UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE QUÍMICA

Programa de Pós-Graduação em Ciências Biológicas
(Bioquímica)

Cristiane Rodrigues Guzzo

**Estudo Estrutural e Funcional das proteínas PilZ e YaeQ do
fitopatógeno *Xanthomonas axonopodis* pv citri.**

São Paulo

Data de Depósito na CPG:

19/01/2010

Cristiane Rodrigues Guzzo

**Estudo Estrutural e Funcional das proteínas PilZ e YaeQ do
fitopatógeno *Xanthomonas axonopodis* pv citri.**

*Tese apresentada ao Instituto de
Química da Universidade de São
Paulo para obtenção do Título de
Doutor em Ciências (Bioquímica)*

Orientador: Prof. Dr. Shaker Chuck Farah

São Paulo
2010

Cristiane Rodrigues Guzzo

Estudo estrutural e funcional de proteínas do fitopatógeno *Xanthomonas axonopodis* pv citri: PilZ, envolvida na biogênese do pilus tipo IV e de YaeQ, pertencente a uma variação do motivo PD-(D/E)XK de endonucleases

*Tese apresentada ao Instituto de
Química da Universidade de São
Paulo para obtenção do Título de
Doutor em Ciências (Bioquímica)*

Aprovado em: _____

Banca Examinadora

Prof. Dr. _____

Instituição: _____

Assinatura: _____

Prof. Dr. _____

Instituição: _____

Assinatura: _____

Prof. Dr. _____

Instituição: _____

Assinatura: _____

Prof. Dr. _____

Instituição: _____

Assinatura: _____

Prof. Dr. _____

Instituição: _____

Assinatura: _____

DEDICATÓRIAS

Eu dedico este trabalho aos anjos da minha vida que sempre estão me guiando e me protegendo.

A Deus, por me iluminar, proteger e me dar forças quando elas não existem mais dentro de mim.

Ao meu marido, Guilherme Tozo de Carvalho, pelo companheirismo, amor e compreensão. Ele faz parte do meu corpo e da minha alma, e é a sustentação da minha vida e da minha alegria.

Aos meus pais, Henrique Guzzo Junior e Dionilde Rodrigues dos Santos Guzzo, pelo incentivo e pelo carinho. Eles são o meu guia para lidar com as adversidades da ciência e da vida. Eles compõem todos os alicerces do meu equilíbrio físico e mental.

Aos meus sogros, Maria Madalena Tozo de Carvalho e Sebastião de Almeida Carvalho, pela compreensão, ajuda e carinho. Eles são meus protetores e guardiões, sempre me ajudando com as dificuldades do dia a dia.

Aos meus irmãos, Fernando Rodrigues Guzzo e Fábio Rodrigues Guzzo, pelo amor e pelo companheirismo. Eles são o meu orgulho e os meus ídolos em que eu sempre me espelho.

Estas pessoas fizeram a total diferença na minha vida!!! Sem elas eu não teria conseguido chegar aonde eu cheguei.

AGRADECIMENTOS

Eu agradeço:

Em especial e eternamente ao meu orientador, Dr. Shaker Chuck Farah, por ter me dado a oportunidade de desenvolver este projeto, pela paciência, dedicação e amizade. Ele me ensinou a fazer ciência e a amá-la.

Ao Dr. Ronaldo A. P. Nagem, pela paciência, dedicação e compromisso. Ele me ensinou os conceitos e como determinar a estrutura tridimensional de proteínas por diferentes técnicas fazendo com que a cristalografia fosse muito mais interessante e contagiante.

Ao Dr. Roberto K. Salinas pela amizade e dedicação. Ele me ensinou os princípios de ressonância magnética nuclear de forma simples e objetiva, mostrando a preciosidade da técnica e sua aplicação em diferentes ensaios.

A Dra. Luci Ribeiro, Dra. Leonor Galvão de Botton, Dra. Pricila Hauk, Dra. Daniela Betton, Dra. Maria C. P. da Silva, Marcelo P. Carpes pela oportunidade de colaborarmos e passarmos excelentes momentos juntos.

Ao Prof. Dr. Elliot Kitajima por ter me acolhido e me ajudado imensamente nos ensaios de microscopia eletrônica na ESALQ de Piracicaba e ao Waldir Caldeira do Instituto de Biologia, USP, pela ajuda em microscopia eletrônica.

Aos meus queridos alunos de iniciação científica: Lucas Yanaguissawa, Clarissa Rocha, Giuliana Giuzio e Pamela Zucker, por terem me dado a oportunidade de passar momentos inesquecíveis e engraçados.

Aos membros do laboratório, passados e presentes, por toda ajuda nas discussões científicas e pela amizade que guardo dentro da minha memória, como bons momentos que vivi na minha iniciação científica e durante o meu doutorado: Dr. Marcos Alegria, Thaís Souza, Dra. Adriana Paulucci, Dr. Luis Holtzhauzen, Dr. Eric Vidsiunas, Ângela Katsuyama, Dr. Fernando Corrêa, Claudia Pereira Rosa, Dra. Cássia Docena, D. Cleusa, Alessandra Paiva, Diorge Paulo de Souza,

Maxuel Andrade, Paola Alejandra Cappelletti, Maycon Oliveira, Raphael Dias Teixeira, Ilda de Souza Costa e Camila Santos Shroeder.

A todos os membros do bloco zero em especial ao Dr. Robson Francisco, que sempre me ajudou com as dificuldades nas instalações dos programas de cristalografia, ao professor Dr. Ronaldo Bento Quaggio, pelas discussões científicas e pelo companheirismo, e aos seus ex-alunos Dra. Elaine Favaro, Dra. Juliana Luz, Dr. Carlos e aos membros do laboratório do Prof. Dr. Frederico José Gueiros Filho e da Prof. Dra. Carla Columbano de Oliveira, pelo bom conviver por todos estes anos.

A todos os membros do bloco zero em especial ao Dr. Robson Francisco, que sempre me ajudou com as dificuldades nas instalações dos programas de cristalografia, ao professor Dr. Ronaldo Bento Quaggio, pelas discussões científicas e pelo companheirismo, e aos seus ex-alunos Dra. Elaine Favaro, Dra. Juliana Luz, Dr. Carlos e aos membros do laboratório do Prof. Dr. Frederico José Gueiros Filho e da Prof. Dra. Carla Columbano de Oliveira, pelo bom conviver por todos estes anos.

A Fernanda Manso Prado por ter me ajudado nos experimentos de HPLC e espectrometria de massa de forma tão prestativa e acolhedora.

A Luci Deise Navarro por ter feito o sequenciamento de todos os genes que trabalhei durante o meu doutorado de forma ágil e eficaz.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa de doutorado. Ao Instituto de Química (IQ) da Universidade de São Paulo (USP) e ao Laboratório Nacional de Luz Síncrotron (LNLS) pelo espaço e pela infra-estrutura.

A todos a minha eterna gratidão e carinho!

“Siga tranqüilamente entre a inquietude e a pressa, lembrando-se de que há sempre paz no silêncio. Tanto quanto possível sem humilhar-se, mantenha-se em harmonia com todos que o cercam. Fale a sua verdade, clara e mansamente. Escute a verdade dos outros, pois eles também têm a sua própria história. Evite as pessoas agitadas e agressivas: elas afligem o nosso espírito. Não se compare aos demais, olhando as pessoas como superiores ou inferiores a você: isso o tornaria superficial e amargo. Viva intensamente os seus ideais e o que você já conseguiu realizar. Mantenha o interesse no seu trabalho, por mais humilde que seja, ele é um verdadeiro tesouro na continua mudança dos tempos. Seja prudente em tudo o que fizer, porque o mundo está cheio de armadilhas. Mas não fique cego para o bem que sempre existe. Em toda parte, a vida está cheia de heroísmo. Seja você mesmo. Sobretudo, não simule afeição e não transforme o amor numa brincadeira, pois, no meio de tanta aridez, ele é perene como a relva. Aceite, com carinho, o conselho dos mais velhos e seja compreensivo com os impulsos inovadores da juventude. Cultive a força do espírito e você estará preparado para enfrentar as surpresas da sorte adversa. Não se desespere com perigos imaginários: muitos temores têm sua origem no cansaço e na solidão. Ao lado de uma sadia disciplina conserve, para consigo mesmo, uma imensa bondade. Você é filho do universo, irmão das estrelas e árvores, você merece estar aqui e, mesmo se você não pode perceber, a terra e o universo vão cumprindo o seu destino. Procure, pois, estar em paz com Deus, seja qual for o nome que você lhe der. No meio do seu trabalho e nas aspirações, na fatigante jornada pela vida, conserve, no mais profundo do seu ser, a harmonia e a paz. Acima de toda mesquinhez, falsidade e desengano, o mundo ainda é bonito. Caminhe com cuidado, faça tudo para ser feliz e partilhe com os outros a sua felicidade”

(Este texto intitulado de desiderata foi encontrado na velha Igreja de Saint Paul, Baltimore, datado de 1692)

RESUMO

Guzzo, C.R. **Estudo Estrutural e Funcional das proteínas PilZ e YaeQ do fitopatógeno *Xanthomonas axonopodis* pv *citri***. 2010.

Número de páginas do trabalho 297. Tese Doutorado - Programa de Pós-Graduação em Ciências Biológicas (Bioquímica), Instituto de Química, Universidade de São Paulo, São Paulo.

O trabalho aqui desenvolvido teve como objeto o estudo estrutural e funcional de várias proteínas do fitopatógeno *Xanthomonas axonopodis* pv *citri* (*Xac*), dentre as quais se destacam as proteínas hipotéticas conservadas YaeQ e SufE, as proteínas RpfC, RpfF e RpfG envolvidas em *quorum sensing* e proteínas PilZ, FimX e PilB envolvidas na biogênese do pilus tipo IV.

Para o desenvolvimento deste trabalho foram utilizadas diferentes técnicas incluindo: clonagem, expressão, purificação, desnaturação térmica, cristalografia, difração de raios-X, RMN, ensaios de 2-híbrido, produção de nocautes, mutação sítio dirigida, Western- e Far- Western, entre outras.

Dentre os resultados mais importantes obtidos temos a determinação estrutural das proteínas YaeQ e PilZ pela técnica MAD. Em ambos os casos, as estruturas representaram topologias inéditas. Com base nos dados estruturais, mostramos que YaeQ pertence à família PD-(D/E)XK presente em endonucleases dependentes de magnésio, e a partir de ensaios funcionais obtivemos evidências que sugerem que YaeQ está envolvida em alguma via de reparo de DNA em *Xac*.

A estrutura tridimensional de PilZ revelou uma inesperada variedade estrutural dentro da família PilZ e mostrou de forma clara porque ortólogos não interagem com o segundo mensageiro bacteriano, c-diGMP. A cadeia principal de PilZ foi assinalada por RMN e a estrutura secundária de PilZ em solução é consistente com aquela determinada por cristalografia.

Duas proteínas que interagem com PilZ foram identificadas: PilB e FimX. Como PilZ, ambos exercem papéis na biogênese do pilus tipo IV (T4P). Mostramos que PilZ interage especificamente com o domínio EAL de FimX e que resíduos conservados na região do C-terminal de PilZ estão envolvidos na interação com PilB, mas não com FimX. Ensaios de mutação sítio dirigida mostraram que a Y22 de PilZ pode estar envolvida na regulação da interação de PilZ com FimX e com PilB. Apesar de PilZ não

interagir com c-diGMP seu parceiro, FimX, interage. PilZ consegue interagir com PilB ao mesmo tempo em que interage com FimX, formando um complexo ternário que é independente da interação de FimX com c-diGMP. Com base em todos estes resultados propusemos possíveis mecanismos de ação de PilZ e FimX no controle da biogênese do T4P.

Além dos resultados acima descritos, determinamos a estrutura de SufE e mostramos que esta aumenta a atividade cisteína dessulfarase de seu parceiro, SufS, em torno de 10 vezes, como ocorre com SufE-SufS de *E.coli*. Clonamos, expressamos, purificamos e fizemos ensaios de cristalização de algumas proteínas envolvidas no controle de *quorum sensing* em *Xac*. Tivemos êxito na cristalização do domínio HPT (histidina fosfotransferase) da proteína chave deste sistema, RpfC.

Palavras-chave: (PilZ, YaeQ, SufE, *quorum sensing*, difração de raios-X, pilus tipo IV)

ABSTRACT

Guzzo, C.R. **Structural and functional studies of PilZ and YaeQ from *Xanthomonas axonopodis* pv *citri* proteins.** 2010.

Number of pages 297. PhD Thesis – Graduate Program in Biochemistry. Instituto de Química, Universidade de São Paulo, São Paulo.

The aim of the project was to perform structural and functional studies of different *Xanthomonas axonopodis* pv *citri* (*Xac*) proteins including the hypothetical proteins YaeQ and SufE; RpfC, RpfF and RpfG involved in the *quorum sensing* and PilZ, FimX and PilB that play roles in type IV pilus (T4P) biogenesis. Several experimental techniques were employed including cloning, expression and purification of recombinant proteins, thermal denaturation, protein crystallography, X-ray diffraction, NMR, two-hybrid assays, Western- and Far-Western Blotting assays, site direct mutagenesis, and the production of *Xac* knockouts strains.

The most important results include the determination of the three-dimensional crystal structures of PilZ and YaeQ using the MAD technique. In both cases, the structures revealed new protein topologies. The comparison of the YaeQ structure with others deposited in public databases revealed that YaeQ proteins represent a new variation within the PD-(D/E)XK magnesium dependent endonucleases superfamily. Functional assays suggest that YaeQ may be involved in DNA repair in *Xac*.

The PilZ three-dimensional structure revealed an unexpected structural variation within the PilZ domain superfamily and showed why PilZ orthologs are not able to bind the important bacterial second messenger, c-diGMP. We assigned the PilZ main chain by NMR and used this information to demonstrate that the PilZ secondary structure in solution is consistent with the PilZ crystal structure. We identified two proteins that interact with PilZ: PilB and FimX. As with PilZ, both PilB and FimX are involved in T4P biogenesis. PilZ binds specifically to the EAL domain of FimX and the conserved residues located in the PilZ unstructured C-terminal region contribute to binding with PilB but not with FimX. Site direct mutagenesis studies showed that PilZ residue Y22 is necessary for its capability to interact with both PilB and FimX. Although PilZ does not bind c-diGMP, her partner, FimX, does. We present evidence that PilZ can bind simultaneously to FimX and PilB, forming a ternary complex that is

independent of c-diGMP. These results allow us to propose possible mechanisms by which PilZ and FimX control T4P biogenesis.

Other results obtained during this period include the resolution of the crystal structure of the SufE protein from *Xac* using the molecular replacement technique. We show that SufE induces a 10-fold increase in the cysteine desulfurase activity of SufS, similar to that observed for the SufE-SufS complex from *E. coli*. Several proteins involved in quorum sensing and c-di-GMP signaling were cloned, expressed and submitted to crystallization trials. Crystals of the HPT (histidine phosphotransferase) domain) of the RpfC sensor histidine kinase were obtained.

Keywords: (PilZ, YaeQ, SufE, quorum sensing, X-ray diffraction, type IV pilus)

LISTA DE ABREVIATURAS E SIGLAS

aa	Aminoácidos
AI	Auto indutor
ATP	Adenosina trifosfato
BMRB	<i>Biological Magnetic Resonance Bank</i>
CD	Dicroísmo Circular
c-diGMP	<i>bis(3'→5')GMP</i> cíclico
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
COSY	<i>Correlation spectroscopy</i>
DF	<i>Diffusible Factor</i>
DGC-A	Atividade diguanilato ciclase
DNA	ácido desoxirribonucleico
DO	Densidade Ótica
DOSY	<i>Diffusion Ordered spectroscopy</i>
DSF	<i>Diffusible Signal Factor</i>
DTT	Diliotreitol
EAL	Motivos conservados EAL, E-glutamato, A-alanina e L-leucina
EDTA	Ácido etilenodiaminotetraacético
FAPESP	Fundação de Amparo à Pesquisa do Estado de São Paulo
FOM	Figure of Merit
GGDEF	Motivos conservados GGDEF, G-Glicina, D-aspartato, E-glutamato e F-fenilalanina
HD-GYP	Motivos conservados HD-GYP, H-histidina, D-aspartato, G-glicina, P-prolina
HK	Histidina quinase
HPT	Histidina fosfotransferase
HSL	Homeserina lactona
HSQC	<i>Heteronuclear Single Quantum Coherence</i>
IPTG	Isopropil β-tiogalactosídeo
KDa	Kilo Daltons
LB	<i>Luria-Bertani</i>
LNLS	Laboratório Nacional de Luz Síncrotron
MAD	<i>Multiple wavelength anomalous dispersion</i>
MCP	<i>Methyl-accepting chemotaxis protein</i>
Δ	nocaute gênico
NOE	<i>Nuclear Overhauser Effect</i>
NOESY	<i>Nuclear Overhauser Effect Spectroscopy</i>
<i>P. aeruginosa</i>	<i>P. aeruginosa</i>
PCR	Reação em Cadeia da polimerase
PDB	<i>Protein Data Bank</i>
PDE-A	Atividade fosfodiesterase
pET	<i>Plasmid for Expression by T7 RNA polymerase</i>
ppm	Partes por milhão
pv	Patovar
Q.S.	<i>Quorum-sensing</i>
RMN	Ressonância magnética nuclear
rms	Raiz quadrada da media dos quadrados
RMSD	Raiz quadrada do quadrado do desvio padrão médio

RNA	ácido ribonucléico
<i>rpf</i>	<i>regulatory of pathogenicity factor</i>
RR	Regulador de Resposta
SDS	Dodecil sulfato de sódio
SelMet	Selênio-metionina
ssDNA	DNA simples fita
T4P	Pilus tipo IV
TAE	Tampão Tris-ácido acético
Tris	Tris (hidroximetil)aminometano
wt	Tipo selvagem
<i>Xac</i>	<i>Xanthomonas axonopodis</i> pv citri
<i>Xcc</i>	<i>Xanthomonas campestris</i> pv campestris

