 4, p. 2910-2917, Jan. 2011.
- 42 NAVARRO, M. et al. Structural basis for c-di-GMP-mediated inside-out signaling controlling periplasmic proteolysis. **Plos Biology**, v. 9, n. 2, p. e1000588, Feb. 2011. doi: 10.1371/journal.pbio.1000588.
- 43 BARAQUET, C. et al. The FleQ protein from *Pseudomonas aeruginosa* functions as both a repressor and an activator to control gene expression from the *pel* operon promoter in response to c-di-GMP. **Nucleic Acids Research**, v. 40, n. 15, p. 7207-7218, Aug. 2012.
- 44 MATSUYAMA, B. Y. et al. Mechanistic insights into c-di-GMP-dependent control of the biofilm regulator FleQ from *Pseudomonas aeruginosa*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 113, n. 2, p. E209-E218, Jan. 2016.
- 45 KRASTEVA, P. et al. *Vibrio cholerae* VpsT regulates matrix production and motility by directly sensing Cyclic di-GMP. **Science**, v. 327, n. 5967, p. 866-868, Feb. 2010.
- 46 CHIN, K. et al. The cAMP receptor-like protein CLP Is a novel c-di-GMP receptor linking cell-cell signaling to virulence gene expression in *Xanthomonas campestris*. **Journal of Molecular Biology**, v. 396, n. 3, p. 646-662, Feb. 2010.
- 47 BARRICK, J.; BREAKER, R. The distributions, mechanisms, and structures of metabolite-binding riboswitches. **Genome Biology**, v. 8, n. 11, p. R239, Nov. 2007.
- 48 WEINBERG, Z. et al. Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline. **Nucleic Acids Research**, v. 35, n. 14, p. 4809-4819, July 2007.
- 49 LEE, E. et al. An Allosteric Self-Splicing Ribozyme Triggered by a Bacterial Second Messenger. **Science**, v. 329, n. 5993, p. 845-848, Aug. 2010.
- 50 CHEN, A.; SUDARSAN, N.; BREAKER, R. Mechanism for gene control by a natural allosteric group I ribozyme. **RNA**, v. 17, n. 11, p. 1967-1972, Nov. 2011.
- 51 DISTRIBUTION of GGDEF, EAL, HD-GYP and PilZ domains in bacterial genomes. 2015. Disponível em: <http://www.ncbi.nlm.nih.gov/Complete_Genomes/c-di-GMP.html>. Acesso em: 01 nov. 2015.
- 52 HOFFMAN, L. et al. Aminoglycoside antibiotics induce bacterial biofilm formation. **Nature**, v. 436, n. 7054, p. 1171-1175, Aug. 2005.
- 53 RYAN, R. P.; DOW, J. M. Intermolecular interactions between HD-GYP and GGDEF domain proteins mediate virulence-related signal transduction in *Xanthomonas campestris*. **Virulence**, v. 1, n. 5, p. 404-408, Sept.-Oct. 2010.

- 54 KULESEKARA, H. et al. Analysis of *Pseudomonas aeruginosa* diguanylate cyclases and phosphodiesterases reveals a role for bis-(3'-5')-cyclic-GMP in virulence. **Proceedings of the National Academy of Sciences of the United States of America**, v. 103, n. 8, p. 2839-2844, Feb. 2006.
- 55 HECHT, G.; NEWTON, A. Identification of a novel response regulator required for the swarmer-to-stalked-cell transition in *caulobacter-crescentus*. **Journal of Bacteriology**, v. 177, n. 21, p. 6223-6229, Nov. 1995.
- 56 TAL, R. et al. Three *cdg* operons control cellular turnover of cyclic di-GMP in *Acetobacter xylinum*: genetic organization and occurrence of conserved domains in isoenzymes. **Journal of Bacteriology**, v. 180, n. 17, p. 4416-4425, Sept. 1998.
- 57 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.
- 58 GALPERIN, M. Y. et al. A specialized version of the HD hydrolase domain implicated in signal transduction. **Journal of Molecular Microbiology and Biotechnology**, v. 1, n. 2, p. 303-305, Nov. 1999.
- 59 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.
- 60 MASSIE, J. et al. Quantification of high-specificity cyclic diguanylate signaling. **Proceedings of the National Academy of Sciences of the United States of America**, v. 109, n. 31, p. 12746-12751, July 2012.
- 61 HOBLEY, L. et al. Discrete Cyclic di-GMP-dependent control of bacterial predation versus axenic growth in *bdellovibrio bacteriovorus*. **Plos Pathogens**, v. 8, n. 2, Feb. 2012.
- 62 LINDENBERG, S. et al. The EAL domain protein YciR acts as a trigger enzyme in a c-di-GMP signalling cascade in *E. coli* biofilm control. **Embo Journal**, v. 32, n. 14, p. 2001-2014, July 2013.
- 63 TSCHOWRI, N.; BUSSE, S.; HENGGE, R. The BLUF-EAL protein YcgF acts as a direct anti-repressor in a blue-light response of *Escherichia coli*. **Genes & Development**, v. 23, n. 4, p. 522-534, Feb. 2009.
- 64 RYAN, R. et al. Dynamic complex formation between HD-GYP, GGDEF and PilZ domain proteins regulates motility in *Xanthomonas campestris*. **Molecular Microbiology**, v. 86, n. 3, p. 557-567, Nov. 2012.
- 65 JENAL, U. Think globally, act locally: How bacteria integrate local decisions with their global cellular programme. **Embo Journal**, v. 32, n. 14, p. 1972-1974, July 2013.
- 66 GUZZO, C. et al. Structure of the PilZ-FimX(EAL)-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.

67 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.

68 MIKKELSEN, H.; MCMULLAN, R.; FILLOUX, A. The *Pseudomonas aeruginosa* Reference Strain PA14 Displays Increased Virulence Due to a Mutation in *ladS*. **Plos One**, v. 6, n. 12, p. e29113, Dec. 2011.

69 BACTERIOMATCH II two-hybrid system vector kit: instruction manual. revisão C. La Jolla, CA: Agilent Technologie, 2011. 30 p.

70 JOUNG, J.; RAMM, E.; PABO, C. A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. **Proceedings of the National Academy of Sciences of the United States of America**, v. 97, n. 13, p. 7382-7387, June 2000.

71 BACTERIOMATCH II two-hybrid system. Disponível em: <<http://www.chem-agilent.com/contents.php?id=300052>>. Acesso em: 05 nov. 2013.

72 ASLANIDIS, C.; DEJONG, P. Ligation-independent cloning of PCR products (LIC-PCR). **Nucleic Acids Research**, v. 18, n. 20, p. 6069-6074, Oct. 1990.

73 CAMILO, C.; POLIKARPOV, I. High-throughput cloning, expression and purification of glycoside hydrolases using Ligation-Independent Cloning (LIC). **Protein Expression and Purification**, v. 99, p. 35-42, July 2014. doi:10.1016/j.pep.2014.03.008

74 CAMILO, C. et al. HTP-oligodesigner: an online primer design tool for high-throughput gene cloning and site-directed mutagenesis. **Journal of Computational Biology**, v. 23, n. 1, p. 27-29, Jan. 2016.

75 SAMBROOK, J.; RUSSELL, D. W. The inoue method for preparation and transformation of competent *E. coli*: "ultra-competent" cells. **CSH Protocols**, v. 2006, n. 1, 2006. doi: 10.1101/pdb.prot3944.

76 HIDALGO, P. et al. Recruitment of the transcriptional machinery through GAL11P: structure and interactions of the GAL4 dimerization domain. **Genes & Development**, v. 15, n. 8, p. 1007-1020, Apr. 2001.

77 WALKER, S. L. et al. ToxR (RegA)-mediated in vitro transcription of *Pseudomonas aeruginosa* *toxA*. **Gene**, v. 150, n. 1, p. 87-92, Dec. 1994.

78 MIKKELSEN, H. et al. Expression of *Pseudomonas aeruginosa* CupD Fimbrial Genes Is Antagonistically Controlled by RcsB and the EAL-Containing PvrR Response Regulators. **Plos One**, v. 4, n. 6, p. e6018, June 2009.

79 RAVICHANDRAN, A. et al. Global regulator MorA affects virulence-associated protease secretion in *Pseudomonas aeruginosa* PAO1. **PLoS One**, v. 10, n. 4, p. e0123805, 2015.

- 80 AN, S.; WU, J.; ZHANG, L. Modulation of *Pseudomonas aeruginosa* biofilm dispersal by a Cyclic-Di-GMP phosphodiesterase with a putative hypoxia-sensing domain. **Applied and Environmental Microbiology**, v. 76, n. 24, p. 8160-8173, Dec. 2010.
- 81 HAY, I. D.; REMMINGHORST, U.; REHM, B. H. MucR, a novel membrane-associated regulator of alginate biosynthesis in *Pseudomonas aeruginosa*. **Applied and Environmental Microbiology**, v. 75, n. 4, p. 1110-1120, Feb. 2009.
- 82 KUCHMA, S. et al. BifA, a cyclic-di-GMP phosphodiesterase, inversely regulates biofilm formation and swarming motility by *Pseudomonas aeruginosa* PA14. **Journal of Bacteriology**, v. 189, n. 22, p. 8165-8178, Nov. 2007.
- 83 ROY, A. B.; PETROVA, O. E.; SAUER, K. The phosphodiesterase DipA (PA5017) is essential for *Pseudomonas aeruginosa* biofilm dispersion. **Journal of Bacteriology**, v. 194, n. 11, p. 2904-2915, June 2012.
- 84 UEDA, A.; WOOD, T. K. Connecting quorum sensing, c-di-GMP, pel polysaccharide, and biofilm formation in *Pseudomonas aeruginosa* through tyrosine phosphatase TpbA (PA3885). **PLoS Pathogens**, v. 5, n. 6, p. e1000483, June 2009.
- 85 GIARDINA, G. et al. Investigating the allosteric regulation of YfiN from *Pseudomonas aeruginosa*: clues from the structure of the catalytic domain. **PLoS One**, v. 8, n. 11, p. e81324, 2013.
- 86 KLEBENSBERGER, J. et al. SiaA and SiaD are essential for inducing autoaggregation as a specific response to detergent stress in *Pseudomonas aeruginosa*. **Environmental Microbiology**, v. 11, n. 12, p. 3073-3086, Dec. 2009.
- 87 DE, N. et al. Determinants for the activation and autoinhibition of the diguanylate cyclase response regulator WspR. **Journal of Molecular Biology**, v. 393, n. 3, p. 619-633, Oct. 2009.
- 88 MERRITT, J. H. et al. SadC reciprocally influences biofilm formation and swarming motility via modulation of exopolysaccharide production and flagellar function. **Journal of Bacteriology**, v. 189, n. 22, p. 8154-8164, Nov. 2007.
- 89 MERRITT, J. H. et al. Specific control of *Pseudomonas aeruginosa* surface-associated behaviors by two c-di-GMP diguanylate cyclases. **MBio**, v. 1, n. 4, p. e00183-10, Oct. 2010.
- 90 PETROVA, O.; CHERNY, K.; SAUER, K. The diguanylate Cyclase GcbA facilitates *Pseudomonas aeruginosa* biofilm dispersion by activating BdlA. **Journal of Bacteriology**, v. 197, n. 1, p. 174-187, Jan. 2015.
- 91 ALM, R. A. et al. Identification of a novel gene, pilZ, essential for type 4 fimbrial biogenesis in *Pseudomonas aeruginosa*. **Journal of Bacteriology**, v. 178, n. 1, p. 46-53, Jan. 1996.
- 92 MERIGHI, M. et al. The second messenger bis-(3'-5')-cyclic-GMP and its PilZ domain-containing receptor Alg44 are required for alginate biosynthesis in *Pseudomonas aeruginosa*. **Molecular Microbiology**, v. 65, n. 4, p. 876-95, Aug. 2007.

93 TMHMM result. Disponível em: <<http://www.cbs.dtu.dk/cgi-bin/webface2.fcgi?jobid=56B369C3000052449ACDC640&wait=20>>. Acesso em: 05 jun. 2012.

94 TMHMM result. Disponível em: <<http://www.cbs.dtu.dk/cgi-bin/webface2.fcgi?jobid=56B369CB00005244D50C22AE&wait=20>>. Acesso em: 05 jun. 2012.

95 PFAM: sequence search results. Disponível em: <<http://pfam.xfam.org/search/>>. Acesso em: 01 fev. 2012.

APÊNDICE A – Tabelas adicionais

Tabela A1 - Lista de primers para amplificação das sequências alvo para clonagem pelo método de dupla-restrição enzimática.