SUMÁRIO

ÍNDICE DE FIGURAS	IV
ÍNDICE DE TABELAS.....	VI
CAPÍTULO 1	
INTRODUÇÃO GERAL	20
1.1 INTRODUÇÃO HISTÓRICA	21
1.2 XANTHOMONAS AXONOPODIS PV CITRI: CAUSADORA DO CANCRO CÍTRICO.....	23
1.3 DETERMINAÇÃO DA ESTRUTURA TRIDIMENSIONAL DE PROTEÍNAS	28
OBJETIVOS	31
CAPÍTULO 2	
MATERIAIS E MÉTODOS	32
2. CLONAGEM DOS GENES	33
2.1 AMPLIFICAÇÃO POR PCR	33
2.2 REAÇÃO DE LIGAÇÃO.....	38
2.3 SEQUENCIAMENTO DOS CLONES (CONFIRMADOS ANTERIORMENTE POR DIGESTÃO DO PLASMÍDIO UTILIZANDO ENZIMAS DE RESTRIÇÃO ESPECÍFICAS).....	40
2.4 MUTAÇÃO SÍTIO DIRIGIDA	41
2.5 OBTENÇÃO DE CEPAS DE XAC NOCAUTES.....	43
2.5.1 <i>Obtenção de cepas nocautes de Xac</i>	43
2.5.1.1 Construção dos oligonucleotídeos para nocaute apolar de YaeQ, SufE, PilZ e SufR em Xac.	43
2.5.1.2 Transformação de Xac – Primeiro Evento de Recombinação.	44
2.5.1.3 Segundo Evento de Recombinação.....	44
2.5.1.4 <i>Análise e complementação de cepas nocautes de Xac</i>	45
2.6 TESTE DE INDUÇÃO E LARGA ESCALA.....	45
2.7 INCORPORAÇÃO DE SELÊNIO-METIONINA EM PROTEÍNAS HETERÓLOGAS.....	46
2.8 EXPRESSÃO DE PROTEÍNAS MARCADAS COM ¹⁵ N E/OU ¹³ C.	47
2.9 LISE CELULAR POR FRENCH PRESS	47
2.10 PURIFICAÇÃO DAS PROTEÍNAS DE INTERESSE.....	48
2.11 CRISTALIZAÇÃO E DETERMINAÇÃO DA ESTRUTURA TRIDIMENSIONAL	53
2.11.1 <i>Teste de cristalização das proteínas purificadas</i>	53
2.11.2 <i>Cristalização das proteínas purificadas: YaeQ, PilZ e os domínios RR e HPT da RpfC</i>	54
2.11.3 <i>Coleta de dados de cristalografia</i>	55
2.11.3.1 <i>Determinação dos comprimentos de onda para coleta de dados MAD utilizando o programa Blu-Ice</i> 56	56
2.11.3.1.1 <i>Inicialmente confirmamos a presença de Selênio na amostra pelo “Excitation scan”</i>	56
2.11.3.1.2 <i>Determinação dos parâmetros f’ e f’’ pelo “MAD scan”</i>	57
2.11.3.1.3 <i>Coleta de Dados MAD</i>	57
2.11.4 <i>Processamento dos dados cristalográficos e obtenção da estrutura tridimensional</i>	57
2.12 ENSAIOS FUNCIONAIS.....	58
2.12.1 <i>Ensaio de 2-híbrido</i>	58
2.12.1.1 Transformação da linhagem PJ69-4a com a isca de interesse.....	58
2.12.1.2 Teste de auto-ativação das células de levedura contendo a Isca.....	59
2.12.1.3 Transformação com a biblioteca genômica de Xac e confirmação dos clones positivos.....	59
2.12.1.4 Purificação do DNA plasmidial de levedura.....	60
2.12.1.5 Transformação em E. coli KC8 e sequenciamento dos clones positivos.....	60
2.12.2 <i>Ensaio de Beta-galactosidase de células de levedura</i>	61
2.12.3 <i>Produção de anticorpos</i>	62
2.12.4 <i>Western Blotting</i>	62
2.12.5 <i>Far-Western</i>	63
2.12.6 <i>RMN</i>	64
2.12.6.1 Ensaio de interação proteína-proteína por ressonância magnética nuclear.....	64
2.12.6.2 Assinalamento da cadeia principal	64
2.12.7 <i>Ensaio de atividades fosfodiesterase e diguanilato ciclase por HPLC</i>	65
2.12.8 <i>Inoculação de Xac em folhas de laranja</i>	65
2.12.9 <i>Ensaio de aderência de Xac em folhas de laranja</i>	66

2.12.10	Ensaio de motilidade de diferentes cepas de <i>Xac</i> em meio semi-sólido.....	66
2.12.11	Dicroísmo Circular e desnaturação térmica	67
2.12.12	Microscopia eletrônica de transmissão	67
2.12.13	Ensaio de resistência a rifampicina para medir a taxa de mutação espontânea.....	67
2.12.14	Gel shift.....	68
2.12.15	Atividade Cisteína dessulfarase	68
2.12.16	Ensaio de resistência a peróxido de hidrogênio e quelante de ferro.....	69

CAPÍTULO 3

INTERAÇÕES ENTRE PILZ, FIMX E PILB, PROTEÍNAS REGULADORAS DA BIOGÊNESE DO

PILUS TIPO IV 70

3.1 INTRODUÇÃO 71

3.1.1	PILUS TIPO IV	71
3.1.2	PILZ E PILUS TIPO IV	73
3.1.3	SINALIZAÇÃO POR C-DIGMP	74

3.2 OBJETIVOS.....77

3.3 RESULTADOS.....78

3.3.1	DETERMINAÇÃO DA ESTRUTURA DE PILZ _{XAC1133} PELA TÉCNICA MAD.	78
3.3.2	ANÁLISE ESTRUTURAL E FUNCIONAL DE PILZ E SEU PAPEL NO FUNCIONAMENTO DO T4P.	88
3.3.3	ENSAIOS DE MOTILIDADE DE DIFERENTES CEPAS DE <i>XAC</i>	123
3.3.4	MICROSCOPIA ELETRÔNICA DAS CÉLULAS NOCAUTES DE PILZ _{XAC1133} E FIMX _{XAC2398}	128
3.3.5	IMUNOLocalização de PILA _{XAC3240(32)_HIS}	132

3.4 DISCUSSÃO.....134

3.5 CONCLUSÕES 158

CAPÍTULO 4

YAEQ PERTENCE À SUPERFAMÍLIA DE PD-(D/E)XK ENDONUCLEASES.....160

4.1 INTRODUÇÃO 161

4.2 OBJETIVOS.....163

4.3 RESULTADOS.....163

4.3.1	CRISTALIZAÇÃO DE YAEQ	163
4.3.2	INCORPORAÇÃO DE SELÊNIO METIONINA NA PROTEÍNA YAEQ	167
4.3.3	COLETA DE DADOS DE DIFRAÇÃO DE RAIOS-X E DETERMINAÇÃO DAS FASES	168
4.3.4	ESTRUTURA TRIDIMENSIONAL DE YAEQ	172
4.3.5	RESÍDUOS ALTAMENTE CONSERVADOS NA FAMÍLIA DE YAEQ.....	174
4.3.6	YAEQ PERTENCE À FAMÍLIA PD-(D/E)XK DE NUCLEASES DEPENDENTES DE MAGNÉSIO	175
4.3.7	ENSAIOS FUNCIONAIS DE YAEQ.....	193
4.3.7.1	Mutantes de YaeQ _{XAC2396} : YaeQ _{D68A} e YaeQ _{D83A}	193
4.3.7.2	Cristalização dos Mutantes de YaeQ: YaeQ _{D68A} e YaeQ _{D83A} , e a determinação da estrutura de YaeQ _{D83A}	196
4.3.7.3	Atividade de Nuclease de YaeQ e dos Mutantes YaeQ _{D68A} e YaeQ _{D83A}	199
4.3.7.4	Ensaio de Ligação ao DNA Marcado com α - ³² P-dATP por “Gel Shift”	201
4.3.8	Ensaio de 2-híbrido usando YaeQ como Isca.....	203
4.3.9	As Proteínas Codificadas pelo mesmo Operon de YaeQ: XAC2397 e XAC2398	203
4.3.10	Nocaute apolar do gene de yaeQ em <i>Xac</i>	205
4.3.11	Mutantes de YaeQ apresentam uma maior taxa espontânea de metagênese	208
4.3.12	Determinação da estrutura de YaeQ em diferentes pHs.....	210

4.4 DISCUSSÃO.....212

4.5 CONCLUSÕES 222

CAPÍTULO 5

OUTROS ALVOS ESTRUTURAIS:

SUFE ESTÁ ENVOLVIDA NA BIOSÍNTESE DE CENTROS DE FERRO-ENXOFRE (FE-S)

& SISTEMA DE “QUORUM SENSING” EM XANTHOMONAS AXONOPODIS PV CITRI 224

5.1 SUFE: COMPONENTE DO SISTEMA CISTEÍNA DESSULFARASE ENVOLVIDA NA BIODISSÍNTese DE CENTROS DE FERRO-ENXOFRE	225
5.1.1 INTRODUÇÃO	225
5.1.2 OBJETIVOS	234
5.1.3 RESULTADOS E DISCUSSÃO	234
5.1.3.1 CRISTALIZAÇÃO DE SUFE	234
5.1.3.2 DETERMINAÇÃO DA ESTRUTURA DE SUFE	238
5.1.3.3 PROPOSTA DE POSSÍVEIS MECANISMOS DE AÇÃO DE SUFE	244
5.1.3.4 COMPARAÇÃO DA ESTRUTURA DE SUFE COM OUTRAS ESTRUTURAS DEPOSITADAS NO PDB.	247
5.1.3.5 SUFE AUMENTA ATIVIDADE CISTEÍNA DESSULFARASE DE SUFS	254
5.1.3.6 POSSÍVEL REGULAÇÃO GÊNICA DO SISTEMA SUF POR SUFR E OS POSSÍVEIS SCAFFOLDS PARA A CONSTRUÇÃO DE CENTROS DE FE-S EM XAC.	256
5.2 PROTEÍNAS DE XANTHOMONAS AXONOPODIS PV CITRI ENVOLVIDAS EM QUORUM-SENSING E SINALIZAÇÃO POR C-DIGMP	259
5.2.1 INTRODUÇÃO	259
5.2.1.1 QUORUM-SENSING MEDIADO POR DSF EM XANTHOMONAS.	259
5.2.2 RESULTADOS E DISCUSSÕES	264
5.2.2.1 PURIFICAÇÃO DAS PROTEÍNAS RPF, RPF, CMF E DOS DOMÍNIOS RR E HPT DA RPF	264
5.2.2.2 CRISTALIZAÇÃO DO DOMÍNIO HPT DA RPF.	266
5.2.3 CONCLUSÕES	271
CONCLUSÕES FINAIS	272
REFERÊNCIAS	275
LISTA DE ANEXOS	292
SÚMULA CURRICULAR.....	293

ÍNDICE DE FIGURAS

FIGURA 1: O CANCRO CÍTRICO.....	26
FIGURA 2: MAPA DO VETOR SUICIDA pNPTS138 ⁵⁷	43
FIGURA 3: PILUS TIPO IV EM <i>XAC</i>	71
FIGURA 4: MECANISMO DE PRODUÇÃO DE PILUS T4P.....	73
FIGURA 5: ESTRUTURA DO SEGUNDO MENSAGEIRO BACTERIANO C-DI-GMP ¹³⁴	76
FIGURA 6: ESPECTRO DE EXCITAÇÃO DE RAIOS-X EM QUE SE MEDE A FLUORESCÊNCIA DE RAIOS-X DO CRISTAL DE PILZ _{XAC1133} CONTENDO SELÊNIO-METININA.....	79
FIGURA 7: DETERMINAÇÃO DOS SÍTIOS DE SELÊNIO DO CRISTAL DE PILZ _{XAC1133}	83
FIGURA 8: RAMACHADRAN PLOT DA ESTRUTURA DE PILZ _{XAC1133}	86
FIGURA 9: ELUIÇÃO DE PILZ _{XAC1133} NA COLUNA DE GEL FILTRAÇÃO SUPERDEX 75.....	87
FIGURA 10: GÉIS DE ELETROFORESE DE SDS-PAGE DAS PROTEÍNAS PURIFICADAS.....	90
FIGURA 11: ENSAIO DE ATIVIDADE FOSFODIESTERASE E DIGUANILATO CICLASE.....	93
FIGURA 12: MOTILIDADE DE <i>XAC</i> EM MEIO SEMI-SÓLIDO.....	127
FIGURA 13: MICROSCOPIA ELETRÔNICA DE <i>XAC</i> EM DIFERENTES MEIOS.....	131
FIGURA 14: ENSAIO DE ADERÊNCIA EM FOLHAS DE LARANJA.....	132
FIGURA 15: ALINHAMENTO MÚLTIPLO ENTRE PILA DE <i>P. AERUGINOSA</i> E OS PILA DE <i>XAC</i> (XAC3241 E XAC3240) E GEL DE SDS-PAGE CONTENDO AMOSTRAS PURAS DE PILA _{XAC3240(32)}	133
FIGURA 16: ANÁLISE DA PRESENÇA DE RESÍDUOS ESSENCIAIS PARA A ATIVIDADE PDE-A DE DOMÍNIOS EAL E DE RESÍDUOS ENVOLVIDOS NA INTERAÇÃO DE DOMÍNIOS PILZ COM C-DGMP.....	137
FIGURA 17: ALINHAMENTO MÚLTIPLO DE MEMBROS ORTÓLOGOS DAS PROTEÍNAS ED21_28508 E ELI_07280.....	139
FIGURA 18: ANÁLISE FUNCIONAL DE TODOS OS DOMÍNIO GGDEF ENCONTRADOS NO GENOMA DE <i>XAC</i>	142
FIGURA 19: ANÁLISE FUNCIONAL DE TODOS OS DOMÍNIO EAL ENCONTRADOS NO GENOMA DE <i>XAC</i>	145
FIGURA 20: ALINHAMENTO MÚLTIPLO DAS SEQUÊNCIAS DE AA DOS DOMÍNIOS HD-GYP.....	147
FIGURA 21: ANÁLISE DO SÍTIO ATIVOS DOS DOMÍNIOS GGDEF E EAL DE FIMX.....	151
FIGURA 22: ESTRUTURA DO DOMÍNIO EAL DE FIMX DE <i>P. AERUGINOSA</i> EM COMPLEXO COM C-DIGMP.....	153
FIGURA 23: MECANISMO DE MONTAGEM DO T4P.....	157
FIGURA 24: OPERON QUE <i>yaeQ</i> PERTENCE EM ESPÉCIES DE <i>XANTHOMONAS</i> ¹⁸⁹	162
FIGURA 25: ESPECTRO DE MASSA DA PROTEÍNA YAEQ.....	168
FIGURA 26: ESPECTRO DE EMISSÃO DE FLUORESCÊNCIA DO CRISTAL DA PROTEÍNA YAEQ _{XAC2396} COM SELÊNIO-METIONINA.....	169
FIGURA 27: GRÁFICO QUE RELACIONA O COEFICIENTE DE CORRELAÇÃO COM O NÚMERO DE CICLOS PARA LOCALIZAR OS SÍTIOS DOS ÁTOMOS PESADOS, PROGRAMA SHELXD ⁶⁹	170
FIGURA 28: ESTRUTURA DE YAEQ COLORIDA SEGUNDO O VALOR DO FATOR-B.....	173
FIGURA 29: CROMATOGRAMA DA ELUIÇÃO DE YAEQ _{XAC2396} NA COLUNA DE GEL FILTRAÇÃO 75.....	174
FIGURA 30: SUPERPOSIÇÃO ESTRUCTURAL DE YAEQ COM OUTRAS PROTEÍNAS ESTRUCTURALMENTE SIMILARES ¹⁹⁵	177
FIGURA 31: MUTANTES DE YAEQ, YAEQ _{D68A} E YAEQ _{D83A}	194
FIGURA 32: PURIFICAÇÃO DOS MUTANTES DE YAEQ: YAEQ _{D68A} E YAEQ _{D83A}	195
FIGURA 33: WESTERN BLOTING DOS MUTANTES DE YAEQ: YAEQ _{D68A} E YAEQ _{D83A}	195
FIGURA 34: CRISTALIZAÇÃO DOS MUTANTES DE YAEQ.....	196
FIGURA 35: ESTRUTURA DE YAEQ _{D83A}	199
FIGURA 36: ATIVIDADE NUCLEÁSICA DE YAEQ E DOS MUTANTES YAEQ _{D68A} E YAEQ _{D83A} EM DIFERENTES TEMPOS DE REAÇÕES.....	200
FIGURA 37: ENSAIO DE “GEL SHIFT” PARA VERIFICAR A INTERAÇÃO DE YAEQ, YAEQ _{D68A} E YAEQ _{D83A} COM DNA, ssDNA E DNA TRATADO COM UV.....	202
FIGURA 38: YAEQ NÃO INTERAGE COM XAC2397 NEM COM FIMX _{XAC2398}	204
FIGURA 39: OBTENÇÃO DE CÉLULAS DE <i>XAC</i> NOCAUTE DE <i>yaeQ</i>	206
FIGURA 40: VERIFICAÇÃO DAS CÉLULAS NOCAUTES DE <i>yaeQ</i>	207
FIGURA 41: CRESCIMENTO DE CÉLULAS DE <i>XAC</i> E DE $\Delta yaeQ$ EM MEIO CONTENDO RIFAMPICINA.....	209
FIGURA 42: ESTRUTURA DE YAEQ EM DIFERENTES PHs.....	211
FIGURA 43: SUPERPOSIÇÃO DE YAEQ COM <i>BAMHI</i> EM COMPLEXO COM DNA.....	214
FIGURA 44: MECANISMO DE REPARO DE DNA POR PAREAMENTO INCORRETO.....	217
FIGURA 45: SOBREPOSIÇÃO ESTRUCTURAL DE PROTEÍNAS PERTENCENTES A FAMÍLIA PD-(D/E)XK, MUTH, VSR E <i>BAMHI</i>	221
FIGURA 46: ORGANIZAÇÃO GÊNICA DOS SISTEMAS <i>ISC</i> E <i>SUF</i> DE <i>E. COLI</i> E DO SISTEMA <i>NIF</i> DA BACTÉRIA FIXADORA DE NITROGÊNIO <i>A. VINELANDII</i> E SUA COMPARAÇÃO COM PROTEÍNAS HOMÓLOGAS ENCONTRADAS EM <i>XAC</i>	226