ID	F / R	Vetor	# pb	Identificação	Primer
PA0707	F	pBT/pTRG	34	PA0707_1F_NotI	CATGGCGGCCGCAATGACTGCGACAGACAGAACG
	R	pBT/pTRG	26	PA0707_780R_XhoI	CATGCTCGAGTCAGCAGGCCGGACTG
PA3825	F	pBT/pTRG	28	PA3825_106F_NotI	CATGGCGGCCGCACAGGCCGAGCGAAGC
	R	pBT/pTRG	28	PA3825_702R_XhoI	CATGCTCGAGTCACGGGTTCTGCGAGCG
	F	pBT/pTRG	29	PA3825_763F_NotI	CATGGCGGCCGCATCCAGGCCGGTTCGCCT
	R	pBT/pTRG	26	PA3825_1572R_XhoI	CATGCTCGAGCTAGCGCCCGGCACTG
PA2133	F	pBT/pTRG	32	PA2133_1F_NotI	CATGGCGGCCGCAAGTGAACAGTTCCCCACAGG
	R	pBT	25	PA2133_858R_BglII_BT	CATGAGATCTTACCCTGGCGGCT
	R	pTRG	25	PA2133_858R_SpeI_TRG	CATGACTAGTTCACCCTGGCGGCT
tnpM	F	pBT/pTRG	29	tnpM_1F_NotI	CATGGCGGCCGCAATGAGCGCTTTCGGGC
	R	pBT/pTRG	31	tnpM_990R_XhoI	CATGCTCGAGTCAATCTGACGGTAGCGAGTC
PA2200	F	pBT	36	PA2200_109F_NotI_BT	CATGGCGGCCGCAACGAACTGGATCGCCAGCACG A
	R	pBT	30	PA2200_747R_BamHI_BT	CATGGGATCCTCAGGCCGCGCTCCGGGCT
	F	pTRG	33	PA2200_109F_BamHI_TRG	CATGGGATCCAACGAACTGGATCGCCAGCACGA
	R	pTRG	30	PA2200_747R_EcoRI_TRG	CATGGAATCTCAGGCCGCGCTCCGGGCT
	F	pBT	31	PA2200_817F_EcoRI_BT	CATGGAATTCACTCAACGACGAAAAAAGTCC
	F	pTRG	32	PA2200_817F_EcoRI_TRG	CATGGAATCTTCTCAACGACGAAAAAAGTCC
	R	pBT/pTRG	25	PA2200_1596R_XhoI	CATGCTCGAGCTAGCGCCCGTCGG
PA3947	F	pBT/pTRG	37	PA3947_1F_NotI	CATGGCGGCCGCAATGAATGATTTGAATGTTCTGGTG
	R	pBT	30	PA3947_1179R_BglII_BT	CATGAGATCTTCAGGATCCGGAGCAATAGT
	R	pTRG	30	PA3947_1179R_SpeI_TRG	CATGACTAGTTCAGGATCCGGAGCAATAGT
pvrR	F	pBT/pTRG	34	pvrR_1F_NotI	CATGGCGGCCGCAATGAGCTGGAATCTCTATCGG
	R	pBT/pTRG	29	pvrR_1200R_XhoI	CATGCTCGAGTTAAACAAGGAGGCTGGGG
PA3311	F	pBT/pTRG	31	PA3311_1F_NotI	CATGGCGGCCGCAATGCCTTTTCTCCCTGGG
	R	pBT/pTRG	31	PA3311_240R_XhoI	CATGCTCGAGTCAGTTCTGCAGCACGTAGCC
	F	pBT/pTRG	31	PA3311_940F_NotI	CATGGCGGCCGCATGGTCGGAGAAACAACGG
	R	pBT/pTRG	27	PA3311_2358R_XhoI	CATGCTCGAGTCAGGCCTGGTTCAAGC
PA5442	F	pBT/pTRG	28	PA5442_595F_NotI	CATGGCGGCCGCACACCGGCTGGTGCAG
	R	pBT	25	PA5442_2853R_BamHI_BT	CATGGGATCCTCAACGAGCCTGGCG
	R	pTRG	25	PA5442_2853R_SpeI_TRG	CATGACTAGTTCAACGAGCCTGGCG
PA4601	F	pBT/pTRG	32	PA4601_127F_NotI	CATGGCGGCCGCACAGATGAGCCAGGAGCTGC
	R	pBT	28	PA4601_735R_BglII_BT	CATGAGATCTTACCAGCACCCGGGTGCT
	R	pTRG	28	PA4601_735R_SpeI_TRG	CATGACTAGTTCACCGCACCCGGGTGCT
	F	pBT/pTRG	28	PA4601_805F_NotI	CATGGCGGCCGCAATGCTGTGCGCCTG
	R	pBT	29	PA4601_4248R_BglII_BT	CATGAGATCTTACGCCCTCGTTGAACATG
	R	pTRG	29	PA4601_4248R_SpeI_TRG	CATGACTAGTTCAGCCCTCGTTGAACATG
PA0861	F	pBT	40	PA0861_1F_EcoRI_BT	CATGGAATTCATTGAATGGTAGATGGAATGAGGCA GAACC
	R	pBT	30	PA0861_618R_BglII_BT	CATGAGATCTTACCAGCAGCCCTCGCCGA
	F	pTRG	41	PA0861_1F_EcoRI_TRG	CATGGAATTCCTTTGAATGGTAGATGGAATGAGGAGAACC
	R	pTRG	30	PA0861_618R_SpeI_TRG	CATGACTAGTTCACCGCAGCCCTCGCCGA
	F	pBT	28	PA0861_688F_EcoRI_BT	CATGGAATTCACGTACCCGGCAGCAACT
	R	pBT	29	PA0861_2472R_BglII_BT	CATGAGATCTTACCAGGAGTTCTGTCCG
	F	pTRG	29	PA0861_688F_EcoRI_TRG	CATGGAATCTTCTGTACCCGGCAGCAACT
PA1727	R	pTRG	29	PA0861_2472R_SpeI_TRG	CATGACTAGTTCACCGAGGTTCTGTCCG
	F	pBT/pTRG	29	PA1727_712F_NotI	CATGGCGGCCGCAAGTTCAGCGCCCGCA
	R	pBT	25	PA1727_2058R_BamHI_BT	CATGGGATCCTCAGGCGACGCTGGC
PA5295	R	pTRG	25	PA1727_2058R_SpeI_TRG	CATGACTAGTTCAGGCGACGCTGGC
	F	pBT/pTRG	28	PA5295_1F_NotI	CATGGCGGCCGCAATTGTGACGCGCCGCA
	R	pBT	25	PA5295_1677R_BglII_BT	CATGAGATCTTCAATTGCCGAGCGG
PA2072	R	pTRG	25	PA5295_1677R_SpeI_TRG	CATGACTAGTTCAAATTGCCGAGCGG
	F	pBT/pTRG	28	PA2072_118F_NotI	CATGGCGGCCGCAATCGCCAGGCAGCAG
	R	pBT	29	PA2072_765R_BglII_BT	CATGAGATCTTCAAGGTCTCCCGCAGCATG
	R	pTRG	29	PA2072_765R_SpeI_TRG	CATGACTAGTTCAGGTCTCCCGCAGCATG
	F	pBT/pTRG	28	PA2072_820F_NotI	CATGGCGGCCGCAAGCCGACGCGCCGTG
	R	pBT	25	PA2072_2595R_BglII_BT	CATGAGATCTTCAAGGCCGCGCGCGC
PA0575	R	pTRG	25	PA2072_2595R_SpeI_TRG	CATGACTAGTTCAGGCCGCGCGCGC
	F	pBT/pTRG	29	PA0575_61F_NotI	CATGGCGGCCGCAAGGCGCCCTCACGCTCA
	R	pBT	29	PA0575_801R_BamHI_BT	CATGGGATCCTCACTCGCGCCAGGCCGGA
	R	pTRG	29	PA0575_801R_SpeI_TRG	CATGACTAGTTCACCTCGCGCCAGGCCGGA
	F	pBT/pTRG	31	PA0575_871F_NotI	CATGGCGGCCGCACGCCGCTGCGCCGCGAG
	R	pBT	35	PA0575_3738R_BamHI_BT	CATGGGATCCTTATACCGCGCTTCTCGGGGAGCG
	R	pTRG	35	PA0575_3738R_SpeI_TRG	CATGACTAGTTTATACCGCGCTTCTCGGGGAGCG

(continua)

(continuação)

Tabela A1 - Lista de primers para amplificação das sequências alvo para clonagem pelo método de dupla-restrição enzimática.

ID	F / R	Vetor	# pb	Identificação	Primer
PA1433	F	pBT/pTRG	27	PA1433_82F_BamHI	CATGGGATCCGAGAAGTCCCGCGAGCA
	R	pBT	28	PA1433_441R_BglII_BT	CATGAGATCTTCAGCTGTCCACAGGCG
	R	pTRG	28	PA1433_441R_SpeI_TRG	CATGACTAGTTCAGCTGTCCACAGGCG
	F	pBT/pTRG	25	PA1433_511F_BamHI	CATGGGATCCCGCCGCAACTGCGC
	R	pBT	27	PA1433_1953R_BglII_BT	CATGAGATCTTCAGGCATCCCCGGACC
	R	pTRG	27	PA1433_1953R_SpeI_TRG	CATGACTAGTTCAGGCATCCCCGGACC
PA0285	F	pBT/pTRG	30	PA0285_184F_NotI	CATGGCGGCCGCACTGGTACGCCAATTCCG
	R	pBT/pTRG	27	PA0285_2283R_XhoI	CATGCTCGAGTCAGTCTTCCGGCAACG
PA2567	F	pBT/pTRG	28	PA2567_1F_NotI	CATGGCGGCCGAGTGGCCGACGCCACC
	R	pBT	27	PA2567_1500R_BglIII_BT	CATGAGATCTCTAGCGCCGACGCCAGT
	R	pTRG	27	PA2567_1500R_SpeI_TRG	CATGACTAGTCTAGCGCCGACGCCAGT
PA3258	F	pBT/pTRG	29	PA3258_1F_NotI	CATGGCGGCCGAGTGAATACGGGGCCG
	R	pBT/pTRG	25	PA3258_1806R_XhoI	CATGCTCGAGTCAGGCGGCCTCGAT
PA4959	F	pBT/pTRG	33	PA4959_1F_NotI	CATGGCGGCCGCAATGGCCATCGAAAAGAAAAAC
	R	pBT	28	PA4959_2076R_BglIII_BT	CATGAGATCTTCATTCTGCTCTCCGAGGA
	R	pTRG	28	PA4959_2076R_SpeI_TRG	CATGACTAGTTCATTCTGCTCTCCGAGGA
PA4367	F	pBT/pTRG	31	PA4367_106F_NotI	CATGGCGGCCGACAGCCTACAAGGCCAAG
	R	pBT/pTRG	33	PA4367_474R_XhoI	CATGCTCGAGTCATTCCGAAGTGGTGACGAAGT
	F	pBT/pTRG	31	PA4367_544F_NotI	CATGGCGGCCGCAATGTGACCAAACCGCTG
	R	pBT/pTRG	25	PA4367_2064R_XhoI	CATGCTCGAGTCAGGGCCGTTCGCT
PA5017	F	pBT/pTRG	33	PA5017_1F_NotI	CATGGCGGCCGCAATGAAAAGTCATCCCGATGC
	R	pBT	26	PA5017_2700R_BglIII_BT	CATGAGATCTTCAGTGCAGGGTGCGG
	R	pTRG	26	PA5017_2700R_SpeI_TRG	CATGACTAGTTCAGTGCAGGGTGCGG
PA1181	F	pBT/pTRG	25	PA1181_853F_BamHI	CATGGGATCCATGGAGCGCAGCCGC
	R	pBT/pTRG	27	PA1181_3363R_XhoI	CATGCTCGAGTCAGCCCAACTCCTGGC
PA5487	F	pBT/pTRG	28	PA5487_1F_BamHI	CATGGGATCCATGAGTCGCGACGACGTC
	R	pBT	27	PA5487_2016R_BglIII_BT	CATGAGATCTTCAGGCCACTTCCAGGC
	R	pTRG	27	PA5487_2016R_SpeI_TRG	CATGACTAGTTCAGGCCACTTCCAGGC
PA1120	F	pBT/pTRG	28	PA1120_130F_NotI	CATGGCGGCCGCAACCCTGCGCGCCTAC
	R	pBT/pTRG	28	PA1120_462R_XhoI	CATGCTCGAGTCAGCGCAGCAGGCTGCC
	F	pBT/pTRG	29	PA1120_532F_NotI	CATGGCGGCCGATCACGGCGGCTGGTAC
	R	pBT/pTRG	30	PA1120_1308R_XhoI	CATGCTCGAGTACTTGGGTGGAGCCTCTG
PA0169	F	pBT/pTRG	25	PA0169_1F_BamHI	CATGGGATCCGTGCGGCTGGAGCGC
	R	pBT	25	PA0169_708R_BglII_BT	CATGAGATCTTCAGCGCGCTGGAGC
	R	pTRG	25	PA0169_708R_SpeI_TRG	CATGACTAGTTCAGCGCGCTGGAGC
PA0847	F	pBT/pTRG	30	PA0847_184F_NotI	CATGGCGGCCGCAAGCCGTTTCGACCGAGA
	R	pBT	28	PA0847_945R_BglII_BT	CATGAGATCTTCAGGCCCTTCTCGCCCTG
	R	pTRG	28	PA0847_945R_SpeI_TRG	CATGACTAGTTCAGGCCCTTCTCGCCCTG
	F	pBT/pTRG	30	PA0847_1015F_NotI	CATGGCGGCCGAGAGCTGTGGGTCCTGCG
	R	pBT	25	PA0847_2046R_BglII_BT	CATGAGATCTTCAGGCTGGCGCCTG
PA4396	R	pTRG	25	PA0847_2046R_SpeI_TRG	CATGACTAGTTCAGGCTGGCGCCTG
	F	pBT/pTRG	28	PA4396_1F_NotI	CATGGCGGCCGCAATGCCCGCCGGTGTG
	R	pBT	28	PA4396_1101R_BglII_BT	CATGAGATCTCTAGCGCAGTGCCGGTAG
PA3177	R	pTRG	28	PA4396_1101R_SpeI_TRG	CATGACTAGTCTAGCGCAGTGCCGGTAG
	F	pBT/pTRG	30	PA3177_1F_NotI	CATGGCGGCCGCAATGGCTCTCCCAAGCA
PA0290	R	pBT/pTRG	25	PA3177_924R_XhoI	CATGCTCGAGTCAGGCGCAGCGCAG
	F	pBT/pTRG	30	PA0290_1F_NotI	CATGGCGGCCGCAATGGACGACCTACCCGG
PA3702	R	pBT/pTRG	26	PA0290_972R_XhoI	CATGCTCGAGTCAGCCCACGACGATG
	F	pBT/pTRG	36	PA3702_1F_NotI	CATGGCGGCCGCAATGCACAACCCTCATGAGAGCAA
PA2870	R	pBT/pTRG	26	PA3702_1044R_XhoI	CATGCTCGAGTCAGCCCAGCGGGCT
	F	pBT/pTRG	28	PA2870_121F_NotI	CATGGCGGCCGACGCCAGGCCAGCAA
	R	pBT	28	PA2870_555R_BglIII_BT	CATGAGATCTTCAGAACGCCCGGTGCAG
	R	pTRG	28	PA2870_555R_SpeI_TRG	CATGACTAGTTCAGAACGCCCGGTGCAG
	F	pBT/pTRG	29	PA2870_625F_NotI	CATGGCGGCCGAGACTTCAAGCGCGCGC
	R	pBT	25	PA2870_1578R_BglII_BT	CATGAGATCTTCAGGCGGGCGCGTA
PA4929	R	pTRG	25	PA2870_1578R_SpeI_TRG	CATGACTAGTTCAGGCGGGCGCGTA
	F	pBT/pTRG	28	PA4929_187F_NotI	CATGGCGGCCGAGCGCCGAGTTTCCA
	R	pBT/pTRG	31	PA4929_666R_XhoI	CATGCTCGAGTCACTGGCTGCGGTATTCCAG
	F	pBT/pTRG	29	PA4929_1276F_NotI	CATGGCGGCCGATCGCGCATCCAGGATC
PA4929	R	pBT/pTRG	29	PA4929_2022R_XhoI	CATGCTCGAGTCAGGAGATACGCTCCGGT