FIGURA 47: SISTEMA DE CISTEÍNA DESSULFARASE CATALISADA POR SufS-SufE.	228
FIGURA 48: MODELO DE BIOGÊNESE DE CENTROS DE Fe-S PELO SISTEMA <i>SUF</i> ^{248,234}	228
FIGURA 49: MODELO DO TRÁFICO DE CENTROS DE Fe-S QUE PODEM LEVAR A MATURAÇÃO DE IspG/H EM <i>E. COLI</i>	230
FIGURA 50: MODELO DE BIOSÍNTESE DE CENTROS DE Fe-S EM BACTÉRIAS, MAQUINARIA ISC E Suf ²³⁴	231
FIGURA 51: ORGANIZAÇÃO GÊNICA DO <i>OPERON SUF</i> E GENES AO SEU REDOR EM <i>XAC</i>	233
FIGURA 52: ESTRUTURA DE SufE ¹⁶⁴	239
FIGURA 53: ALINHAMENTO MÚLTIPLO DE ORTÓLOGOS DA FAMÍLIA DE SufE.	241
FIGURA 54: OS AMINOÁCIDOS CONSERVADOS NA FAMÍLIA DE SufE SE LOCALIZAM EM UMA FACE DA ESTRUTURA E FORMAM UMA REDE DE INTERAÇÕES ENTRE ELES.	242
FIGURA 55: OS AMINOÁCIDOS CONSERVADOS NA FAMÍLIA DE SufE SE LOCALIZAM EM UMA FACE DA ESTRUTURA E FORMAM UMA REDE DE INTERAÇÕES ENTRE ELES.	243
FIGURA 56: POSSÍVEL MECANISMO DE AÇÃO DE SufE (MECANISMO 1).	246
FIGURA 57: ALINHAMENTO DAS SEQÜÊNCIAS DAS PROTEÍNAS COM SIMILARIDADE ESTRUTURAL COM SufE DE <i>XAC</i> ²⁷⁴	249
FIGURA 58: SOBREPOSIÇÃO DA ESTRUTURA DE SufE COM OUTRAS ESTRUTURAS DEPOSITADAS NO PDB COM SIMILARIDADE ESTRUTURAL ²⁷⁴	251
FIGURA 59: SOBREPOSIÇÃO ESTRUTURAL DOS AMINOÁCIDOS PERTENCENTES AO SÍTIO ATIVO.	254
FIGURA 60: ALINHAMENTO MÚLTIPLO DAS SEQÜÊNCIAS PRIMÁRIAS DAS CISTEÍNAS DESSULFARASES ENCONTRADAS EM <i>E. COLI</i> : SufS, CsdA e IscS COM SufS DE <i>XAC</i> ¹⁶⁴	255
FIGURA 61: ATIVIDADE CISTEÍNA DESSULFARASE DE SufS NA PRESENÇA DE SufE.	256
FIGURA 62: ENSAIO DE PATOGENICIDADE EM PLANTA DE CEPAS DE <i>XAC</i> QUE CONTEM O GENE DE <i>SufR</i> TRUNCADO.	258
FIGURA 63: PURIFICAÇÃO DE SufR.	258
FIGURA 64: SISTEMAS DE Q.S ²⁹⁷	260
FIGURA 65: MODELO DO FUNCIONAMENTO DO SISTEMA <i>RPF</i> POR DSF NA REGULAÇÃO DA PRODUÇÃO DE FATORES DE PATOGENICIDADE EM <i>XANTHOMONADS</i>	262
FIGURA 66: ESQUEMA DOS DOMÍNIOS ENCONTRADOS NAS PROTEÍNAS DE <i>XAC</i> QUE POSSUEM O DOMÍNIO GGDEF.	263
FIGURA 67: GEL DE 15%SDS-PAGE REFERENTE A ÚLTIMA ETAPA DE PURIFICAÇÃO DAS PROTEÍNAS RpfF, RpfG, RpfC e CMF E DOS DOMÍNIOS RR e HPT DA RpfC.	265
FIGURA 68: CRISTAIS DO DOMÍNIO HPT DA RpfC.	268
FIGURA 69: CRISTAIS DO DOMÍNIO HPT DA RpfC SEM CAUDA DE HISTIDINA E INCORPORADA COM SELÊNIO METIONINA.	268
FIGURA 70: ALINHAMENTO DA SEQÜÊNCIA DA RpfC _{HIS-HPT(584-729)} COM A ARCB _{HPT} DE <i>E. COLI</i>	269
FIGURA 71: ESTRUTURAS DE DOMÍNIOS HPT DAS PROTEÍNAS ARCB, BARA e YojN DE <i>E. COLI</i> E Ypd1 DE <i>S. CEREVISIAE</i>	270

ÍNDICE DE TABELAS

TABELA 1: SEQUÊNCIA DOS OLIGONUCLEOTÍDEOS UTILIZADOS PARA AMPLIFICAR OS GENES EM ESTUDO.	34
TABELA 2: SEQUÊNCIA DE OLIGONUCLEOTÍDEOS PARA A AMPLIFICAÇÃO DAS REGIÕES PROMOTORAS DE XAC2934, XAC2935 E XAC2936, ALVOS DE LIGAÇÃO DE <i>SufR</i>	35
TABELA 3: SEQUÊNCIA DOS OLIGONUCLEOTÍDEOS UTILIZADOS PARA AMPLIFICAR OS FRAGMENTOS 1 E 2 PARA A CONSTRUÇÃO DE CEPAS NOCAUTES DE <i>YAEQ</i> , <i>SufE</i> , <i>PilZ</i> E <i>SufR</i>	36
TABELA 4: SEQUÊNCIA DOS OLIGONUCLEOTÍDEOS UTILIZADOS PARA AMPLIFICAR OS GENES DE INTERESSE AFIM DE COMPLEMENTAR CEPAS DE <i>XAC</i> NOCAUTES.	36
TABELA 5: A SEQUÊNCIA DE AMINOÁCIDOS DOS FRAGMENTOS DE DIFERENTES PROTEÍNAS ESTUDADAS NESTE PROJETO.	37
TABELA 6: A SEQUÊNCIA DE AMINOÁCIDOS CODIFICADAS PELA JANELA DE LEITURA QUE RESTA DOS GENES TRUNCADOS DAS CEPAS NOCAUTES DE <i>XAC</i>	38
TABELA 7: CLONAGEM E CONDIÇÕES DE EXPRESSÃO DE CADA GENE.	39
TABELA 8: SEQUÊNCIA DOS OLIGONUCLEOTÍDEOS PARA A OBTENÇÃO DE GENES MUTADOS DE <i>YAEQ</i> E <i>PilZ</i>	42
TABELA 9: RESUMO DAS CONDIÇÕES DE PURIFICAÇÃO DAS PROTEÍNAS.	48
TABELA 10: LINHAS DE RAIOS-X UTILIZADAS PARA COLETA DOS DADOS DE DIFRAÇÃO DOS CRISTAIS DE <i>SufE</i> , <i>PilZ</i> , HPT, <i>YAEQ</i> E <i>YAEQ_{D83A}</i>	56
TABELA 11: DADOS ESTATÍSTICOS DO PRIMEIRO MAPA DE DENSIDADE ELETRÔNICA.	84
TABELA 12: RESUMO DOS RESULTADOS MOSTRADOS NA FIGURA 13.	129
TABELA 13: RESULTADO DO ENSAIO DE 2-HÍBRIDO DOS DOMÍNIOS GGDEF DAS PROTEÍNAS XAC0258, XAC1420 D XAC1940.	143
TABELA 14: PREVISÃO DA ATIVIDADE DGC-A E/OU PDE-A DE TODAS AS PROTEÍNAS DE <i>XAC</i> QUE CONTEM DOMÍNIO GGDEF, EAL E HD-GYP.	148
TABELA 15: AS POSIÇÕES DOS SELÊNIO REFINADAS PELO PROGRAMA SHARP.	171
TABELA 16: DADOS ESTATÍSTICOS DO REFINAMENTO DAS FASES DO CRISTAL DE <i>YAEQ_{XAC2396}</i> -SELMET.	172
TABELA 17: DADOS ESTATÍSTICOS DO PROCESSAMENTO DOS DADOS DE DIFRAÇÃO DE RAIOS-X DO CRISTAL DE <i>YAEQ_{D83A}</i> . (COLETADOS NA LINHA MX1 DO LNLS).	197
TABELA 18: RESULTADOS DO ENSAIO DE 2-HÍBRIDO UTILIZANDO <i>YAEQ</i> COMO ISCA.	203
TABELA 19: DADOS ESTATÍSTICOS DO PROCESSAMENTO DE DADOS DE DIFRAÇÃO DE RAIOS-X DE <i>YAEQ</i> EM DIFERENTES pHs.	210
TABELA 20: ÁRVORE DE ESPÉCIES MOSTRANDO A OCORRÊNCIA DOS GENES <i>YAEQ</i> , <i>MUTH</i> E <i>VSR</i> EM DIFERENTES ESPÉCIES ¹²⁰	218
TABELA 21: DADOS DO REFINAMENTO DA ESTRUTURA DE <i>SufE</i>	238
TABELA 22: COMPARAÇÃO DA ESTRUTURA DE <i>SufE</i> COM OUTRAS ESTRUTURAS DEPOSITADAS NO PDB.	248
TABELA 23: DADOS ESTATÍSTICOS DO PROCESSAMENTO DOS DADOS DE DIFRAÇÃO DE RAIOS-X DO DOMÍNIO HPT DA RPF _C , (RPF _C _{HPT(584-725)}).	269

CAPÍTULO 1

Introdução Geral

Introdução Histórica

A diversidade de temas discutidos nesta tese merece uma explicação que somente faz sentido num contexto histórico. Nesta secção conto brevemente como chegamos a estudar as proteínas PilZ, FimX, PilB, YaeQ, SufE, RpfC e outras abordadas nesta tese.

Introdução a Biologia Estrutural durante a Iniciação Científica: Triagem de proteínas alvos para estudos estruturais.

Este projeto se originou na rede SMOLBNet (“Structural Molecular Biology Network” da FAPESP), que era composta por dezesseis laboratórios no estado de São Paulo e tinha como objetivo ensinar e disseminar a prática de biologia estrutural (cristalografia e ressonância magnética nuclear (RMN) de proteínas).

A Dra. Leonor Galvão de Botton, que nesta época era aluna de doutorado do laboratório, e a aluna de iniciação científica Ângela M. Katsuyama fizeram uma triagem de 35 proteínas alvos de *Xanthomonas axonopodis* pv citri (*Xac*) para estudo de proteômica estrutural¹. Estes foram selecionados *in silico* por serem proteínas de no máximo 300 aminoácidos (aa), hipotéticas conservadas, sem ortólogo com estrutura depositada, sem segmentos transmembranares e com mais de 1% de metioninas. Os Drs. Ana Paula Valente e Fábio Almeida, da Universidade Federal do Rio de Janeiro, haviam recentemente desenvolvido uma nova metodologia utilizando o antibiótico rifampicina, um inibidor da RNA polimerase bacteriana, que possibilitava a coleta de espectros ¹H-¹⁵N HSQC (Heteronuclear Single Quantum Coherence) de proteínas recombinantes expressas em *E. coli* antes que elas fossem purificadas. Análise destes espectros permitiam selecionar proteínas que melhor se adequavam a estudos estruturais ou por cristalografia de raios-X ou por RMN^{1,2}.

Usando esta técnica, em colaboração com Drs. Almeida e Valente, identificamos algumas proteínas como ótimas candidatas a estudos estruturais por serem solúveis, estáveis e apresentarem um espectro de HSQC com deslocamentos químicos bastante dispersos. Algumas destas proteínas, identificadas, pela Leonor e Ângela viraram objetos de análise no meu projeto de iniciação científica (2003-2004) e me deram continuidade durante o meu doutorado. Durante este período, determinamos a estrutura tridimensional da proteína hipotética conservada YaeQ_{XAC2396} por difração de raios-X, utilizando a técnica MAD (Multiple wavelength Anomalous Dispersion)^{1,2}. Estes resultados estão descritos no **Capítulo 4**. Neste período também adquirimos cristais da

proteína hipotética conservada YajQ_{XAC3671} que difratavam apenas até 4Å e por isso não obtivemos êxito na determinação desta estrutura. Outra proteína identificada nesta triagem inicial, SufE, foi cristalizada e sua estrutura resolvida por substituição molecular². Estes resultados estão descritos no **Capítulo 5.1**.

Determinação estrutural de PilZ. Após os dois primeiros anos do doutorado, ainda não tínhamos obtido êxito em experimentos desenhados para determinar a função de YaeQ. Por isso, começamos a estudar as proteínas codificadas pelo *operon* do qual o gene de *yaeQ* pertence (XAC2396 (*yaeQ*), XAC2397 e XAC2398 (*fimX*)), a fim de testar a hipótese de que existisse algum tipo de correlação entre eles. Ensaio de 2-híbridos envolvendo triagens de YaeQ como isca contra uma biblioteca de presas de *Xac* não forneceram resultados relevantes, mas os ensaios usando FimX como isca revelou que esta interage com PilZ_{XAC1133}, uma proteína envolvida no controle da biogênese do pilus tipo IV (T4P) em outras bactérias³. Esta interação entre FimX e PilZ era muito interessante por várias razões. O laboratório já estava interessado em estudar a sinalização por c-diGMP e a proteína PilZ pertence a uma superfamília de receptores ao c-diGMP⁴. Além disso, FimX possui os domínios potencialmente envolvidos na síntese e degradação de c-diGMP, GGDEF e EAL, respectivamente⁵. Por isso, começamos a estudar este sistema mais profundamente.

Determinamos a estrutura tridimensional de PilZ_{XAC1133} pela técnica MAD, sendo a primeira estrutura determinada por MAD usando dados coletados na linha MX2 do LNLS^{4,5}. Também assinalamos a cadeia principal de PilZ_{XAC1133} utilizando RMN, a fim de analisar a dinâmica e a estrutura secundária da proteína em solução. Finalmente, identificamos mais um novo parceiro de PilZ, PilB, envolvido diretamente na biogênese do T4P. Baseado na caracterização das interações de PilZ com PilB e com os domínios de FimX, propusemos um mecanismo de regulação da biogênese do T4P bacteriano envolvendo estas proteínas⁶. Estes resultados estão descritos mais detalhadamente no **Capítulo 3**.

Estudos de outras proteínas envolvidas em quorum sensing e sinalização por di-GMP cíclico: Cristalização do domínio HPT da proteína RpfC. Desde o começo do meu doutorado, um dos objetivos era tentar resolver a estrutura das proteínas do complexo RpfF-RpfC-RpfG que nosso grupo tinha identificado através de ensaios de dois-híbridos⁷. Também eram objetos de estudo proteínas de *Xac* que interagem com RpfG e que possuem o domínio GGDEF com atividade diguanilato ciclase (catalisa a conversão de duas moléculas de GTP em c-diGMP e pirofosfato). Todas as proteínas

e seus domínios propostos para estudo foram purificados e submetidos a ensaios de cristalização. Obtivemos apenas cristais do domínio HPT (histidina fosfotransferase) da RpfC que difrataram até 2.2Å. Como não há um ortólogo do domínio HPT com estrutura conhecida não pudemos resolver as fases pela técnica de substituição molecular e tentamos, portanto, resolver as fases via incorporação de selênio metionina (SelMet) e tentar coletar um conjunto de dados de difração de raios-X MAD desta amostra. No entanto, os cristais de HPT-SelMet eram ainda menores e demoravam, pelo menos seis meses, para começar a aparecer nas placas de cristalização. Estes resultados estão descritos de forma resumida no **Capítulo 5.2**.

A tese é composta por 5 capítulos. O que segue neste **Capítulo 1** são secções sobre a biologia da bactéria *Xac* e uma breve discussão sobre as técnicas de determinação estrutural de proteínas. O **Capítulo 2** descreve os Materiais e Métodos e os **Capítulos 3, 4 e 5** mencionados acima possuem Introdução, Objetivos, Resultados, Discussão e Conclusões separados sobre os temas específicos tratados. Em alguns casos, as publicações relevantes foram anexados e mencionamos às figuras relevantes destes manuscritos no corpo do texto. Por fim, a tese finaliza com uma conclusão geral do trabalho de doutorado.

Trabalhos em colaboração com outros grupos de pesquisa. Além dos resultados descritos acima, trabalhei em colaboração com Dra. Pricila Hauk, na época aluna de doutorado do laboratório do Prof. Dr. Paulo Lee Ho, do Instituto do Butantã, na cristalização⁸ e na determinação da estrutura tridimensional por MAD da proteína mais abundante na superfície de *Leptospira interrogans*, LipL32⁹. Trabalhei também em colaboração com a aluna de doutorado Daniela Betton, do laboratório do Prof. Dr. Walter R. Terra na determinação da estrutura tridimensional das proteínas catepsina 2 e 3, cisteína proteinases que possuem especificidades diferentes ao substrato, encontradas no sistema digestivo de *Tenebrio molitor larvae* (resultados ainda não publicados).

1.1 *Xanthomonas axonopodis* pv *citri*: causadora do cancro cítrico

O papel da laranja na história do Brasil. A laranjeira é nativa da Ásia, mas a sua origem é motivo de controvérsias. A descrição de citros mais antiga aparece na literatura chinesa, por volta do ano 2000 antes de Cristo¹⁰. Aparentemente, a disseminação de lavouras de laranja ao redor do mundo partiu da Ásia e foi levada para o Norte da África, posteriormente para o Sul da Europa e foi trazida para as

Américas por volta de 1500 durante os (re)descobrimientos das Américas (pelos europeus). Atualmente, os pomares mais produtivos são encontrados em regiões com climas tropicais e sub-tropicais, destacando os países Brasil, Estados Unidos, países do Mediterrâneo, México, China e África do Sul¹¹, sendo, São Paulo e Flórida as principais regiões produtoras de laranja.

A laranja foi trazida para o Brasil nos barcos dos colonizadores portugueses, que tinham um abastecimento de vitamina C para evitar o escorbuto que dizimava a maioria das tripulações neste período. As primeiras mudas de laranjeiras e as técnicas utilizadas para o cultivo eram espanholas e foram trazidas pelos portugueses no período de 1530-1540¹⁰. Nos meados do século XIX o Brasil possuía uma grande variedade de laranjas adaptadas ao clima e ao solo brasileiro destacando-se a laranja Bahia, baiana ou "de umbigo", por ser muito sucosa e de excelente qualidade. Em 1873, os técnicos em citricultura de Riverside, na Califórnia, receberam três mudas de laranja Bahia, e estas se espalharam pelos EUA e outras partes do mundo com o nome de Washington Navel¹¹. Na década de 80 o Brasil se tornou o maior produtor mundial de laranja, sendo destinada a produção de suco concentrado. Só o Estado de São Paulo produzia 70% das laranjas e 98% do suco concentrado produzido no Brasil^{10,12}.

O papel da citricultura na economia brasileira: O Brasil tem mais de 1 milhão de hectares de terra com cultivo de plantas cítricas e produz, aproximadamente, um terço dos frutos cítricos mundiais, respondendo atualmente por nada menos que 80% do suco de laranja concentrado negociado mundialmente^{10,12}. A cultura de citros tem uma fundamental importância na economia brasileira, envolve mais de 420 mil empregos diretos e indiretos, é a principal atividade de 330 municípios e gera em torno de US\$ 5 bilhões de dólares em receita bruta anual. Estas informações revelam a importância da cadeia agroindustrial cítrica no Brasil¹³.

Apesar do grande potencial da citricultura Brasileira, o país vem perdendo um espaço significativo no comércio internacional, devido a problemas fitossanitários. A citricultura brasileira é alvo de uma grande variedade de doenças, tais como: Cancro Cítrico, Clorose Variegada dos Citros (CVC), Declínio, Gomose de Phytophthora, Greening, Leprose do Citros, Mancha da Alternária, Morte Súbita dos Citros (MSC), Pinta Preta, Podridão Floral dos Citros, Rubelose, Citrus Tristeza Vírus (CTV); além de inúmeras pragas como: Ácaro da Ferrugem, Bicho Furão, Cigarrinha, Minador dos Citros, Mosca Negra dos Citros, Moscas das Frutas e Ortézia. Estes problemas

causam grandes prejuízos às lavouras de citros no país, pondo em risco os benefícios econômicos e sociais gerados pela agricultura.