(continua)

(conclusão)

Tabela A1 - Lista de primers para amplificação das sequências alvo para clonagem pelo método de dupla-restrição enzimática.

ID	F / R	Vetor	# pb	Identificação	Primer
PA1851	F	pBT/pTRG	29	PA1851_1F_NotI	CATGGCGGCCGCAATGCTTGCACGCGACG
	R	pBT/pTRG	28	PA1851_153R_XhoI	CATGCTCGAGTCAACGCCGCTGCATCAG
	F	pBT/pTRG	29	PA1851_637F_NotI	CATGGCGGCCGCGAGACTACGCCCAGCGCG
	R	pBT/pTRG	28	PA1851_1206R_XhoI	CATGCTCGAGCTATTCGAGGCGGTCTCGTG
PA0338	F	pBT/pTRG	28	PA0338_1F_NotI	CATGGCGGCCGCGAGTGCAGCGCTTACGC
	R	pBT	26	PA0338_1131R_BglII_BT	CATGAGATCTTCAGCAACAGGCCACG
	R	pTRG	26	PA0338_1131R_SpeI_TRG	CATGACTAGTTCAGCAACAGGCCACG
PA4332	F	pBT/pTRG	28	PA4332_544F_NotI	CATGGCGGCCGCAATGCGGCAACGCATG
	R	pBT	28	PA4332_1128R_BglII_BT	CATGAGATCTTCAGGCACTGGTGACCTC
	R	pTRG	28	PA4332_1128R_SpeI_TRG	CATGACTAGTTCAGGCACTGGTGACCTC
PA1107	F	pBT/pTRG	29	PA1107_592F_NotI	CATGGCGGCCGCGACAGTGCAGCAACCTGC
	R	pBT/pTRG	27	PA1107_1197R_XhoI	CATGCTCGAGTTACCGCAGGCTTTCCG
PA4843	F	pBT/pTRG	32	PA4843_1F_NotI	CATGGCGGCCGCAATGATGACCGAGCAGCATG
	R	pBT	25	PA4843_1629R_BglII_BT	CATGAGATCTTTACCGGCGTCGGG
	R	pTRG	25	PA4843_1629R_SpeI_TRG	CATGACTAGTTTACCGGCGTCGGG
PA3343	F	pBT/pTRG	33	PA3343_1F_NotI	CATGGCGGCCGCGAGTGTGCGTGACACAGAAGGA
	R	pBT/pTRG	29	PA3343_153R_XhoI	CATGCTCGAGTCAGCGTTTCAGCGCGAAG
	F	pBT/pTRG	30	PA3343_619F_NotI	CATGGCGGCCGCAAGTCGCGCGAGCATTT
	R	pBT/pTRG	28	PA3343_1170R_XhoI	CATGCTCGAGTCAGGCCACACCTGATT
PA2960	F	pBT/pTRG	33	PA2960_1F_NotI	CATGGCGGCCGCAATGAGTTTGCACCCCAATCT
	R	pBT/pTRG	28	PA2960_357R_XhoI	CATGCTCGAGTTACATCGTGTGGGTCTCG
PA4608	F	pBT/pTRG	32	PA4608_1F_NotI	CATGGCGGCCGCAATGAGTGACCAGCAGCAGC
	R	pBT	26	PA4608_378R_BglII_BT	CATGAGATCTTCAGTCGTCGTGGGCG
	R	pTRG	26	PA4608_378R_SpeI_TRG	CATGACTAGTTCAGTCGTCGTGGGCG
PA2989	F	pBT/pTRG	29	PA2989_1F_NotI	CATGGCGGCCGCAATGGCTGCGACGCTTTG
	R	pBT/pTRG	29	PA2989_765R_XhoI	CATGCTCGAGTCAGGTGCAAGAAGGTTGG
PA0012	F	pBT/pTRG	36	PA0012_1F_NotI	CATGGCGGCCGCAATGGATAATCAACGACATAACCC
	R	pBT/pTRG	29	PA0012_267R_XhoI	CATGCTCGAGTCAGGCTTCGCTGAGAAAG
PA3542	F	pBT/pTRG	33	PA3542_1F_BamHI	CATGGGATCCATGAATACAGCCGTCACGTC
	R	pBT	28	PA3542_474R_BglII_BT	CATGAGATCTTCACCGGACCCGCGCCGAA
	R	pTRG	28	PA3542_474R_SpeI_TRG	CATGACTAGTTCACCGGACCCGCGCCGAA
	F	pBT/pTRG	35	PA3542_535F_BamHI	CATGGGATCCAACAGATGTACAACCTGTACTTCG
	R	pBT	25	PA3542_1170R_BglII_BT	CATGAGATCTTCAGCGAGCGGTGGC
	R	pTRG	25	PA3542_1170R_SpeI_TRG	CATGACTAGTTCAGCGAGCGGTGGC
PA2799	F	pBT/pTRG	29	PA2799_1F_NotI	CATGGCGGCCGCAATGCAGCCCATCGAGC
	R	pBT/pTRG	28	PA2799_300R_XhoI	CATGCTCGAGTCAGTGGATCGCGACGAT
PA3353	F	pBT/pTRG	34	PA3353_1F_NotI	CATGGCGGCCGCGAGTGCTATCATTGAGGCAATTCG
	R	pBT/pTRG	32	PA3353_792R_XhoI	CATGCTCGAGTCAGAACAGTTCGCTCTTCTCG
PA4108	F	pBT/pTRG	28	PA4108_1F_BamHI	CATGGGATCCGCTGCTGAAACGCATTCCG
	R	pBT	28	PA4108_1245R_BglII_BT	CATGAGATCTCTACGGCTTTCCGGTTCC
	R	pTRG	28	PA4108_1245R_SpeI_TRG	CATGACTAGTCTACGGCTTTCCGGTTCC
PA2572	F	pBT/pTRG	32	PA2572_1F_NotI	CATGGCGGCCGCAATGAACGATAGCGCACCTC
	R	pBT/pTRG	28	PA2572_1344R_XhoI	CATGCTCGAGCTAGGTCTGCTGACTCCGG
PA4781	F	pBT/pTRG	32	PA4781_1F_NotI	CATGGCGGCCGCAATGGAGAGCATGCTGGACA
	R	pBT	26	PA4781_1182R_BglII_BT	CATGAGATCTCTAGGCTGGTGGCGCG
	R	pTRG	26	PA4781_1182R_SpeI_TRG	CATGACTAGTCTAGGCTGGTGGCGCG

Fonte: Elaborada pela autora.

Tabela A2 - Lista de primers para amplificação das sequências alvo para clonagem pela técnica LIC.

Construção	Primers diretos (extensão LIC: 5'-CAGGGCGCCA-3')	#pb	T _m (°C)	Primers reversos (extensão LIC: 5'-GACCCGACGCGG-3')	#pb	T _m (°C)
PA0707_A1	CAGGGCGCCATGACTGCGACAGACAGAACGC	31	63,2	GACCCGACGCGGTACAGAGCCGGAC	26	63,1
PA3825_B1	CAGGGCGCCATGCAGGCCGAGCGAAGC	27	63,6	GACCCGACGCGGTACAGGGTTCTGCGAG	28	62,1
PA3825_C1	CAGGGCGCCATGTCCAGGCGGGTCG	25	58,5	GACCCGACGCGGTACAGCGCCCGGC	24	61,8
PA3825_D1	CAGGGCGCCATGCAGGCCGAGCGAAGC	27	63,6	GACCCGACGCGGTACAGCGCCCGGC	24	61,8
PA2133_E1	CAGGGCGCCATGGTGAACAGTTCCCCACAGG	31	61,9	GACCCGACGCGGTACCCCTGGCGG	25	59,9
tnpM_F1	CAGGGCGCCATGAGCGCTTTCGGGC	26	62,8	GACCCGACGCGGTCAATCTGACGGTAGCGAG	31	63,4
PA2200_G1	CAGGGCGCCATGAACGAACTGGATCGCCAG	30	62,7	GACCCGACGCGGTACAGGCCGCGC	24	62,0
PA2200_H1	CAGGGCGCCATGCTCAACGACGAAAACTGCC	32	63,7	GACCCGACGCGGTACAGCGCCCGTC	25	62,9
PA3947_A2	CAGGGCGCCATGAATGATTGAATGTTCTGGTGTG	36	63,3	GACCCGACGCGGTACAGGATCCGGAGCAATAG	31	63,2
pvrR_B2	CAGGGCGCCATGAGCTGGAATCTATCGGG	31	61,8	GACCCGACGCGGTCAAACAAGGAGGCTGGG	30	62,8
PA3311_C2	CAGGGCGCCATGCCTTTTCTCCCTGGGAAG	30	63,8	GACCCGACGCGGTACAGTTCTGACAGCAGTAGC	32	63,2
PA3311_D2	CAGGGCGCCATGTGGTCGGAGAAACAACGG	30	63,8	GACCCGACGCGGTACAGGCTGGTTCAGG	28	59,5
PA5442_E2	CAGGGCGCCATGCACCGGCTGGTGACG	27	62,8	GACCCGACGCGGTCAACGAGCCTGGCG	27	61,0
PA4601_F2	CAGGGCGCCATGCAGATGAGCCAGGAGCTG	30	62,1	GACCCGACGCGGTACCCGACCCGG	25	61,5
PA4601_G2	CAGGGCGCCATGCTGCTGCGCTGCAC	27	63,5	GACCCGACGCGGTACGCCCTCGTTGAACATG	31	61,3
PA0861_H2	CAGGGCGCCATGTTGAATGGTAGATGGAATGAGG	34	61,5	GACCCGACGCGGTACCCGAGGCC	25	60,4
PA0861_A3	CAGGGCGCCATGCTACCCGGCAGCAAC	28	61,3	GACCCGACGCGGTACCCGAGGTTCTGTCC	30	62,8
PA1727_B3	CAGGGCGCCATGGACTCGCGCTGGAG	27	61,8	GACCCGACGCGGTACAGCGACGCTGG	26	60,7
PA5295_C3	CAGGGCGCCATGTTGTCAGCGCCCGC	26	63,7	GACCCGACGCGGTCAATTGCCGAGCGG	27	62,2
PA2072_D3	CAGGGCGCCATGGCCAGGACGAGGAC	27	61,8	GACCCGACGCGGTACAGGTTCCCGCAGC	28	62,8
PA2072_E3	CAGGGCGCCATGAGCCGACGCGCC	24	60,8	GACCCGACGCGGTCAAGGCCGCGC	25	63,6
PA0575_F3	CAGGGCGCCATGGGCGCCCTCACGC	25	62,9	GACCCGACGCGGTACTCGCGCCAGG	26	60,7
PA0575_G3	CAGGGCGCCATGCGCCGCTGCG	23	59,1	GACCCGACGCGGTACACCGCGCTTCTC	29	63,4
PA1433_H3	CAGGGCGCCATGGAGAATCCCGCAGC	28	60,9	GACCCGACGCGGTACAGTGTCCACAGGC	29	62,6
PA1433_A4	CAGGGCGCCATGCGCCGCAACTGC	25	61,0	GACCCGACGCGGTACAGGCATCCCGG	26	61,9
PA0285_B4	CAGGGCGCCATGCTGGTACGCCAATCCG	29	60,2	GACCCGACGCGGTACAGTCTCCGCAACG	29	61,3
PA2567_C4	CAGGGCGCCATGGTGCCGACGCCA	25	62,8	GACCCGACGCGGTACAGCGCCGAGC	25	62,3
PA3258_D4	CAGGGCGCCATGGTGAATACGGGGCCG	28	60,0	GACCCGACGCGGTACAGGCGGCTCG	25	58,9
PA4959_E4	CAGGGCGCCATGGCCATCGAAAAGAAAACCATC	33	62,2	GACCCGACGCGGTATTCGTCTCCCGAGG	29	61,2
PA4367_F4	CAGGGCGCCATGGACGCTACAAGGCCAAG	30	64,0	GACCCGACGCGGTATTCGAAGTGGTGACG	31	62,5
PA4367_G4	CAGGGCGCCATGCTGACCAAAACCGCTGTCG	30	63,4	GACCCGACGCGGTACAGGCCGTTCGC	26	62,1
PA5017_H4	CAGGGCGCCATGAAAAGTCATCCCGATGCC	30	61,3	GACCCGACGCGGTACAGTGCAGGGTGCG	27	58,9
PA1181_A5	CAGGGCGCCATGGAGCGCAGCCGCC	25	63,5	GACCCGACGCGGTACGCCAACTCTGG	28	59,5