Sintomas do Cancro Cítrico e sua erradicação: O Cancro Cítrico, causado por *Xanthomonas axonopodis* pv *citri*⁴, causa prejuízos de aproximadamente R\$ 300 milhões de reais por ano além de ter sido responsável pela perda de quase quatro milhões de árvores e mudas desde 1997¹⁴. Por este motivo, *Xac* têm sido alvo de grande interesse científico, econômico e político¹⁵. O cancro cítrico foi inicialmente detectado no começo do século XIX na Índia ou Ásia e se espalhou para o Japão, África, Austrália, Nova Zelândia, Sul da América e Estados Unidos^{16,17}. A primeira constatação de cancro cítrico no Brasil ocorreu no município paulista de Presidente Prudente em 1957 e espalhou-se rapidamente no estado de São Paulo e no sul do estado de Minas Gerais. Posteriormente, foi considerado endêmico nos estados do Rio Grande do Sul, Santa Catarina, Mato Grosso e mais recentemente em Rondônia^{18,19}.

O grande problema do cancro cítrico é a sua difícil erradicação e a sua fácil e rápida disseminação. Todos os tecidos novos de citros são suscetíveis a *Xac*²⁰, que não necessitam de um vetor de transmissão específico, podendo se espalhar entre vários cultivares através do vento, da chuva e do próprio homem. *Xac* entra na planta através dos estômatos (**Figura 1a**)^{15,17}, por lesões nos tecidos causados normalmente pelo material de colheita ou pela larva minadora dos citros (*Phyllocnitis citrella*) (**Figura 1b**)^{15,17}.

Os primeiros sintomas do cancro cítrico se manifestam nas folhas, surgindo lesões necróticas arredondadas, salientes e de coloração amarronzada e amarelada ao redor, nos dois lados das folhas. Posteriormente estas lesões aparecem nas frutas e nos ramos (**Figura 1c-e**)^{15,17,21} diminuindo a quantidade e a qualidade das frutas e, portanto, inviabilizando comercialmente a planta.

Devido à gravidade da doença e os prejuízos causados por *Xac*²²⁻²⁴, sua erradicação tem tido considerável atenção na comunidade agrícola, política e científica. No entanto, não foi descoberta nenhuma forma eficaz para sua erradicação e o procedimento adotado para evitar uma possível epidemia em toda lavoura é a detecção precoce da doença e a erradicação de plantas contaminadas a um raio de 30 metros do foco de contaminação (**Figura 1f-g**).

Figura 1: O Cancro Cítrico.

a) Imagem de SEM, *Scanning Electron Microscopy*, do estômato de uma folha de laranjeira com bactérias de *Xac* entrando no canal do estômato¹⁷, **b)** Larva minadora dos citros (*Phyllocnistis citrella*) sobre uma folha de laranjeira, e o caminho formado pela larva contaminado com *Xac*. A larva se alimenta da planta ocasionando lesões nas folhas, galerias, que facilitam a colonização por *Xac*¹⁵. Lesões ocasionadas por *Xac* na folha, **c)**¹⁷, fruto, **d)**¹⁵, e caule, **e)**¹⁷. **f,g)** Vistas aéreas da erradicação de árvores de laranja em um raio de 30 metros em relação ao centro do foco de contaminação por *Xac*¹⁵.

Sequenciamento e anotação do genoma de *Xac*: *Xac* é um fitopatógeno Gram-negativo, baciforme, de respiração aeróbica²⁵. Em maio de 2002 o seu genoma foi publicado e apresentou um único cromossomo circular de 5.2 Mb e dois plasmídios adicionais de 34 e 65 Kb²⁵. Sua anotação revelou 4313 ORFs (Open Reading Frames) sendo, na época 62,8% com função assinalada e o restante sem função conhecida²⁵.

A maioria dos patógenos e fitopatógenos utilizam sistemas de secreção tipo III (T3SS) e/ou tipo IV (T4SS) durante o processo de infecção com o objetivo de injetarem macromoléculas, importantes para o desenvolvimento da doença, dentro da célula hospedeira²⁶⁻²⁸. O T3SS é responsável pela secreção de fatores de patogenicidade essenciais para a virulência de diferentes patógenos e fitopatógenos²⁶⁻³⁰. O T4SS também inclui sistemas envolvidos no processo de conjugação bacteriana e também no processo de secreção de proteínas e complexos proteína-DNA^{28,31,32}. *Xac* tem dois T4SS codificados por dois *loci* vir, um localizado no cromossomo e outro localizado no plasmídio, e também um T3SS codificado pelo locus *hrp*^{25,33}. Além dos T3SS e T4SS, os fitopatógenos utilizam várias outras estratégias moleculares para invadir e colonizar as células hospedeiras, como sistemas de secreção do tipo I, II, V e VI, a produção de adesinas, pili, fatores de sinalização bacteriana, receptores de fatores externos e a produção de proteínas envolvidas no processo de transdução de sinal, entre outras³⁴⁻³⁷.

As bactérias possuem domínios sinalizadores que percebem de forma direta ou indireta estímulos ambientais. Em muitos casos, estes domínios sensores estão associados a domínios HK (histidina quinase) com os domínios RR (Regulador de Resposta) para formar os chamados sistemas de dois componentes por serem módulos em vias que usam a transferência de grupos fosfato (histidina fosfato no casos de domínios HK e aspartil fosfato no caso dos domínios RR) para oscilar entre estados regulatórios que regulam a atividade dos domínios efetores, ocasionando mudanças fisiológicas na bactéria³⁸. Os domínios sensores/sinalizadores mais comumente encontrados associados aos domínios HK são: GAF, PAS, BluF, HNOB, KdpD, Cachê, CHASE, PASTA, CitAP, NIT, ETR1, MHYE, MHYT, MASE e Pup-like; e os domínios efetores comumente encontrados associados aos domínios pertencentes aos sistema de dois componentes são: ACyc, EAL, GGDEF, HD-GYP, MCP, PP2A, PP2C, PTP, OmpR e S_TKc^{39,40}. O genoma de *Xac* codifica 58 proteínas contendo domínios HK e 60 proteínas contendo domínios RR, sendo que 21 delas possuem tanto o domínio HK quanto o domínio RR⁴⁰. Algumas destas proteínas são objetos de

estudo desta tese: a) RpfC, proteína quimérica com 4 domínios, Sensor-HK-RR-HPT, que forma um sistema de dois componentes com a RpfG que contém dois domínios, RR-HDGY⁴¹, descritos no **Capítulo 5.2**, e b) FimX, composta por 4 domínios, RR-PAS-GGDEF-EAL, descrita no **Capítulo 3**.

O estudo da biologia de *Xac* através de ensaios funcionais e estruturais de seus componentes moleculares, guiadas por análises genômicas, além de poder ajudar a compreender melhor a fisiologia da bactéria e os mecanismos moleculares envolvidos na patogenicidade, pode ajudar a desvendar os mecanismos biológicos de ação de outras bactérias do gênero *Xanthomonas*, que causam infecções em mais de 124 famílias de plantas monocotiledôneas e 268 plantas dicotiledôneas^{15,17,21,37,42}. Muitas delas têm papel importante na sociedade e na agricultura mundial, como *Xanthomonas oryzae* pv *oryzae* que ataca arroz, *Xanthomonas axonopodis* pv *phadeoli* que ataca feijão e *Xanthomonas translucens* que ataca cereais e gramíneas^{15,17,21}. Desta forma, novos métodos de controle destas doenças poderão ser desenvolvidos.

1.2 Determinação da estrutura tridimensional de proteínas

Projetos Genomas: Os projetos de seqüenciamento de genomas criaram um vasto banco de dados de seqüências de DNA de vários organismos, mas não forneceram informações funcionais para muitos dos genes codificados por estes organismos. Algumas proteínas podem ter a sua função anotada a partir da similaridade em seqüência primária com outras proteínas com função conhecida. No entanto, uma grande porcentagem dos genes depositados em banco de dados públicos não possuem homólogos com função conhecida, tornando difícil o estudo destas proteínas.

Projetos pós-genômicas foram iniciados para entender a função desta gama enorme de seqüências de DNA e proteínas. Neste contexto, os projetos de genômica estrutural tem o âmbito de determinar a estrutura tridimensional das proteínas codificadas por estes genomas em larga escala e a partir da estrutura inferir sua função. Como é inviável determinar as estruturas de todas as proteínas de todos os organismos conhecidos, o objetivo destes programas é de pelo menos resolver a estrutura de um número mínimo de membros de cada família protéica (definida por homologia ou similaridade de seqüência primária) que permite a fiel modelagem de todos os outros membros da família⁴³.

Técnicas para determinação estrutural de proteínas: Existem várias técnicas para a determinação estrutural de proteínas, sendo que a difração de raios-X, RMN e microscopia eletrônica⁴⁴ estão entre as principais. Atualmente há mais de 56 mil estruturas depositadas no PDB, das quais 85,7% foram determinadas por difração de raios-X, 13,7% por RMN, 0,4% por microscopia eletrônica e 0,2% por outras técnicas, como difração de elétrons, modelagem molecular, entre outras⁴⁴.

Apesar da maioria das estruturas depositadas no PDB terem sido determinadas por cristalografia de proteínas, ainda é uma técnica que tem muitas limitações, como por exemplo a obtenção de cristais de proteína com uma boa organização para que tenha uma difração em alta resolução. Outro grande problema é trabalhar com proteínas de membrana que além de serem insolúveis e necessitarem de detergentes para serem solubilizadas, possuem regiões flexíveis e pequenas alças hidrofílicas características desta classe de proteínas. Estas características reduzem significativamente as chances de formar contatos cristalinos e conseqüentemente diminuem a taxa de sucesso no processo de cristalização⁴⁵.

Com os projetos de genômica estrutural, os kits de cristalização foram otimizados e os ensaios de cristalização (triagens de identificar condições iniciais que produzem cristais) podem ser feitas por robôs utilizando pequenas quantidades de proteína, testando milhares de condições. Os programas de computadores utilizados para determinar a estrutura da proteína a partir dos padrões de difração foram otimizados e conseqüentemente, o processamento dos dados podem ser feitos durante a coleta. Em casos favoráveis, no final da coleta dos dados, já se tem um modelo da estrutura da proteína.

A técnica de RMN, além de possibilitar a determinação estrutural da proteína, também pode fornecer muitas outras informações que a técnica de difração de raios-X não fornece. As informações mais importantes fornecidas por RMN são parâmetros que descrevem a dinâmica da proteína em solução e o mapeamento de resíduos envolvidos na interação com pequenas moléculas ou com outras proteínas. Além disso, RMN pode ser utilizada para testar o estado de enovelamento da proteína antes da proteína ter sido purificada, a fim de averiguar se é um bom alvo para estudo estrutural¹. Recentemente, a técnica foi empregada para resolver a estrutura de proteínas em células vivas incluindo células de *E. coli*⁴⁶, células humanas⁴⁷ e oócitos de *Xenopus laevis*⁴⁸. Modificações pós-traducionais também foram detectadas por RMN em células vivas⁴⁹.

O conhecimento da estrutura tridimensional em alta resolução é indispensável para o desenho racional de drogas específicas que são projetadas para atuar especificamente na proteína em estudo. Estes medicamentos desenhados diminuem os efeitos colaterais quando comparados com os remédios formulados tradicionalmente. Alguns exemplos de drogas comerciais que foram projetadas são i: A droga para tratamento da AIDS “Agenerase” e “Viracept”, desenhada a partir da estrutura da proteinase do HIV⁵⁰, ii) A droga “Releza” para gripe, desenhada a partir da estrutura da neuraminidase do vírus influenza⁵¹, iii) A droga “Gleevec” foi desenhada a partir do domínio quinase da proteína c-Abi e é utilizada no tratamento de câncer⁵². Outros exemplos incluem o “Viagra” e vários remédios para tratamento de hipertensão e problemas cardiovasculares^{53,45}.

Objetivos

O objetivo geral do doutorado era determinar a estrutura tridimensional de várias proteínas do fitopatógeno *Xac*, que eram de interesse do laboratório:

1. Estudo Estrutural e funcional de PilZ_{XAC1133}, suas interações com FimX_{XAC2398} e PilB_{XAC3239} e o papel destas proteínas na biogênese do pilus tipo IV.
2. Estudo estrutural e funcional de YaeQ_{XAC2396}.
3. Estudo funcional e estrutural de proteínas pertencentes ao sistema de biossíntese de centros de Fe-S. Em especial as proteínas SufE_{XAC2355}, SufS_{XAC2938} e SufR_{XAC2934}.
4. Determinação da estrutura do domínio HPT da RpfC.

A técnica empregada para a determinação estrutural destas proteínas era cristalografia e difração de raios-X. Outras técnicas como RMN, ensaios de 2-híbrido, produção de nocautes gênicos, Western- e Far-Western Blotting, e mutagênese sítio-dirigida foram empregadas para melhor entender as relações estrutura – função destas proteínas na biologia da *Xac*. Objetivos específicos para cada um destes projetos estão apresentados nos **Capítulos 3, 4 e 5**.

CAPÍTULO 2
Materiais e Métodos

2. Clonagem dos genes

2.1 Amplificação por PCR

A amplificação dos genes foram feitas por PCR a partir de DNA genômico de *Xac*. Os oligonucleotídeos mostrados na **Tabela 1** foram usados nas clonagens dos genes que foram subsequentemente expressos em *E. coli*. A **Tabela 2** mostra as seqüências dos oligonucleotídeos utilizados para amplificar regiões promotoras dos genes XAC2934, XAC2935 e XAC2936. A **Tabela 3** mostra as seqüências dos oligonucleotídeos para a construção de plamídios necessários para a obtenção de cepas de *Xac* nocautes e a **Tabela 4** apresenta as seqüências dos oligonucleotídeos utilizados para a construção de plasmídios para a complementação de cepas nocautes de *Xac*. Foi utilizado para reação de PCR 2.5µl de DMSO; 5.0µl de tampão da enzima (10x concentrado, Fermentas); 0,5 µL de *Pfu* DNA Polymerase (25 U/µl, Fermentas); 1 µL de 25 mM de dNTP_{mix}; 2µL de solução 10 pmol oligonucleotídeo sense (+); 2 µL de solução 10 pmol oligonucleotídeo anti-sense (-); 6 ng de DNA genômico de *X. citri*; água deionizada para um volume final de 50 µL; 2.5µl de 25mM Sulfato de magnésio. Utilizando-se o termociclador Applied Biosystems GeneAmp® 9700, foram ajustadas as seguintes temperaturas e tempos de reação: 95°C por 2 min. (1 ciclo); 95°C por 1 min., temperatura de anelamento I por 30 seg. 70°C por 1 -3 min (5 ciclos); 95°C por 1 min., temperatura de anelamento II por 30 seg., 70°C por 1-3 min. (30 ciclos); 70°C por 10 min. As temperaturas de anelamento de cada par de oligonucleotídeos estão mostrados nas **Tabelas 1-4**. A seqüência de aminoácidos de cada domínio estão mostrados na **Tabela 5**. A seqüência codificada pela janela de leitura que resta dos genes truncados nas cepas nocautes de $\Delta sufR$, $\Delta sufE$, $\Delta yaeQ$ e $\Delta pilZ$ estão mostrados na **Tabela 6**.

Tabela 1: Sequência dos oligonucleotídeos utilizados para amplificar os genes em estudo.

Temperaturas de anelamento utilizadas na amplificação assim como o tamanho do fragmento do PCR estão mostrados na tabela. Os sítios de restrição incorporados encontram-se discriminados em marrom (*Pvu* II), salmão (*Sma* I), laranja (*Xho* I) e violeta (*Nco* I), em verde escuro (*Nde* I), verde claro (*EcoR* I), em azul (*BamH* I), azul água (*Hind* III), rosa (*Bgl* II), verde água (*EcoRV*). XAC0258* é a proteína completa e XAC0258 é apenas o domínio GGDEF. Os oligonucleotídeos antisense de 2493* e 0350* foram utilizados para clonar os fragmentos para estudos estruturais.