PA5487_B5	CAGGGCGCCATGAGTCGCGACGACGTCC	28	62,9	GACCCGACGCGGTACAGGCCACTTCCAGG	28	59,5
PA1120_C5	CAGGGCGCCATGACCTTGCGCGCC	24	60,4	GACCCGACGCGGTACAGCGCAGCAGGC	26	60,9
PA1120_D5	CAGGGCGCCATGTCACGGCGGCTGG	25	61,9	GACCCGACGCGGTCACTTGGGTGGAGCC	28	64,0
PA0169_E5	CAGGGCGCCATGGTGCGGTGGAGCG	26	61,8	GACCCGACGCGGTACAGCGCGTGGAG	26	62,3
PA0847_F5	CAGGGCGCCATGAGCCGTTTCGACCGAG	28	62,5	GACCCGACGCGGTACAGGCTTCTCGCC	27	59,8
PA0847_G5	CAGGGCGCCATGGAGCTGTGGGTCTGCG	29	63,7	GACCCGACGCGGTACAGGCTGGCGCC	25	60,6
PA4396_H5	CAGGGCGCCATGCCCGCGGTGTGG	25	64,0	GACCCGACGCGGTACAGCGAGTGCCG	26	60,7
PA3177_A6	CAGGGCGCCATGGCTCTCCCAAGCAGTC	29	61,9	GACCCGACGCGGTACAGGCGCAGCGC	25	62,3
PA0290_B6	CAGGGCGCCATGGACGACCTACCCGGC	27	63,2	GACCCGACGCGGTACGCCACGACGATG	28	63,1
PA3702_C6	CAGGGCGCCATGCACAACCTCATGAGAGCA	31	62,2	GACCCGACGCGGTACGCCCGCGG	24	61,8
PA2870_D6	CAGGGCGCCATGCGCCAGGCCAGC	25	62,9	GACCCGACGCGGTACAGAACGCCCGGTG	27	63,1
PA2870_E6	CAGGGCGCCATGGACTTCAAGCGCGCG	27	60,6	GACCCGACGCGGTACAGGCCGGCGC	24	61,8
PA4929_F6	CAGGGCGCCATGGCGCCGAGTTTC	26	60,2	GACCCGACGCGGTCACTGGTGCGGTATTC	30	63,3
PA4929_G6	CAGGGCGCCATGTGCGCATCCAGGATC	28	63,7	GACCCGACGCGGTACAGGAGATACGTTCCGG	30	63,0
PA1851_H6	CAGGGCGCCATGCTTGCACGCGACGG	26	61,3	GACCCGACGCGGTCAACGCCGCTGC	25	63,9
PA1851_A7	CAGGGCGCCATGGACTACGCCAGCGC	27	60,3	GACCCGACGCGGTATTGAGGCGGTCTG	28	62,5
PA0338_B7	CAGGGCGCCATGGTGCGGCGTTACG	26	58,4	GACCCGACGCGGTACAGCAACAGGCCACG	28	60,8
PA4332_C7	CAGGGCGCCATGCGGCAACGCATGC	25	58,8	GACCCGACGCGGTACAGGCACTGGTGACCTC	30	63,6
PA1107_D7	CAGGGCGCCATGCAGTGCGCAACCTG	27	62,2	GACCCGACGCGGTACCCGAGGCTTTCC	28	62,7
PA4843_E7	CAGGGCGCCATGATACCGAGCACGATGAC	30	62,4	GACCCGACGCGGTACCCGGCGTCCG	25	60,0
PA3343_F7	CAGGGCGCCATGGTGTGCGTGACAGAAAG	31	62,0	GACCCGACGCGGTACAGGTTTACGCGC	27	59,3
PA3343_G7	CAGGGCGCCATGAAGTCGCGCGAGC	25	61,4	GACCCGACGCGGTACAGGCCACACCTG	27	63,8

(continua)

(conclusão)

Tabela A2 - Lista de primers para amplificação das sequências alvo para clonagem pela técnica LIC.

Construção	Primers diretos (extensão LIC: 5'-CAGGGCGCCA-3')	#pb	T_m (°C)	Primers reversos (extensão LIC: 5'-GACCCGACGCGG-3')	#pb	T_m (°C)
PA2960_H7	CAGGGCGCCATGAGTTTGCCACCCAATCTGG	31	63,9	GACCCGACGCGGTCACATCGTGTGGGTCG	29	60,3
PA4608_A8	CAGGGCGCCATGAGTGACCAGCACGACGAAC	31	62,7	GACCCGACGCGGTCAGTCGTCTGGGCG	28	64,0
PA2989_B8	CAGGGCGCCATGGCTGCGACGTTTGGC	27	61,6	GACCCGACGCGGTCAGGTCGAAGAAGTTGG	31	61,7
PA0012_C8	CAGGGCGCCATGGATAATCAACGACAATACCCACG	35	62,0	GACCCGACGCGGTCAGGCTTCGCTGAGAAAG	31	64,0
PA3542_D8	CAGGGCGCCATGAATACAGCCGTCAACGTCAAC	33	61,4	GACCCGACGCGGTCACCGGACCCGG	25	59,7
PA3542_E8	CAGGGCGCCATGAACCAGATGTACAACCTGTACTTCG	37	62,1	GACCCGACGCGGTCAGCGAGCGGTGG	26	60,7
PA2799_F8	CAGGGCGCCATGCAGCCATCGAGCACAAC	30	63,3	GACCCGACGCGGTCAGTGGATCGCGACG	28	60,8
PA3353_G8	CAGGGCGCCATGGTGCTATCATTGAGGCATTTCG	33	62,2	GACCCGACGCGGTCAGAACAGTTCGTCTTTCTCG	34	61,2
PA4108_H8	CAGGGCGCCATGGTGCTGAAACGCATTCCG	30	63,3	GACCCGACGCGGTCACGGCTTCCGGTTC	29	63,7
PA2572_A9	CAGGGCGCCATGAACGATAGCGCACCTCC	29	60,1	GACCCGACGCGGTCAGGTCGTGACTCCG	29	61,3
PA4781_B9	CAGGGCGCCATGGAGAGCATGCTGGACAGG	30	61,0	GACCCGACGCGGTCAGGCTGGTGCG	26	58,8