XAC	Apelido	OLIGO	pb	Tm1/Tm2	Sequência dos Oligonucleotídeos
Clonagem em pET3a/pET28a (ensaio de expressão) e/ou em pOBD (Duplo Híbrido)					
3240 <i>pilA</i>	PilAA	3240 (+) 3240 (-)	437	61/66	5' TTTTCCATGGTTCATATGTTGGCCAACAAGGTGATTAAAGGC 3' 5' TTGGATCCCTCGAGCTAACGGCAGGATGCC 3'
3241 <i>pilA</i>	PilAB	3241 (+) 3241 (-)	437	62.9/67.9	5' TTTTCCATGGTTCATATGAAGAAGCAACAGGGTTTCACGC 3' 5' TTGGATCCCTCGAGTTACTGGCAGGTAGCTGGAGCG 3'
3240 ₃₂₋₁₄₆ <i>pilA</i> ₃₂	PilAA ₃₂	3240 (+) 3240 (-)	341	62/67	5' TTTTCCATGGTTCATATGCAGATCTACGTTGCCAAATCGC 3' 5' TTGGATCCCTCGAGCTAACGGCAGGATGCC 3'
3241 ₃₀₋₁₄₆ <i>pilA</i> ₃₀	PilAB ₃₀	3241 (+) 3241 (-)	347	65/68	5' TTTTCCATGGTTCATATGCAGAACTATGTTGCCAAGTCGCAGG 3' 5' TTGGATCCCTCGAGTTACTGGCAGGTAGCTGGAGCG 3'
1133 <i>pilZ</i>	PilZ	1133 (+) 1133 (-)	351	52/64,7	5' TTC TCC ATG GAC CAT ATG AGT GCA ATG AAT 3' 5' CG GGA TCC CTC GAG TTA CAT CGT ATG CGT CG 3'
2398 FimX	G6 inteiro	2398 (+) 2398 (-)	2067	58/62.2	5' CTAGGGAGCCACCATGGAAAAAGGCAAAGATC 3' 5' GCGGATCCCTCGAGCTTTTCAGCCGAAGTCG 3'
2398 ₂₆₇₋₆₈₉ GGDEF-EAL	G6	2398 (+) 2398 (-)	1266	52.6/67.3	5' TATACCATGGTTCATATGTTGCGCCAGCGCGAC 3' 5' GCGGATCCCTCGAGCTTTTCAGCCGAAGTCG 3'
2398 ₂₆₇₋₄₃₂ GGDEF	GG	GGDEF (+) GGDEF (-)	489	55/67.3	5' TATACCATGGTTCATATGTTGCGCCAGCGCGAC 3' 5'CGGATCCCTCGAGTCAGCCGGATCGTAGATGC 3'
2398 ₄₃₆₋₆₈₉ EAL	EAL	EAL (+) EAL (-)	759	55/67.3	5'GATCCCATGGTGCATATGCGGGCCGAGGAAGAAGC 3' 5' GCGGATCCCTCGAGCTTTTCAGCCGAAGTCG 3'
3239 PilB	PilB	PilB (+) PilB (-)	1733	65/68.4	5' TTTTCCATGGTTCATATGAATAGTGTCGTGACCGCAAATTTGG 3' 5' TTTCTCGAGAAAGCTTTAGTCCTTCGTACACGATTGATCTCGG3'
2396 <i>yaeQ</i>	5	YaeQ (+) YaeQ (-)	546	51.6/68.5	5' A ATA CAG CTG CCC GGG ATC CAG ATG GCC CTC AC 3' 5' CA CGT CTC GAG GCA TGC TCA TTC GGC CGG 3'
2397	2397	2397 (+) 2397 (-)	1179	58.3/65.3	5' GCGAATTCATATGAGCATGCGTTGCGACG 3' 5' CTATAAGCTTCTCGAGTTACGCCACGCTCCCCGC 3'
1303 <i>mutS</i>	MutS	1303 (+) 1303 (-)	2705	62/69	5' TTTCCCGGGCATATGGGTGTCGCAATCGCCG 3' 5' TTGGATCCCTCGAGTCAAAGCAACGCCTTCAACCG 3'
2405 <i>mutL</i>	MuL	2405 (+) 2405 (-)	1874	69.7/69.9	5' CGCGCGTGCCGCACTGGAGCCCGGCATATGGCG 3' 5' TTGGATCCCTCGAGTCAACGCCCCCTCAAAAACCAA 3'
2355 <i>sufE</i>	17	2355 (+) 2355 (-)	435	51.0/66.5	5' CTATCA CAG CTG GCC ATG GGA TGTTAGATGACCACTC 3' 5' G GGG CTC GAG TGC TCA CTG CTG CGC 3'
3671 <i>yajQ</i>	1	3671 (+) 3671 (-)	483	62.2/70	5' CCT ACC ATG GCG GCT GTT TTT CAA TTG CAG GAG C 3' 5' ATC ACC CTC GAG CAC TCA ATC GCG GAA GTT GTC 3'
2934 <i>sufR</i>	SufR	2934 (+) 2934 (-)	456	50/66.6	5' TATTCATGGTTCATATGTTGCGCGTCAACAACTC 3' 5' AAGGATCCCTCGAGGCCTACAGGTCAGGTGGTCG 3'
2117 <i>nifA</i> ^{like}	NifU	2117 (+) 2117 (-)	597	52/67.4	5' GGTACCATGGCCCATATGATCCAGATATCCGACAAAGC 3' 5' AAGGATCCCTCGAGTCAGGCCGCGCTGTGC 3'

0499 <i>erpA</i>		0499 (+) 0499 (-)	384	52/67.4	5' TTTTCCATGGGCCATATGAGTACGCTCGTTTCGCTTCC 3' 5' TGAATTCCTCGAGTCAGACGCTGAAGCTGCTCCCGC 3'
1619 <i>sufA</i>	SufA	1619 (+) 1619 (-)	336	50/66	5' CTG TTT CAG CAT ATG TCC ATG GCC GTC AGC 3' 5' TT GAA TTC CTC GAG GAA TCA GCG GTG CTC 3'
3230	3230	3230 (+) 3230 (-)	888	60/67	5' CTGACCATGGTTCATATGGGGTTATGCACTTCAAAGCC 3' 5' TTTCCCGGGGGATCCAGAATGCCTACGCGATCTG 3'
2938 <i>sufS</i>	SufS	2938 (+) 2938 (-)	1242	56.5/69.0	5'TTCGCCGGGCATATGATGAACGTGTCCGC 3' 5'ATGAATTCCTCGAGTCAGCCCAACAGTCGACG 3'
1877 HD-GYP	$\beta_{199-233}$	1877 (+) 1877 (-)	418	54.8/66.7	5' C GGC ACA CAT ATG TTC CTG GAG CGC ATG TCG 3' 5' GGA GAA TTC TCA ATC CAT CGT CCA GGC ATC CTT G 3'
0350 HD-GYP	$\alpha\text{-d}_{171-419}$	0350 (+) 0350 (-)	750	53.2/67	5' A CTT ACC ATG GTG CAT ATG CTC AAG ACC GCC GAC 3' 5' G CTC CTC GAG TGT GAA TTC CGC TCA GGG CAG C 3'
	$\alpha_{171-311}$	0350* (-)	426	53.2/65.7	5' AT GAA TTC TCA ATC CCA ACC GCG TTT GTA AGG CCG 3'
2493 HD-GYP	$\gamma\text{-d}_{146-328}$	2493 (+) 2493 (-)	554	56/67	5' GAA ACC ATG GGT CAT ATG CAG GAC GCC GAC ACC 3' 5' AG GAA TTC CTC GAG CAG CGC CTA CGC CAC TG 3'
	$\gamma_{146-288}$	2493* (-)	432	55.9/66.9	5' TT GAA TTC TCA GCT GTC GAG CTT GCG CAT GGC 3'
0258*	G1	0258* (+) 0258* (-)	1056	56,4/68	5' CCTAAGATCTAACATATGACTTGC CAC AAC GCC GCA GG 3' 5' TTT AAG CTT CTC GAG CGC TCA GCG GTG ACC CGA TG 3'
0258 ₁₈₂₋₃₃₉ GGDEF	G2	0258 (+) 0258 (-)	477	55/67,1	5' CACACCATG GAACATATG CGT ACC GAC GAG CTC ACC G 3' 5' TA GGA TCC CTC GAG TCA CTC CTG CAT GCA CAG CAC 3'
0424 ₃₈₁₋₅₄₂ GGDEF	G3	0424 (+) 0424 (-)	486	58,3/67,3	5' GAC GCC ATG GGC CAT ATG CGC GAA GAC CAA CTC 3' 5' AA GGA TCC CTC GAG TTC GCC GCC GGT CTA ATC 3'
1420 ₃₃₀₋₄₉₇ GGDEF	G4	1420 (+) 1420 (-)	507	52,6/67,3	5' TTA GCC ATG GAA CAT ATG CGC GAT GCG TTG ACC 3' 5' TT GGA TCC CTC GAG TCA GTC GGT CGC ATC GTA AG 3'
1940 ₈₁₄₋₉₇₂ GGDEF	G5	1940 (+) 1940 (-)	480	52,6/67,3	5' CTA TCCATGGAGCATATG CAC GAC GAA CTCACCGG 3' 5' TG GAA TTC CTCGAG TCACACGTGGTCGTAATCGACAC 3'
3768 ₄₇₆₋₇₁₀ MCP	MCP	3768 (+) 3768 (-)	2124	50/64	5' GATCCCATGGTTCATATGAAACAACAGGCCGCCAACC 3' 5' TTGGATCCCTCGAGTCAGGCGACGTGTTTTTCC 3'

Tabela 2: Sequência de oligonucleotídeos para a amplificação das regiões promotoras de XAC2934, XAC2935 e XAC2936, alvos de ligação de SufR.

Temperaturas de anelamento utilizadas na amplificação assim como o tamanho do fragmento do PCR estão mostrados na tabela.

Regiões promotoras de SufR

Gene	Oligo	pb	Tm1/Tm2	Sequência dos Oligonucleotídeos
proXAC2934	proXAC2934 (+) proXAC2934 (-)	287	67	5' GCGTAGCGGGGCGAGGGCGAACATGC 3' 5' CATGGGGAAGACGAAATCCGAAAGCGGACCG 3'
proXAC2935	proXAC2935 (+) proXAC2935 (-)	276	69	5' AGCGATGGAAGGTCCGTTGGCGATCACCG 3' 5' CATTGGGCTGCCTACAGGTCAGGTGGTCGC 3'
proXAC2936	proXAC2936 (+) proXAC2936 (-)	283	67	5' TGCAAGCAGGTGTTCCGCGAACTGCCG 3' 5' GGGATTTTCCATCAGATTCTTTAGCCCTCTCC 3'

Tabela 3: Sequência dos oligonucleotídeos utilizados para amplificar os fragmentos 1 e 2 para a construção de cepas nocautes de *yaeQ*, *sufE*, *pilZ* e *sufR*.

Temperaturas de anelamento utilizadas na amplificação estão mostrados na tabela. Os sítios de restrição incorporados encontram-se discriminados em marrom (*Pvu* II), salmão (*Sma* I), laranja (*Xho* I) e violeta (*Nco* I), em verde escuro (*Nde* I), verde claro (*EcoR* I), em azul (*BamH* I), azul água (*Hind* III), rosa (*Bgl* II), verde água (*EcoRV*). As enzimas de restrição utilizadas para digerir o fragmento proveniente da amplificação do PCR (em parênteses) assim como as enzimas utilizadas para digerir o vetor estão mostradas na última coluna.

Oligonucleotídeos para construção de cepas nocautes					
Nocaute	OLIGO	Tm1/Tm2		Sequência dos Oligonucleotídeos	pNPTS138
$\Delta yaeQ$ XAC2396	2396 (+) 1°	55.3/61.3	Fragmento 1	5' GCCACAGC AAGCTT GAACGGGAATTATTGA 3'	HindIII/BamHI (HindIII/ EcoRI)
	2396 (-) 1°			5' GC GAA TTC GAAGTGTGGCGGTGAGGGC 3'	
	2396 (+) 2°	58.9/63.9	Fragmento 2	5' G CAG GAA TTC CCG GCC GAA TGA GCA TG 3'	HindIII/BamHI (EcoRI)
	2396 (-) 2°			5' AT GGATCC AGCTGCCGCACCGGCC 3'	
$\Delta sufE$ XAC2355	2355 (+) 1°	59.4/63.4	Fragmento 1	5' GCTGCG AAGCTT CTGGGCCCCACCATC 3'	HindIII/BamHI (HindIII/ EcoRI)
	2355 (-) 1°			5' GC AGT GAA TTC AAG CGG GAA GGG GGA GG 3'	
	2355 (+) 2°	58.1/63.2	Fragmento 2	5' TG GAATTC CATCTGTGCGCGACCCGCAG 3'	HindIII/BamHI (EcoRI/BamHI)
	2355 (-) 2°			5' AT GGATCC CCAGAAATGGCGCGACAGTCC 3'	
$\Delta sufR$ XAC2934	2934 (+) 1°	62.0/68.0	Fragmento 1	5' TGTTG AAGCTT CCAGCTGCCGATCGACCGC 3'	HindIII/EcoRI (HindIII/BamHI)
	2934 (-) 1°			5' GGT GAATTC GGTGACGCCAACATGGGGAA 3'	
	2934 (+) 2°	62.0/67.0	Fragmento 2	5' TT GAATTC CGCGTCGCGACCCACCTGACCTGTAG 3'	HindIII/EcoRI (BamHI/EcoRI)
	2934 (-) 2°			5' CT GGATCC AGGTCACCTTGCTACGGGTACTTCCAGG 3'	
$\Delta pilZ$ XAC1133	1133 (+) 1°	62.0/67.0	Fragmento 1	5' CAC GATATC GACGGAGCCTATGCCATGACCGC 3'	EcoRV/BamHI (EcoRV/BglII)
	1133 (-) 1°			5' GA AGATCT TGCACTCATCGAGTTCCTAGAACGTGGC 3'	
	1133 (+) 2°	62.0/67.0	Fragmento 2	5' GA AGATCT TTCGACAAGCCGACGCATACGATGTAAGC 3'	EcoRV/BamHI (BglII/BamHI)
	1133 (-) 2°			5' TT GGATCC GCCCGTGCTCAGCTTGCCG 3'	

Tabela 4: Sequência dos oligonucleotídeos utilizados para amplificar os genes de interesse afim de complementar cepas de *Xac* nocautes.

Temperaturas de anelamento utilizadas na amplificação estão mostrados na tabela. Os sítios de restrição incorporados encontram-se discriminados em marrom (*Pvu* II), salmão (*Sma* I), laranja (*Xho* I) e violeta (*Nco* I), em verde escuro (*Nde* I), verde claro (*EcoR* I), em azul (*BamH* I), azul água (*Hind* III), rosa (*Bgl* II), verde água (*EcoRV*). As enzimas de restrição utilizadas para digerir o fragmento proveniente da amplificação do PCR (em parênteses) assim como as enzimas utilizadas para digerir o vetor estão mostradas na última coluna.

Complemenação	Tm1/Tm2	Nome da construção	Sequência do Primer	Enzimas de Restrição
PilZ _{XAC1133}	52.0/66.0	pUFR047_ <i>pilZ</i>	5' CAC GGATCC GACGGAGCCTATGCCATGACCGC 3' 5' CG GGATCC CTCGAG TTACATCGTATGCGTCG 3'	BamHI/ Sall (BamHI/ XhoI)
PilZ _{XAC1133} e PilZ _{mutantes}	50.0/52.5	pMR20_ <i>pilZ</i> e pMR20_ <i>pilZ</i> _{Δ107-117}	5' GA GGATCC ATACATATGAGTGCAATGAATG 3' 5' CG GAATTC TTACATCGTATGCGTCG 3'	BamHI/EcoRI (BamHI/EcoRI)

Tabela 5: A sequência de aminoácidos dos fragmentos de diferentes proteínas estudadas neste projeto.

XAC	SEQUÊNCIA	aa
2398 ²⁶⁷⁻⁶⁸⁹ Domínios GGDEF-EAL de FimX (G6)	MQRDQVTGLLNRPTFMVALEQAVAQAGRSEGGSGFLLIEPDHYARILPEFGLDSDADAL IAALAAHVASVLDSDVLAARFGEHSFALLNGNYARTHALAETVRNAFAQHVSFGTR SATVTVSIGGVQIGEKIASIGQVLNRGTEAVRTTAEELGGNAVSIYDPAADRAEEERIER WVEQLREALVGDGFLHYQPVNLQGEPELEYQAFLRLERNGEMMSPNAFMAIAEEH DLVTEIDRWVVARAIRQLGERQRAGHKTHLLVRIGPNSFSDPQMIDTIREQLAVYGVPG ERLWLQTPESKVFTHLRNAQQFLAAVSAMDCKVGLEQFGSGLDSFQLLAHFHPAFLK LDRGITGDIASARDSQEKIREITSRAQPAGILTMAEFVADAQSMSSFFSAGVDYVQGDF VAPTGPLMNYEFG	421
2398 ²⁶⁷⁻⁴³² Domínio GGDEF de FimX	LRQRDQVTGLLNRPTFMVALEQAVAQAGRSEGGSGFLLIEPDHYARILPEFGLDSDADAL LIAALAAHVASVLDSDVLAARFGEHSFALLNGNYARTHALAETVRNAFAQHVSFGTR SATVTVSIGGVQIGEKIASIGQVLNRGTEAVRTTAEELGGNAVSIYDPA	165
2398 ⁴³⁶⁻⁶⁸⁹ Domínio EAL de FimX	AAEEERIERWVEQLREALVGDGFLHYQPVNLQGEPELEYQAFLRLERNGEMMSPNAF MAIAEEHDLVTEIDRWVVARAIRQLGERQRAGHKTHLLVRIGPNSFSDPQMIDTIREQLA VYGVPGERLWLQTPESKVFTHLRNAQQFLAAVSAMDCKVGLEQFGSGLDSFQLLAHFH PAFLKLDRGITGDIASARDSQEKIREITSRAQPAGILTMAEFVADAQSMSSFFSAGVDYV QGDFVAPTGPLMNYEFG	253
3240 ³²⁻¹⁴⁶ <i>pilA</i> ₃₂	MQIYVAKSQVAAGLAEVSPGKTRYEILVNDSEGSTLTSASAIGLTATATSRCSISVLAP DANGAQAHAIIDCTLRGNPRVNGAKLTWARDANGIWTCTDLPLVDKDRFAPASCR	116
3241 ³⁰⁻¹⁴⁶ <i>pilA</i> ₃₀	MQNYVAKSQVTAGLAELSPGKTQYETLLNEGKGTTLTTADAVGLQATSTTNRCSAITV LAPDATTGEQTNALSCTLAGSPRIVGKKITWARDTNGVWTCSTDIATADKAKFAPATCQ	118
1877 Domínio HD-GYP de RpfG β ₁₉₉₋₂₃₃	MFLERMSHVAGLIAEQLGLSEEEVRIIEMAAPLHDMGKIAIPDAVLLKPGKLTDDMSV MRRHPRIGYELLSGSQNRFIQVGALIALRHHERYDGSQYDPDGLVGEAIPLEARIVAVAD VFDALLSPRPYKDAWTMD	136
0350 HD-GYP α ₁₇₁₋₃₁₁	MLKTADDYTYLHSMAGVAGLMSGLARQLGLPESQVVEAAMGGLLHDMGKAVTPLEILN KPGRLTNEEMEVMRKHPLDGHRLLVAGGVQNAVVEIALHHHEKMDGSGYPNQLVG ENISLMSRMGAVCDVYDAISSDRPYKRGW	142
0350 HD-GYP α -d ₁₇₁₋₄₁₉	MLKTADDYTYLHSMAGVAGLMSGLARQLGLPESQVVEAAMGGLLHDMGKAVTPLEIL NKPGRILTNEEMEVMRKHPLDGHRLLVAGGVQNAVVEIALHHHEKMDGSGYPNQLV GENISLMSRMGAVCDVYDAISSDRPYKRGWDPAESLKQMASWKGHFDPVVFQAFVRR LGIYPVGSVLRLESQKLAVVIEQTPDALLKPKVRVFYSAKLRSHVLVQDIDLSRPDCHDR IAQMESPAEWGF RDLEKLWLP	250
2493 HD-GYP γ ₁₄₆₋₂₈₈	MQDADTGAIHWRMAVYARVLAEAVGWSPQQAQTLQDAAPLHDAGKIAVPGSILQKPD ALDEHERSVVRQHPQAGHDLLRHGRGPVFRLLAAEIALHHHECWGDSGYPQGLSGDAI PESARIVAVADVFDALCGRRAYKAPWTVED	144
2493 HD-GYP γ -d ₁₄₆₋₃₂₈	MQDADTGAIHWRMAVYARVLAEAVGWSPQQAQTLQDAAPLHDAGKIAVPGSILQKPD ALDEHERSVVRQHPQAGHDLLRHGRGPVFRLLAAEIALHHHECWGDSGYPQGLSGDAI PESARIVAVADVFDALCGRRAYKAPWTVEDAMRKLDSMAGKQLEPRLVQHFHEALPQI LQIKDAWDSPVA	184
0258 ¹⁸²⁻³³⁹ GGDEF (G2)	MRTDELTLGLHNRRLALHLQECRRAGHVPGVGLILIDCDRFKRNDEFGLHAGDRVLRQ NVAHALRSVIAEGDLACRLGGDEFVLVAGQGTSELVHHIGECIVRAVEARLCNGQGEQ GCSVSVGSCMIDADAAWNWDWYTLADSALYEAKRQGGNVLCMQE	159
0424 ³⁸¹⁻⁵⁴² GGDEF (G3)	MREDQLTGALNRRGFEEFLQREAVRTQRSGQPLCVAMLDLDDFRRLNETHGHAGGD AALRHTVEVAKAVLRTTDAIARFGGEEFVLLLPDSTIFEASAIVRLQRALSQRSLLHDD VRVFISFSAGVALRGTDQDEVIRRADRAMYDAKAAGKNRVINAD	162
1420 ³³⁰⁻⁴⁹⁷ GGDEF (G4)	MRDALTGVFNRHQLGTALTHEFELASRQGWPLSIAFIDLDDFKRVNDAHGHVGVDEV LRFNAQTLQRMVRTSDLVARYGGEEFLVLPNTPADAAALMVIERILAETARTPMAQLQG GPLHITFSAGLATHGGAERQFDSIDHLLQAADSTLYRSKHRGRNRVTAYDATD	169
1940 ⁸¹⁴⁻⁹⁷² GGDEF G5	MHDELTGLLNRAGILALREMLLHAEHQGRPLAVVLIDLDHFKLVNDQHGHLGDAVL AGVGRMDTLMRGDDRIGRYGGEELLALLPGLNLDATHRLDALHRGICGDYPIDGGW LHVTCSVGVAWFQRGETLEQLLARADAALYRAKRSGRNRVDYDHV	160
3768 ⁴⁷⁶⁻⁷¹⁰ MCP	MEQQAANLEETAASMEELTATVKQNAEHARQANELARSAAGVASKGGEVVSQVVG MSGIETSSKKIAEHSVIDGIAFQTNILALNAAVEAARAGEQGRGFVAVASEVRSQAQRSS GAAKEIKDLIDASVGKVAEGSALVHKAGTTMTEIVASVNRVTDIMGQITAASIEQSSGIE QVNQTVIQMDETTQQNAALVEEAMAAARAMEKQSSTLTQLVSLFQLEPATAQHLEKH VA	236
1878 ⁵⁸⁴⁻⁶⁹⁹ HPT	MVRQTDAAIGKRGPATNGAVLDPVLDLDELAEELGMGEFEFEFIQQCLSDATAACVVEL G KAAADEDWGRLREHAHAVKGVSSNLGLVILAEQGGELMRLPEAQLRSEWRQRLEQL KTGLAQGREALAQARRKTEGGASNGDP	137

Tabela 6: A sequência de aminoácidos codificadas pela janela de leitura que resta dos genes truncados das cepas nocautes de *Xac*.

$\Delta yaeQ$	MALTATVRRAEFPAD
$\Delta sufE$	MTTSPFPLEFICDTARAQQ
$\Delta pilZ$	MSARSSDKPHTM
$\Delta sufR$	MLRVTEFRVATT

2.2 Reação de Ligação

Os genes amplificados foram purificados com o Kit de purificação de PCR da Qiagen. Tanto os vetores quanto os genes purificados foram digeridos com as enzimas adequadas (**Tabela 3, 4 e 7**) e nas temperaturas ideais de cada enzima durante 24 horas. Posteriormente, as enzimas foram inativadas a 85 °C por 20 minutos. Os produtos digeridos foram purificados através de eletroforese em gel de agarose utilizando o kit de Extração em gel da Qiagen.

Os fragmentos 1 e 2 de $\Delta pilZ$, $\Delta sufR$, $\Delta sufE$ e $\Delta yaeQ$ (**Tabela 3**) foram inicialmente digeridos com as enzimas adequadas e estas enzimas foram inativadas a 85 °C por 20 minutos. Os produtos digeridos foram purificados com o Kit de purificação de PCR da Qiagen. Estes fragmentos foram ligados da seguinte forma: 5 µL de tampão para ligase (10x concentrado, New England BioLabs); 5 µL de 10mM ATP; 2 µL de T4 DNA ligase (400.000 U/mL, New England BioLabs); 15 µL do fragmento 1; 15 µL do fragmento 2; água deionizada para um volume final de 50 µL. A reação foi encubada a 25 °C durante 4 horas e posteriormente encubada a 16 °C por 12 horas. O fragmento ligado foi purificado através de eletroforese em gel de agarose utilizando o kit de Extração em gel da Qiagen. Os fragmentos posteriormente foram digeridos com as enzimas adequadas e nas temperaturas ideais de cada enzima durante 24 horas. Posteriormente, as enzimas foram inativadas a 85 °C por 20 minutos. Os produtos digeridos foram purificados através de eletroforese em gel de agarose utilizando o kit de Extração em gel da Qiagen.

A ligação do inserto nos vetores corretamente digeridos foi feita da seguinte forma: 2 µL de tampão para ligase (10x concentrado, New England BioLabs); 2 µL de 10mM ATP; 1 µL de T4 DNA ligase (400.000 U/mL, New England BioLabs); 5 µL de inserto; 1 µL de vetor; água deionizada para um volume final de 20 µL. A reação foi encubada a 25 °C durante 4 horas e posteriormente encubada a 16 °C por 12 horas. Estas amostras foram utilizadas para transformar células de *E. coli* da cepa DH5α por

cloreto de cálcio segundo o método de Sambrook *et al.* (2001)⁵⁴. As colônias obtidas na transformação foram inoculadas em 3ml de 2xTY (16g/L Triton X-100 a; 10 g/L extrato de levedo; 5 g/L NaCl; pH 7,4) contendo 100µg/mL do antibiótico adequado. O DNA plasmidial das colônias foram purificadas utilizando o kit de extração de plasmídeo da Qiagen.

Tabela 7: Clonagem e condições de expressão de cada gene.

A tabela dispõe de todos os genes estudados que foram clonados tanto no vetor de expressão (pET3a e/ou pET28a) e/ou no vetor para ensaios de 2-híbrido (pOBD). As enzimas de restrição utilizadas para digerir o fragmento proveniente da amplificação do PCR (em parênteses) assim como as enzimas utilizadas para digerir o vetor estão mostradas. As condições de expressão das proteínas que se restringem a cepa de *E.coli* e a temperatura de indução estão discriminadas para cada proteína assim como a quantidade de aminotriazol necessária para cada gene nos ensaios de 2-híbrido. (-) insolúvel, (+) solúvel, (±) parcialmente solúvel e N.D não disponível.

XAC	Apelido	pET3a	pET28a	Solub.	Temperatura BI21(DE3)	pOBD	Aminotriazol
3240 <i>pilA</i>	PilAA	NdeI/BamHI (NdeI/BamHI)	NdeI/BamHI (NdeI/BamHI)	-	37°C RP	NcoI/XhoI (NcoI/Sall)	
3241 <i>pilA</i>	PilAB	NdeI/BamHI (NdeI/BamHI)		-	37°C RP	NcoI/XhoI (NcoI/Sall)	
3240 ₃₂₋₁₄₆ <i>pilA</i> ₃₂	PilAA ₃₂		NdeI/BamHI (NdeI/BamHI)	-	37°C STAR	NcoI/XhoI (NcoI/Sall)	
3241 ₃₀₋₁₄₆ <i>pilA</i> ₃₀	PilAB ₃₀		NdeI/BamHI (NdeI/BamHI)	N.D	Não	NcoI/XhoI (NcoI/Sall)	
PilZ mutantes					37°C DE3		5mM
1133 <i>pilZ</i>	PilZ	NdeI/BamHI (NdeI/BamHI)		+	37°C DE3	NcoI/XhoI (NcoI/Sall)	
2398 FimX	G6 inteiro		NcoI/BamHI (NcoI/BamHI)	+	28°C STAR		
2398 ₂₆₇₋₆₈₉ GGDEF- EAL	G6		NdeI/BamHI (NdeI/BamHI)	+	32°C RP	NcoI/XhoI (NcoI/Sall)	5mM
2398 ₂₆₇₋₄₃₂ GGDEF	GG	NdeI/BamHI (NdeI/BamHI)		+	37°C RIL	NcoI/XhoI (NcoI/Sall)	5mM
2398 ₄₃₆₋₆₈₉ EAL	EAL	NdeI/BamHI (NdeI/BamHI)		+	37°C RIL	NcoI/XhoI (NcoI/Sall)	5mM
3239 PilB	PilB	NdeI/HindIII (NdeI/HindIII)	NdeI/HindIII (NdeI/HindIII)	-	37°C STAR	NcoI/XhoI (NcoI/Sall)	5mM
YaeQ _{D68A} e YaeQ _{D83A}					37°C RP		
2396 <i>yaeQ</i>	5			+	37°C DE3 ou STAR	PvuII/XhoI (PvuII/Sall)	10 mM
2397	2397		NdeI/HindIII (NdeI/HindIII)	+	37°C DE3 ou RP	EcoRI/XhoI (EcoRI/Sall)	
1303 <i>mutS</i>	MutS	NdeI/BamHI (NdeI/BamHI)		N.D	37°C RIL	SmaI/XhoI (SmaI/Sall)	
2405 <i>mutL</i>	MutL	NdeI/BamHI (NdeI/BamHI)		N.D	37°C RIL	SmaI/XhoI (SmaI/Sall)	
2355 <i>sufE</i>	17			+	37°C DE3	NcoI/XhoI (NcoI/Sall)	10 mM

3671 <i>yajQ</i>	1			+	37°C DE3	NcoI/XhoI (NcoI/Sall)	5 mM
2934 <i>sufR</i>	SufR	NdeI/BamHI (NdeI/BamHI)		+	37°C STAR ou RP	NcoI/XhoI (NcoI/Sall)	
2117 <i>nifA^{like}</i>	NifU	NdeI/BamHI (NdeI/BamHI)		N.D	37°C STAR	NcoI/XhoI (NcoI/Sall)	
0499 <i>erpA</i>		NdeI/EcoRI (NdeI/EcoRI)		N.D	37°C STAR	NcoI/XhoI (NcoI/Sall)	
1619 <i>sufA</i>	SufA	NdeI/EcoRI (NdeI/EcoRI)		N.D		NcoI/XhoI (NcoI/Sall)	
3230	3230	NdeI/BamHI (NdeI/BamHI)			N.D		
2938 <i>sufS</i>	SufS	NdeI/EcoRI (NdeI/EcoRI)	NdeI/EcoRI (NdeI/EcoRI)	+	37°C STAR	NcoI/XhoI (NcoI/Sall)	
1877 HD-GYP	β	NdeI/EcoRI (NdeI/EcoRI)		-	37°C STAR		
0350 HD-GYP	α -d					NcoI/XhoI (NcoI/Sall)	5 mM
	α	NdeI/EcoRI (NdeI/EcoRI)		-	37°C STAR		
2493 HD-GYP	γ -d					NcoI/XhoI (NcoI/Sall)	5 mM
	γ	NdeI/EcoRI (NdeI/EcoRI)		-	37°C STAR		
0258*	G1	NdeI/HindIII (NdeI/HindIII)	NdeI/HindIII (NdeI/HindIII)	+	28°C DE3 ou RIL		
0258 GGDEF	G2	NdeI/BamHI (NdeI/BamHI)	NdeI/BamHI (NdeI/BamHI)	-	28°C RP	NcoI/XhoI (NcoI/Sall)	5 mM
0424 GGDEF	G3	NdeI/BamHI (NdeI/BamHI)	NdeI/BamHI (NdeI/BamHI)	+	28°C RP	NcoI/XhoI (NcoI/Sall)	5 mM
1420 GGDEF	G4	NdeI/BamHI (NdeI/BamHI)	NdeI/BamHI (NdeI/BamHI)	+	28°C RP	NcoI/XhoI (NcoI/Sall)	5 mM
1940 GGDEF	G5	NdeI/EcoRI (NdeI/EcoRI)	NdeI/EcoRI (NdeI/EcoRI)	+	28°C RP	NcoI/XhoI (NcoI/Sall)	5 mM
3768 ₄₇₆₋₇₁₀	MCP	NdeI/BamHI (NdeI/BamHI)		$\pm$	37°C RIL ou RP	NcoI/XhoI (NcoI/Sall)	

2.3 Sequenciamento dos clones (confirmados anteriormente por digestão do plasmídeo utilizando enzimas de restrição específicas)

O sequenciamento dos plasmídios foram feitos com base no método de Sanger *et al.* (1977)⁵⁵ utilizando o kit ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit III (Applied Biosystems). Os oligonucleotídeos flanqueadores do inserto *T7 forward primer* (+) de seqüência: 5' CACTATAGGGAGACCACAACGGTTTCC 3'.para o vetor pET3a, *T7 forward primer* (+) de seqüência: 5'TAATACGACTCACTATAGG3' para o vetor pET28a, *primer* (+) de sequencia: 5' CAAAGACAGTTGACTGTATCGAACTATC 3'.para o vetor pODB.

A reação de sequenciamento foi preparada da seguinte forma: 5 μ L de DNA de mini-preparação, ; 1 μ L de solução 4 pmol oligonucleotídeo sense (+) ou 1 μ L de

solução 10 μmol oligonucleotídeo anti-sense (-); 6 μL de tampão (200 mM Tris-HCl pH 9,0; 5 mM MgCl_2); 2 μL de Big Dye terminator; água deionizada para um volume final de 20 μL . Utilizando o termociclador Applied Biosystems GeneAmp® 9700, procedeu-se às reações de sequenciamento utilizando as seguintes temperaturas e tempos: 96°C por 45 segundos, 50 °C por 30 seg, 60 °C por 4 min. (40 ciclos). Terminada a reação, o DNA foi precipitado com isopropanol 75% por 15 minutos. Depois o DNA foi centrifugado por 40 minutos a 4000 rpm e o sobrenadante foi retirado pela inversão do tubo e subsequente centrifugação por 5 segundos a 1000 rpm. O seqüenciamento é feito utilizando o seqüenciador automático ABI 377 da Perkin-Elmer Applied Biosystems.

2.4 *Mutação sítio dirigida*

A preparação de plasmídios contendo genes de interesse com mutações sítio dirigida foram feitas utilizando o kit QuickChange Site-Directed Mutagenesis (STRATAGENE). O protocolo baseia-se de uma única etapa de PCR contendo oligos específicos (**Tabela 8**) e DNA molde que é o próprio plasmídio contendo o gene a ser mutado. A reação de PCR foi feita com 5.0 μL de tampão de reação; 1 μL de *Pfu* DNA Polymerase (25 U/ μL , Fermentas); 1 μL de dNTP_{mix}; 1 μL de solução 10 μmol oligonucleotídeo sense (+); 1 μL de solução 10 μmol oligonucleotídeo anti-sense (-); 2 μL de DNA plasmidial (5-50ng) e água deionizada para um volume final de 50 μL . Utilizando-se o termociclador Applied Biosystems GeneAmp® 9700, foram ajustadas as seguintes temperaturas e tempos de reação: 95°C por 30 segundo (1 ciclo); 95°C por 30 segundo, temperatura de anelamento de 55 °C por 1 minuto e 68°C por 1 minuto multiplicado pelo número de Kpb do plasmídio (5 ciclos); e 4°C por 10 minutos.

Após a reação de PCR foram adicionados 1 μL de Dpn1 em cada amostra e as amostras foram encubadas a 37 °C por 1 hora. Estas amostras foram utilizadas para transformar a cepa de *E. coli* XL1-Blue (empresa STRATAGENE) pelo método de cloreto de cálcio, descrito acima, com 5, 15 e 30 μL de reação e plaqueadas em meio LB contendo o antibiótico específico.

As colônias obtidas na transformação foram inoculadas em 3mL de 2xTY (16g/L Triton X-100 a; 10 g/L extrato de levedo; 5 g/L NaCl; pH 7,4) contendo 100 $\mu\text{g/mL}$ do antibiótico adequado. O DNA plasmidial das colônias foram purificadas utilizando o kit de extração de plasmídio da Qiagen e submetidas à sequenciamento (descrito acima).

A obtenção dos plasmídios para complementar células de *Xac* nocautes de *pilZ* com diferentes plasmídios contendo *pilZ* mutado foi feito utilizando os plasmídios de *pilZ* mutados no vetor pET3a, obtidos da forma descrita acima. Foram desenhados oligos específicos (**Tabela 4**) que contem o sítio BamHI e mais 7 pb antes do códon iniciador do gene de *pilZ*. Este fragmento foi amplificado e clonado no vetor pRM20 (expressão constitutivo em *Xac*)⁵⁶.

Tabela 8: Sequência dos oligonucleotídeos para a obtenção de genes mutados de *yaeQ* e *pilZ*.

Genes clonados no pET3a e/ou pOAD	Apelido	OLIGO	Sequência dos Oligonucleotídeos
YaeQ D68A	D68A	D68A (+) D68A (-)	5' AGTAACGAGACGAGCCT GCC CTATGGCGCCGCGACTAACC 3' 5' GGTGTAGTCGCGGCGCCATAGGGCAGGCTCGTCGTCGTTACT 3'
YaeQ D83A	D83A	D83A (+) D83A (-)	5' CCGGATCTGTGGATC GCC CTCGGCCAGCCTGACGAAAGC 3' 5' GCTTTCGTCAGGCTGGCCGAGGGCGATCCACAGAT CGG 3'
PilZ Y22A	1	Y22A (+) Y22A (-)	5' TGAAAGACAAGCCAGCGCTC GCC AGCGCCTACATGCCGTTTCGT 3' 5' ACGAACGGCATGTAGGCGCTGGCGAGCGCTGGCTTGTCTTTCA 3'
PilZ Y25A	2	Y25A (+) Y25A (-)	5' AGCCAGCGCTCTACAGCGCC GCC ATGCCGTTTCGTGAAAGGCGG 3' 5' CCGCCTTTCACGAACGGCATGGCGGCGCTGTAGAGCGCTGGCT 3'
PilZ F28A	3	F28A (+) F28A (-)	5' TCTACAGCGCCTACATGCCG GCC GTGAAAGGCGGCGGCATCTT 3' 5' AAGATGCCGCCGCCTTTCACGGCCGGCATGTAGGCGCTGTAGA 3'
PilZ Y22Y25F28A	4	Y22Y25F28A (+) Y22Y25F28A (-)	5' GCCAGCGCTC GCC AGCGCC GCC ATGCCG GCC GTGAAAGGCG 3' 5' CGCCTTTCACGGCCGGCATGGCGGCGCTGGCGAGCGCTGGC 3'
PilZ F35A	5	F35A (+) F35A (-)	5' TCGTGAAAGGCGGCGGCATC GCC GTGCCACGCCCAAGCGCTA 3' 5' TAGCGCTTGGGCGTGGGCACGGCGATGCCGCCGCCTTTCACGA 3'
PilZ W69A	6	W69A (+) W69A (-)	5' CGGTCGCCGGCAAGGTCATC GCG ACCACGCCTGCCGGCGCACA 3' 5' TGTGCGCCGGCAGGCGTGGTTCGCGATGACCTTGCCGGCGACCG 3'
PilZ R97A	7	R97A (+) R97A (-)	5' GCCCGGAAGGCGAGGCGGT GCCA ACAAGATCGAGACGCTGCT 3' 5' AGCAGCGTCTCGATCTTGTGGCCACCGCCTCGCCTTCCGGGC 3'
PilZ L104A	8	L104A (+) L104A (-)	5' AGGCGGTGCGCAACAAGATC GCG ACGCTGCTGGCCGGATTGAC 3' 5' GTCAATCCGGCCAGCAGCGTCGCGATCTTGTGCGCACCGCCT 3'
PilZ E101A	9	E101A (+) E101A (-)	5' GCAACAAGATCGAGACGCTG GCG GCCGGATTGACCACCTCGGA 3' 5' TCCGAGGTGGTCAATCCGGCCGCGCAGCGTCTCGATCTTGTTCG 3'
PilZ STOP107	10	STOP107 (+) STOP107 (-)	5' TCGAGACGCTGCTGGCCGGAT GAA CCACCTCGGACAAGCCGAC 3' 5' GTCGGCTTGTCCGAGGTGGTTTCATCCGGCCAGCAGCGTCTCGA 3'

2.5 Obtenção de cepas de *Xac* nocautes

2.5.1 Obtenção de cepas nocautes de *Xac*

2.5.1.1 Construção dos oligonucleotídeos para nocaute apolar de *YaeQ*, *SufE*, *PilZ* e *SufR* em *Xac*.

O vetor **pNPTS138** (**Figura 2**) é um vetor suicida, capaz de se transferir por conjugação em outras bactérias. Este plasmídeo contém o gene de resistência a kanamicina⁵⁷. Basicamente, amplificamos dois fragmentos: o primeiro contém mais ou menos 40 pb do começo da região codificadora do gene e mais 1000 pb antes dele; o segundo contém mais ou menos 40 pb do final do gene de interesse mais 1000 pb que o prosseguem.