Fonte: Elaborada pela autora.

Tabela A3 - Composição dos oito grupos de “iscas” formados. Na coluna “Descrição”, são dados os critérios de formação de cada grupo – as proteínas a que se faz referência nessa coluna são as proteínas que originaram as construções que compõe o grupo em questão.

Grupo	Componentes	Arquitetura	Descrição
11	pBT-PA0707_A1		Todas as proteínas do grupo contêm domínio EAL.
	pBT-PA2133_E1		
	pBT- (tnpM) _F1		
	pBT-PA3947_A2		
	pBT- (pvrR) _B2		
12	pBT-PA3825_B1		Todas as proteínas do grupo contêm domínio EAL e domínio CSS-motif, com exceção de pBT-PA3353_G8, que contém domínio PilZ.
	pBT-PA3825_C1		
	pBT-PA3825_D1		
	pBT-PA2200_G1		
	pBT-PA2200_H1		
	pBT-PA3353_G8		
13	pBT-PA3311_C2		Todas as proteínas do grupo contêm os domínios GGDEF e EAL.
	pBT-PA3311_D2		
	pBT-PA1727_B3		
	pBT-PA4601_F2		
	pBT-PA4367_F4		
	pBT-PA4367_G4		
14	pBT-PA0575_G3		Todas as proteínas do grupo contêm os domínios GGDEF e EAL e domínio PAS.
	pBT-PA0285_B4		
	pBT-PA4959_E4		
	pBT-PA5017_H4		
	pBT-PA1181_A5		
	pBT-PA2072_D3		
15	pBT-PA5487_B5		Todas as proteínas do grupo contêm domínio GGDEF.
	pBT-PA0169_E5		
	pBT-PA1107_D7		
	pBT-PA3343_F7		
	pBT-PA3343_G7		
16	pBT-PA1120_C5		Todas as proteínas do grupo contêm domínio GGDEF acompanhado de diferentes domínios N-terminais.
	pBT-PA1120_D5		
	pBT-PA3702_C6		
	pBT-PA4929_F6		
	pBT-PA4929_G6		
	pBT-PA0338_B7		
17	pBT-PA4608_A8		Todas as proteínas do grupo contêm domínio PilZ.
	pBT-PA2989_B8		
	pBT-PA0012_C8		
	pBT-PA3542_D8		
	pBT-PA3542_E8		
	pBT-PA2799_F8		
18	pBT-PA0861_H2		Grupo diversificado. Tanto pBT-PA4108_H8 quanto pBT-PA2572_A9 contêm domínio HD-GYP.
	pBT-PA1433_H3		
	pBT-PA2870_D6		
	pBT-PA1851_H6		
	pBT-PA4108_H8		
	pBT-PA2572_A9		

Fonte: Elaborada pela autora

Tabela A4 - Composição dos oito grupos de “presas” formados. Na coluna “Descrição”, são dados os critérios de formação de cada grupo – as proteínas a que se faz referência nessa coluna são as proteínas que originaram as construções que compõe o grupo em questão.

Grupo	Componentes	Arquitetura	Descrição
P1	pTRG-PA0707_A1		Todas as proteínas do grupo contêm domínio EAL.
	pTRG-PA2133_E1		
	pTRG- (tnpM) _F1		
	pTRG-PA3825_B1		
	pTRG-PA3825_D1		
	pTRG-PA2200_G1		
P2	pTRG-PA3311_C2		Todas as proteínas do grupo contêm os domínios GGDEF e EAL.
	pTRG-PA3311_D2		
	pTRG-PA1727_B3		
	pTRG-PA5295_C3		
	pTRG-PA4367_F4		
	pTRG-PA4367_G4		
P3	pTRG-PA5442_E2		Todas as proteínas do grupo contêm os domínios GGDEF e EAL e domínio PAS.
	pTRG-PA0861_H2		
	pTRG-PA0861_A3		
	pTRG-PA0575_G3		
	pTRG-PA0285_B4		
	pTRG-PA4959_E4		
P4	pTRG-PA4601_F2		Todas as proteínas do grupo contêm os domínios GGDEF e EAL acompanhados de diferentes domínios N-terminais, com exceção de pTRG-PA4601_F2.
	pTRG-PA2072_D3		
	pTRG-PA5017_H4		
	pTRG-PA1181_A5		
	pTRG-PA5487_B5		
P5	pTRG-PA0169_E5		Todas as proteínas do grupo contêm domínio GGDEF.
	pTRG-PA1107_D7		
	pTRG-PA3343_F7		
	pTRG-PA3343_G7		
P6	pTRG-PA0847_F5		Todas as proteínas do grupo contêm domínio GGDEF acompanhado de diferentes domínios comumente envolvidos em sinalização.
	pTRG-PA0847_G5		
	pTRG-PA4396_H5		
	pTRG-PA4929_F6		
	pTRG-PA4929_G6		
	pTRG-PA0338_B7		
P7	pTRG-PA2960_H7		Todas as proteínas do grupo contêm domínio PilZ.
	pTRG-PA2989_B8		
	pTRG-PA0012_C8		
	pTRG-PA3542_D8		
	pTRG-PA3542_E8		
P8	pTRG-PA3947_A2		Grupo diversificado. Tanto pTRG-PA3947_A2 quanto pTRG-(pvrR)_B2 contêm domínio EAL; pTRG-PA2572_A9 e pTRG-PA4781_B9 contêm domínio HD-GYP e pTRG-PA2799_F8 e pTRG-PA3353_G8 contêm domínio PilZ. As quatro primeiras contêm ainda domínio REC.
	pTRG- (pvrR) _B2		
	pTRG-PA2572_A9		
	pTRG-PA4781_B9		
	pTRG-PA2799_F8		
	pTRG-PA3353_G8		

Fonte: Elaborada pela autora.