Posteriormente, os dois fragmentos são ligados e clonados no vetor pNPTS138, como descrito mais detalhadamente acima (**Tabela 3 e 6**). Este plasmídeo contendo o fragmento irá inserir no genoma de *Xac* por recombinação homóloga, por dois eventos de recombinação descritos a baixo. Portanto, o fragmento contido no plasmídeo irá trocar com o do cromossomo. Assim sendo, apenas um pequeno fragmento (**Tabela 6**) do gene irá ser expresso.

Figura 2: Mapa do vetor suicida pNPTS138⁵⁷.

2.5.1.2 Transformação de *Xac* – Primeiro Evento de Recombinação.

Xac é encubada em 1mL de meio 2xTY durante 12 horas sob agitação em 28°C. Este pré-inóculo é centrifugado por 10 minutos a 6000 r.p.m. As células são ressuspensas em 50 mL de meio Circle-grow e encubadas sob agitação em 28 °C até D.O_{600nm} entre 0.5-1.0. As células são centrifugadas por 10 minutos em 6000 r.p.m e lavada três vezes com água estéril. As células são ressuspensas no volume residual e 50µL destas células são utilizadas para eletroporar. São adicionados 3µL de DNA plasmidial aos 50µL de células eletrocompetentes. A transformação é feita em cubetas de 0,2 cm e são utilizados 2KV, 25uF e 200V para eletroporar as células. As células são recuperadas em 1mL de meio 2xTY por 4 horas a 28 °C. As células são centrifugadas por 10 minutos em 5000 rpm e plaqueadas em placas grandes contendo LB agar, 100µg/ml de ampicilina e 100µg/ml kanamicina. Os clones obtidos da transformação de *Xac* são picados em duas placas LB contendo 1,5% agar. Uma das placas contem 100µg/ml de ampicilina e 100µg/ml kanamicina e a outra contem 100µg/ml de ampicilina, 100µg/ml kanamicina e 5% sacarose.

2.5.1.3 Segundo Evento de Recombinação

As colônias que cresceram em meio ampicilina e kanamicina e não cresceram em meio contendo sacarose são encubadas em 1mL de 2xTY com 100µg/ml de ampicilina, sob agitação em 28°C por 12 horas. As células do pré-inóculo são diluídas (1: 10⁴, 1: 10⁵ e 1: 10⁶) e plaqueadas em placas contendo LB agar e 100µg/ml de ampicilina.

As colônias são picadas em três condições: 1- LB agar contendo 5% sacarose e 100µg/ml de ampicilina; 2- LB agar contendo 5% sacarose, 100µg/ml de ampicilina e 100µg/ml de kanamicina e 3- LB agar contendo 100µg/ml de kanamicina. As colônias que cresceram em meio LB agar contendo 5% sacarose e 100µg/ml de ampicilina são repicadas novamente nas três condições descritas acima. As colônias que cresceram apenas na presença de ampicilina e sacarose são encubadas em 1mL de 2xTY com 100µg/ml de ampicilina, sob agitação em 28°C por 12 horas e armazenadas com 30% glicerol no freezer -80°C. A confirmação das colônias nocautes são feitas por PCR de colônia utilizando como par de oligonucleotídeos o Forward do gene e o reverse do fragmento 2 utilizado para a construção do gene truncado. A reação de PCR é feita picando cada colônia de *Xac* em 20µL de água destilada. Desta amostra utiliza-se 5µL

para a reação de PCR, que é feita pela adição de 2,5 µL de tampão 10X da enzima DNA polimerase, 0,5 µL de dNTP_{mix} (25mM), 0,5 µL de primer sense (+) e antisense (-) (10pMOL/µL), 0,25 µL de enzima DNA polimerase, 0,5 µL de tampão contendo magnésio, 15,25 µL de água autoclavada. A reação de PCR é equivalente ao descrito acima no item 1.1.

Os nocautes dos genes *pilB*_{XAC3239}, *fliC*_{XAC1975} e *fimX*_{XAC2398} foram feitos pelo aluno de doutorado do laboratório, Mestre Maxuel de Oliveira Andrade, e suas produções estão descritos nos Materiais e Métodos do Anexo 2.

2.5.1.4 Análise e complementação de cepas nocautes de Xac

Cepas nocautes de *pilZ*_{XAC1133} foram complementadas com plasmídios contendo o gene tipo selvagem clonado nos vetores pUFR047 (sob controle do promotor endógeno)⁵⁸ e/ou pMR20 (expressão constitutivo em *Xac*)⁵⁶. A transformação de cepas de *Xac* nocauteadas com os plasmídios para a complementação foi feita conforme descrito no item 1.5.1.2 em que as células foram plaqueadas em meios contendo antibióticos adequados para selecionar os plasmídios pUFR047 ou pMR20. Cepas nocautes de *pilZ*_{XAC1133} foram complementadas também com genes mutantes de *pilZ* clonados no pMR20 que foram construídos a partir dos plasmídios pET3a (vetores de expressão)^{59,60} que continham *pilZ* mutados.

A complementação das células de *Xac* nocautes do gene *fimX*_{XAC2398} foi feita pelo aluno de doutorado do laboratório, Mestre Maxuel de Oliveira Andrade, e está descrita nos materiais e métodos do Anexo 2.

2.6 Teste de Indução e Larga Escala

Inicialmente foram feitos testes de indução de cada proteína em diferentes cepas de *E. coli*, BL21(DE3), BL21(DE3)STAR, BL21(DE3)RIL e BL21(DE3)RP⁶⁰. As cepas nas quais as proteínas induziram em maior quantidade e a temperatura ideal de indução estão esquematizados na **Tabela 7**. Também estão indicados na **Tabela 7** os vetores nos quais os genes de interesse estão clonados. As células transformadas foram plaqueadas em meio sólido 2xTY (10g/l Triton X-100 a; 5 g/l extrato de levedo; 8 g/l NaCl; 7,5 g/l agar) contendo 100 µg/ml de antibiótico adequado, ampicilina para os genes clonados no vetor de expressão pET3a e kanamicina para os genes clonados no vetor de expressão pET28a e cloranfenicol, com exceção da cepa BL21(DE3).

Algumas colônias foram usadas para inocular 4 mL de meio 2xTY contendo os antibióticos adequados a 37 °C.

Quando a densidade ótica ($D.O_{600nm}$) se encontrava entre 0,8 e 1,0, estas células foram transferidas para 500 mL de meio 2xTY em frasco de Erlenmeyer de 2L e encubadas a 37 °C e a 200 r.p.m. Na $D.O_{600nm}$ entre 0,8 a 1,0 recolheu-se 1 mL de amostra não induzida e procedeu-se à indução, adicionando 250 μ L de 1 M IPTG, após 15 minutos de indução a temperatura do shaker foi transferida para a temperatura ideal de expressão (**Tabela 7**).

Após 6h de indução recolheu-se 1 mL de amostra induzida e as células foram recolhidas por centrifugação durante 15 min., a 8000 rpm, 4°C. O precipitado foi ressolubilizado em 25 mL de solução de lise por litro de cultura (1 mM de EDTA, 25 % sacarose, 50 mM Tris-HCl pH = 8,0, 0.03% glicerol, 0.03% Triton X-100 e 0.03% Tween), adicionou-se 14mM β mercaptoetanol e as células foram armazenadas a -20°C.

As amostras, não induzida e induzida, foram centrifugadas a 14000 rpm a temperatura ambiente por 20 min., ressuspensas em tampão de amostra (10% p/v glicerol; 2% SDS p/v; 50 mM Tris-HCl pH 6,8; 0.1% Bromofenol blue, 100 mM β -mercaptoetanol) e encubadas a 100 °C por 5 min. As amostras foram submetidas à análise por SDS-PAGE (Sodium Dodecyl Sulphate – Polyacrylamide gel Electrophoresis) em géis de 15% acrilamida e 0.4% bisacrilamida para análise de proteínas maiores de 20KDa e géis de 18% acrilamida e 0.4% bisacrilamida para análise de proteínas menores que 20KDa.

2.7 Incorporação de selênio-metionina em proteínas heterólogas.

Células transformadas com o plasmídio de expressão contendo o gene a ser expresso são encubadas em 5 mL de meio 2xTY com antibiótico adequado à 37°C, 200 r.p.m durante a noite. Este pré-inoculo é centrifugado e as células são ressuspensas em 1 mL de meio mínimo (6 g/L de Na_2HPO_4 pH = 7,35; 3 g/L de KH_2PO_4 ; 0,5 g/L NaCl; 2 mM $MgSO_4$, 0,4% glicose; 100 μ M $CaCl_2$; 1 g/L NH_4Cl). Estas células são adicionadas em 500 mL de meio mínimo contendo antibiótico adequado e encubadas a 37°C, sob agitação de 200 r.p.m, até D.O de 0,8. Retira-se 1mL de cultura para controle de não induzido e adiciona-se os seguintes aminoácidos: 100mg/L de lisina, fenilalanina e treonina; 50 mg/L de isoleucina e valina e 60mg/L de selênio-metionina⁶¹. Ao final de quinze minutos dar-se-á a expressão dos genes de

interesse através da adição de 1mM de IPTG. As células são recolhidas por centrifugação e ressolubilizadas em 25 mL de solução de lise desgaseificada (1 mM de EDTA, 25 % sacarose e 50 mM Tris-HCl pH = 8,0, deixada por 45 minutos no vácuo) e posteriormente submetidas à lise celular e as etapas de purificação da proteína.

2.8 Expressão de proteínas marcadas com ^{15}N e/ou ^{13}C .

Células transformadas com o plasmídeo de expressão contendo o gene a ser expresso são encubadas em 5 mL de meio 2xTY com antibiótico adequado à 37°C, 200 r.p.m durante a noite. Este pré-inoculo é centrifugado e as células são ressuspensas em 1 mL de meio mínimo (6 g/L de Na_2HPO_4 pH = 7,35; 3 g/L de KH_2PO_4 ; 0,5 g/L NaCl; 2 mM MgSO_4 , 0,4% glicose; 100 μM CaCl_2 ; 1 g/L NH_4Cl). Estas células são adicionadas em 500 mL de meio mínimo contendo antibiótico adequado e encubadas a 37°C, sob agitação de 200 r.p.m, até D.O de 0,8. Retira-se 1mL de cultura para controle (não induzido) e centrifuga as células em condições estéreis e as ressuspende em meio mínimo marcado com ^{15}N (6 g/L de Na_2HPO_4 pH = 7,35; 3 g/L de KH_2PO_4 ; 0,5 g/L NaCl; 2 mM MgSO_4 , 0,4% glicose; 100 μM CaCl_2 ; 1 g/L $^{15}\text{NH}_4\text{Cl}$) ou duplamente marcado ^{15}N e ^{13}C (6 g/L de Na_2HPO_4 pH = 7,35; 3 g/L de KH_2PO_4 ; 0,5 g/L NaCl; 2 mM MgSO_4 , 0,4% glicose marcada com ^{13}C ; 100 μM CaCl_2 ; 1 g/L $^{15}\text{NH}_4\text{Cl}$) com antibióticos adequados e adiciona 500 μL de 1M IPTG em cada 500mL de cultura. Após 4 horas de indução retira-se 1mL de cultura para controle induzido e o restante é recolhido por centrifugação e ressuspendido em 25 mL de solução de lise (1 mM de EDTA, 25 % sacarose e 50 mM Tris-HCl pH = 8,0) e posteriormente submetidas à lise celular e as etapas de purificação da proteína.

2.9 Lise Celular por French Press

As células induzidas passaram cinco vezes no French Press® sob pressão de 16000psi. O lisado obtido foi centrifugado por 1 hora, 4 °C e 13000 rpm. O precipitado, que contem as proteínas insolúveis, foi separado do sobrenadante, que contem as proteínas solúveis, e ressuspendido no mesmo volume que o sobrenadante em tampão discriminado no item de purificação, a seguir. Esta amostra ressolubilizada foi submetida à centrifugação por 1 hora, 4 °C e 13000 rpm e o sobrenadante foi separado. As amostras do sobrenadante da lise celular e do precipitado devidamente ressolubilizado, foram analisadas por SDS-PAGE para averiguar se a proteína é solúvel ou insolúvel.

2.10 Purificação das Proteínas de Interesse

As etapas de purificação das proteínas estudadas neste trabalho estão esquematizadas na **Tabela 9**. As frações mais puras de cada proteína no final do processo de purificação foram utilizadas para teste de cristalização e em outros ensaios funcionais. As amostras de proteínas puras foram armazenadas no freezer à -20°C, algumas proteínas após serem congeladas precipitavam e estas foram congeladas com 50% glicerol para evitar precipitação após o descongelamento.

A dosagem das proteínas foram feitas utilizando o método de espectrofotometria, que relaciona a concentração da proteína com a absorbância em 280 nm menos o valor de dispersão em 330 nm e seu respectivo coeficiente de absorvidade molar (Absorb(Prot)), lei de Bouguer-Lambert-Beer, calculado pelo programa ProtParam⁶² que utiliza as seguintes equações:

$$\text{Absorb(Prot)} = \text{Ext(Prot)}/\text{Massa\_molar\_da\_proteína}$$

$$\text{Ext(Prot)} = \text{N(Tyr)} \times \text{Ext(Tyr)} + \text{N(Trp)} \times \text{Ext(Trp)} + \text{N(Cystine)} \times \text{Ext(Cystine)}.$$

Onde N é o número de resíduos do tipo indicado e Ext é a coeficiente de extinção da proteína ou daquele resíduo em 280 nm. Em 280nm os valores de Ext(Tyr) é igual a 1490, Ext(Trp) é igual a 5500, Ext(Cystine) é igual a 125⁶².

Tabela 9: Resumo das condições de purificação das proteínas.

Proteínas envolvidas em “Quorum Sensing” em Xac			
Proteínas	1ª Etapa	2ª Etapa	3ª Etapa
CMF _{his} -XAC0330	<u>Solubilização do Precipitado da Lise Celular:</u> 6M uréia 50mM Tris pH 8,6 1mM EDTA 14mM βmercaptoetanol <u>Coluna Aniônica-Q-Sepharose-FF</u> 6M uréia 50mM Tris pH 8,6 1mM EDTA 14mM βmercaptoetanol <u>Gradiente</u> 0-400mM NaCl. 12CV <u>Diálise</u> 6M ureia 50mM Tris pH 8,0 <u>Concentração</u> Amicon com filtro 10KDa.	<u>Coluna de Afinidade-HisTrap™(GE)</u> 6M uréia 50mM Tris pH 8,0 300mM NaCl, 20mM imidazol <u>Gradiente</u> 0-1M imidazol. 12CV <u>Diálise</u> 5mM Acetato de Sódio pH 5,0 <u>Concentração</u> Amicon com filtro 10KDa.	<u>Coluna de Exclusão por Tamanho-Superdex-75 prepgrade</u> 5mM Acetato de Sódio pH 5,0 <u>Concentração</u> Amicon com filtro 10KDa. Armazenadas com 50% glicerol

RpfF _{XAC1879}	<u>Solubilização do Precipitado da Lise Celular:</u> 8M uréia 50mM Tris pH 8,0 1mM EDTA 14mM βmercaptoetanol <u>Coluna Aniônica-Q-Sepharose-FF</u> 8M ureia 50mM Tris pH 8,0 1mM EDTA 14mM βmercaptoetanol <u>Gradiente</u> 0-400mM NaCl. 12CV <u>Concentração</u> Amicon com filtro 10KDa.	<u>Coluna de Exclusão por Tamanho-Superdex-200 prepgrade</u> 8M uréia 50mM Tris pH 8,0 1mM EDTA 14mM βmercaptoetanol <u>Diálise</u> 5mM Acetato de Sódio pH 5,0 <u>Concentração</u> Amicon com filtro 10KDa.	
RpfG _{XAC1877}	<u>Solubilização do Precipitado da Lise Celular:</u> 6M uréia 50mM Tris pH 7,0 1mM EDTA 14mM βmercaptoetanol <u>Coluna Aniônica-Q-Sepharose-FF</u> 6M ureia 50mM Tris pH 7,0 1mM EDTA 14mM βmercaptoetanol <u>Gradiente</u> 0-400mM NaCl. 12CV <u>Concentração</u> Amicon com filtro 10KDa.	<u>Coluna de Exclusão por Tamanho-Superdex-200 prepgrade</u> 6M uréia 50mM Tris pH 7,0 1mM EDTA 14mM βmercaptoetanol <u>Diálise</u> 5mM Tris pH 7,0 1mM EDTA 14mM βmercaptoetanol <u>Concentração</u> Amicon com filtro 10KDa.	<u>Coluna de Exclusão por Tamanho-Superdex-200 prepgrade</u> 5mM Tris pH 7,0 7mM βmercaptoetanol <u>Concentração</u> Amicon com filtro 10KDa. Armazenadas com 50% glicerol
RpfC _{XAC1878_HPT (584-699)}	<u>Coluna de Afinidade-HisTrap™(GE)</u> 50mM Tris pH= 8,0 300mM NaCl, 20mM imidazol <u>Gradiente</u> 0-1M imidazol. 20CV <u>Concentração</u> Amicon com filtro 10KDa.	<u>Coluna de Exclusão por Tamanho-Superdex-75 prepgrade</u> 5mM Tris pH 7.0 <u>Diálise</u> 50mM Tris pH 8,0 10mM CaCl ₂ <u>Concentração</u> Amicon com filtro 10KDa Tratamento com Trombina do Kit Thrombin CleanCleave™	<u>Coluna de Exclusão por Tamanho-Superdex-75 prepgrade</u> 5mM Tris pH 7.0 <u>Concentração</u> Amicon com filtro 10KDa

RpfC_{his-RR(415-613)} e RpfC_{his-HPT(584-699)}	<u>Coluna de Afinidade-HisTrap™ (GE)</u> 50mM Tris pH= 8,0 300mM NaCl, 20mM imidazol <u>Gradiente</u> 0-1M imidazol. 20CV <u>Diálise</u> 50mM Tris pH 8.0 14mM βmercaptoetanol <u>Concentração</u> Amicon com filtro 10KDa.	<u>Coluna de Exclusão por Tamanho-Superdex-75 prepgrade</u> 5mM Tris pH 7.0 <u>Diálise</u> 5mM Tris pH 8.0 1mM DTT 1mM MgCl ₂ 5mM NaCl <u>Concentração</u> Amicon com filtro 10KDa.	
XAC0258_{his}	<u>Coluna de Afinidade-HisTrap™ (GE)</u> 50mM Tris pH 8,0 300mM NaCl, 20mM imidazol <u>Gradiente</u> 0-1M imidazol. 12CV <u>Concentração</u> Amicon com filtro 10KDa.	<u>Logo que sai da HisTrap</u> Adiciona: 28mM βmercaptoetanol e encuba 30°C por 30 minutos.	<u>Diálise</u> 5mM Tris pH 8,0 3mM DTT
XAC1420_{his-GGDEF(330-497)}	<u>Coluna de Afinidade-HisTrap™ (GE)</u> 50mM Tris pH 8,0 300mM NaCl, 20mM imidazol <u>Gradiente</u> 0-1M imidazol. 20CV <u>Diálise</u> 50mM Tris pH 8,0 14mM βmercaptoetanol <u>Concentração</u> Amicon com filtro 10KDa.	<u>Coluna de Exclusão por Tamanho-Superdex-75 prepgrade</u> 5mM Tris pH 7,0 1mM DTT 1mM MgCl ₂ <u>Concentração</u> Amicon com filtro 10KDa.	<u>Diálise</u> 5mM Tris pH 8,0 3mM DTT
Proteínas contendo Domínios HD-GYP			
Proteínas	1ª Etapa	2ª Etapa	3ª Etapa
<u>Domínios HD-GYP</u> RpfG_{his-HDGY(199-233)} XAC0350_{his-HDGY(171-311)}	<u>Solubilização do Precipitado da Lise Celular:</u> 6M uréia 50mM Tris pH 8,6 1mM EDTA 14mM βmercaptoetanol <u>Coluna de Afinidade-HisTrap™ (GE)</u> 6M uréia	<u>Diálise</u> 5mM Tris pH 8,0 1mM DTT 1mM MgCl ₂ <u>Concentração</u> Amicon com filtro 10KDa.	<u>Coluna de Exclusão por Tamanho-Superdex-75 prepgrade</u> 5mM Tris pH 8,0 1mM DTT 1mM MgCl ₂ <u>Concentração</u> Amicon com filtro

XAC2493 _{his-} HDGYP(146-288)	50mM Tris pH 8,0 300mM NaCl, 20mM imidazol <u>Gradiente</u> 0-1M imidazol. 12CV		10KDa.
Proteínas envolvidas na Biogênese do T4P em Xac			
Proteínas	1ªEtapa	2ªEtapa	3ªEtapa
FimX _{XAC2398}	<u>Coluna Aniônica-Q-Sepharose-FF</u> 50mM Tris pH 6,5 1mM EDTA 14mM βmercaptoetanol <u>Gradiente</u> 0-1M NaCl. 12CV <u>Concentração</u> Amicon com filtro 10KDa.	<u>Coluna de Exclusão por Tamanho-Superdex-200 prepgrade</u> 50mM Tris pH 8,0 14mM βmercaptoetanol 250mM NaCl. <u>Concentração</u> Amicon com filtro 10KDa.	<u>Coluna de Exclusão por Tamanho-Superdex-200 prepgrade</u> 50mM Tris pH 8,0 14mM βmercaptoetanol 250mM NaCl. <u>Concentração</u> Amicon com filtro 10KDa.
FimX _{EAL(436-689)}	<u>Coluna Aniônica-Q-Sepharose-FF</u> 50mM Tris pH 7,0 1mM EDTA 14mM βmercaptoetanol <u>Gradiente</u> 0-0.5M NaCl. 12CV <u>Concentração</u> Amicon com filtro 10KDa.	<u>Coluna de Exclusão por Tamanho-Superdex-200 prepgrade</u> 5mM Tris pH 7,0 1mM βmercaptoetanol 10 mM NaCl. <u>Concentração</u> Amicon com filtro 10KDa.	
FimX _{GGDEF (267-432)}	<u>Coluna Aniônica-Q-Sepharose-FF</u> 50mM Tris pH 7,0 1mM EDTA 14mM βmercaptoetanol <u>Gradiente</u> 0-0.5M NaCl 12CV <u>Concentração</u> Amicon com filtro 10KDa.	<u>Coluna de Exclusão por Tamanho-Superdex-200 prepgrade</u> 50mM Tris pH 8,0 14mM βmercaptoetanol 25mM NaCl. <u>Concentração</u> Amicon com filtro 10KDa.	<u>Coluna de Exclusão por Tamanho-Superdex-75 prepgrade</u> 50mM Tris pH 8,0 14mM βmercaptoetanol 500mM NaCl <u>Concentração</u> Amicon com filtro 10KDa.
FimX _{his-GGDEF-EAL(267-689)}	<u>Coluna de Afinidade-HisTrap™ (GE)</u> 50mM Tris pH 8,0 300mM NaCl, 20mM imidazol <u>Gradiente</u> 0-1M imidazol. 12CV	<u>Coluna de Exclusão por Tamanho-Superdex-200 prepgrade</u> 10mM Tris pH 8,0 5 mM βmercaptoetanol 10 mM NaCl. <u>Concentração</u> Amicon com filtro	

	<u>Concentração</u> Amicon com filtro 10KDa.	10KDa.	
PilZ _{XAC1133}	<u>Coluna Catiônica-SP-Sepharose</u> 50mM Tris pH 8,0 1mM EDTA <u>Gradiente</u> 0-1M NaCl. 12CV <u>Concentração</u> Amicon com filtro 3KDa.		
PilA _{XAC2340(32)}	<u>Coluna de Afinidade-HisTrap™(GE)</u> 50mM Tris pH 8,0 300mM NaCl, 20mM imidazol 6M uréia <u>Gradiente</u> 0-1M imidazol. 20CV <u>Concentração</u> Amicon com filtro 3KDa.	<u>Coluna de Exclusão por Tamanho-Superdex-75 prepgrade</u> 50mM Tris pH 8,0 14mM βmercaptoetanol 500mM NaCl 6M uréia	
YaeQ _{XAC2396} e XAC2397			
Proteínas	1ªEtapa	2ªEtapa	3ªEtapa
XAC2397 _{his}	<u>Coluna de Afinidade-HisTrap™(GE)</u> 50mM Tris pH= 8,0 300mM NaCl, 20mM imidazol <u>Gradiente</u> 0-1M imidazol. 20CV	<u>Coluna de Exclusão por Tamanho-Superdex-200 prepgrade</u> 5mM Tris pH = 8.0 25 mM NaCl <u>Concentração</u> Amicon com filtro 10KDa.	
YaeQ _{XAC2396} , YaeQ _{XAC2396-SelMet} YaeQ _{D68A} e YaeQ _{D83A}	<u>Coluna Aniônica-Q-Sepharose-FF</u> 50mM Tris pH 7,0 1mM EDTA 14mM βmercaptoetanol <u>Gradiente</u> 0-1M NaCl. <u>Concentração</u> Amicon com filtro 10KDa.	<u>Coluna de Exclusão por Tamanho-Superdex-75 prepgrade</u> 5mM Tris pH 7,0 <u>Concentração</u> Amicon com filtro 10KDa.	
Proteínas envolvidas na biossíntese de centros de Fe-S			
Proteínas	1ªEtapa	2ªEtapa	3ªEtapa
SufE _{XAC2355}	<u>Coluna Aniônica-Q-Sepharose-FF</u>	<u>Coluna de Exclusão por Tamanho-Superdex-75 prepgrade</u>	<u>Coluna de Exclusão por Tamanho-Superdex-200</u>

	50mM Tris pH 8,0 1mM EDTA <u>Gradiente</u> 0-300mM NaCl. <u>Concentração</u> Amicon com filtro 10KDa.	5mM Tris pH 8,0 <u>Concentração</u> Amicon com filtro 10KDa.	<u>prepgrade</u> 5mM Tris pH 8,0 <u>Concentração</u> Amicon com filtro 10KDa.
SufS _{XAC2938-his}	<u>Coluna de Afinidade- HisTrap™ (GE)</u> 50mM Tris pH= 8,0 300mM NaCl, 20mM imidazol 10µM piridoxalfosfato <u>Gradiente</u> 0-1M imidazol. 20CV	<u>Dialise</u> 5mM Tris pH 8,0 150µM piridoxalfosfato	<u>Coluna de Exclusão por Tamanho- Superdex-200 prepgrade</u> 5mM Tris pH = 8.0 150µM piridoxalfosfato 1mM DTT <u>Concentração</u> Amicon com filtro 10KDa.
SufR _{XAC2934}	<u>Coluna Aniônica- Q-Sepharose-FF</u> 50mM Tris pH 6,0 1mM EDTA 28mM βmercaptoetanol <u>Gradiente</u> 0-500mM NaCl. <u>Concentração</u> Amicon com filtro 3KDa.	<u>Coluna de Exclusão por Tamanho- Superdex-75 prepgrade</u> 5mM Tris pH 8,0 5mM βmercaptoetanol <u>Concentração</u> Amicon com filtro 3KDa.	

2.11 Cristalização e determinação da estrutura tridimensional

2.11.1 Teste de cristalização das proteínas purificadas

Inicialmente foram feitas triagens de cristalização utilizando a técnica de difusão de vapor por gota sentada. Teste de cristalização utiliza diferentes "kits" da empresa "Hampton Research", "Qiagen" e "Molecular Dimensions Limited", que foram desenvolvidos especialmente para testar várias condições de cristalização de proteínas solúveis e em complexo. Foram utilizadas caixas de gota sentada (sitting drop) contendo 24 poços. No reservatório foi aplicado 300 µL de solução do Kit e no pedestal foi adicionado 1,5 µL de proteína pura concentrada e 1,5 µL da solução do poço. A placa foi fechada com plástico aderente da empresa Hampton.

Posteriormente, as condições de cristalização foram otimizadas variando o pH, concentração da proteína e do agente precipitante e a temperatura, a fim de obter

cristais com bons padrões de difração quando submetidos à análise por difração de raios-X, vide a seguir.

2.11.2 Cristalização das proteínas purificadas: YaeQ, PilZ e os domínios RR e HPT da RpfC.

RpfC_{his-HPT(584-699)} e RpfC_{his-RR(415-613)}: Inicialmente foram feitas triagens utilizando o Kit da empresa Hampton Research, Crystal Screen I e II, Index, como descrito anteriormente. Obtivemos cristais pequenos dos domínios RR (5,5g/L) e HPT (6,32g/L) da RpfC. RpfC_{his-HPT(584-699)} cristalizou no poço 9 do Kit Crystal Screen I (2M Sulfato de amônio, 0,1M citrato de sódio pH = 5,6 e 0,2M tartarato de sódio potássio). A otimização das condições de cristalização foi feita variando a concentração de agente precipitante, a concentração da proteína e o pH. Os cristais de RpfC_{his-HPT(584-699)} são muito pequenos e por enquanto não conseguimos cristais bons o suficiente para determinar a sua estrutura.

RpfC_{his-RR(415-613)} cristalizou em 3 condições, poço 1, 6 e 14 do kit Crystal Screen I. Poço 1: 0,02M cloreto de cálcio, 0,1M acetato de sódio pH=4,6 e 30% 2-metil-2,4 pentanediol; poço 6: 0,2M cloreto de magnésio, 0,1M Tris pH = 8,5 e 30% PEG 4000 e poço 9: 0,2M acetato de amônio, 0,1M citrato de sódio pH = 5,6 e 30% PEG 4000. Em nenhuma destas condições os cristais de RpfC_{his-RR(415-613)} difratou.

PilZ_{XAC1133}: PilZ_{XAC1133} cristalizou em 12 condições: 1 – 30% PEG 4000, 0.1M Tris HCl pH = 8.5 e 0,2M MgCl₂, 2- 30% PEG 4000, 0.1M citrato de sódio pH = 5,6 e 0,2M Acetato de amônio, 3- 30% PEG 4000, 0.1M acetato de sódio pH = 4,6 e 0,2M Acetato de amônio, 4- 30% PEG 8000, 0.1M cocadilato de sódio pH = 6,5 e 0,2M sulfato de amônio, 5- 30% PEG 4000, 0.1M Tris HCl pH = 8,5 e 0,2M Sulfato de Lítio, 6- 20% PEG 8000, 0.1M cocadilato de sódio pH = 6,5 e 0,2M Acetato de magnésio, 7- 20% PEG 4000, 0.1M acetato de sódio pH = 4,6 e 0,2M sulfato de amônio, 8- 30% PEG 8000, 0.1M cocadilato de sódio pH = 6,5 e 0,2M Acetato de sódio, 9- 30% PEG 8000 e 0,2M sulfato de amônio, 10 - 30% PEG 4000 e 0,2M sulfato de amônio, 11- 18% PEG 8000, 0.1M cocadilato de sódio pH = 6,5 e 0,2M acetato de zinco, 12- 2M sulfato de amônio.

Todas as condições em que PilZ_{XAC1133} cristalizou foram refinadas variando a porcentagem de PEG, pH e temperatura, 18°C e 4°C. Para as condições em que não havia tampão foi refinado a porcentagem de PEG e a concentração do sal, e para as condições que havia apenas o sal, a concentração do sal foi refinada.

A melhor condição de cristalização de PilZ_{XAC1133} foi na placa nomeada D em que foi refinado o a concentração de PEG 4000 de 22 a 32%, (cada passo era de 2%) e o pH de 0.1M de Tris-HCl de 8,9 a 7,6 (variando 0,4 unidades de pH) e mantendo a concentração de 0.2M de cloreto de magnésio. PilZ_{XAC1133} difratou até 1,8 Å na condição de: 24% PEG 4000, 0,1M Tris-HCl pH 8,0 e 0,2M cloreto de magnésio.

YaeQ_{XAC2398}, YaeQ_{XAC2398-SelMet}, YaeQ_{D68A} e YaeQ_{D83A}: YaeQ_{XAC2398} na concentração 8 mg/mL cristalizou em 10 condições diferentes durante a triagem inicial. Utilizamos a técnica de gota pendurada e os kits da empresa Hampton research – Crystal Screen e Crystal screen II e também o Kit da empresa Molecular Dimensions Limited – Clear Strategy Screen 1 tampões 1,2,3,4 e 5 para fazer os testes de cristalização iniciais. As condições de cristalização iniciais são: 1- 0,2M acetato de sódio, 0,1M Tris pH 8,5 e 30% W/V PEG 4000; 2- 0,2M acetato de sódio, 0,1M Tris pH 8,5 e 30% W/V PEG 4000; 3- 0,2 M MgCl₂, 0,1M Tris pH 8,5 e 25% PEG 2K MME; 4- 0,2 M MgCl₂, 0,1M Tris pH 7,5 e 25% PEG 2K MME ; 5- 0,3 M acetato de sódio, 0,1M Tris pH 8,5 e 25% PEG 2K MME ; 6- 0,2M acetato de sódio triidratado, 0,1M cocadilato de sódio pH 8,5 e 30% W/V PEG 8000; 7- 0,2M acetato de sódio triidratado, 0,1M cocadilato de sódio pH 8,5 e 30% W/V PEG 8000; 8- 0,2M cloreto de cálcio Diidratado, 0,1 M HEPES pH 7,5 e 28% W/V PEG 400; 9- 0,2M acetato de amônio Triidratado, 0,1M citrato de sódio diidratado pH 5,6 e 30% PEG 4000; 10- 0,3M acetato de sódio, 0,1M acetato de sódio pH 5,5 e 25% W/V PEG 2000 MME.

Todas as condições em que YaeQ_{XAC2398} cristalizou foram refinadas variando a porcentagem de PEG e pH. As melhores condições de cristalização de YaeQ_{XAC2398}, YaeQ_{XAC2398-SelMet} e dos mutantes de YaeQ, YaeQ_{D68A} e YaeQ_{D83A} foram na faixa de 30-32% PEG 4000, 0,2M acetato de amônio e 100mM Tris pH 8,5-9,0 ou 30% PEG8000, 100mM cocadilato de sódio pH 6,7 e 200mM acetato de sódio.

Os testes de cristalização e o refinamento das condições de cristalização da proteína SufE_{XAC2355} foram feitos pela aluna de pós doutorado do laboratório, Dra. Lucicleide Ribeiro, e estão descritos nos materiais de métodos do Anexo 5.

2.11.3 Coleta de dados de cristalografia

A **Tabela 10** mostra de forma resumida a linha de luz utilizada para coleta dos dados de difração dos cristais de SufE_{XAC2355}, PilZ_{XAC1133}, Domínio HPT da RpfC (RpfC_{his-HPT(584-699)}), YaeQ_{XAC2396} e YaeQ_{D83A} e a metodologia utilizada para resolver as fases.

Tabela 10: Linhas de raios-X utilizadas para coleta dos dados de difração dos cristais de SufE, PilZ, HPT, YaeQ e YaeQ_{D83A}.

<i>Proteína</i>	<i>Linha de Raios-X</i>	<i>Técnica empregada para resolver as fases</i>
YaeQ	MX1 LNLS, Campinas	-----
YaeQ-SelMet	“Cristallography Beam Line” Advanced Light Source (ALS), Berkeley, Califórnia	MAD
YaeQ_{D83A}	MX2 LNLS, Campinas	Substituição Molecular
SufE	MX1 LNLS, Campinas	Substituição Molecular
PilZ-SelMet	MX2 LNLS, Campinas	MAD
HPT	MX1 LNLS, Campinas	-----

2.11.3.1 Determinação dos comprimentos de onda para coleta de dados MAD utilizando o programa Blu-Ice

A determinação dos valores dos parâmetros do espalhamento anômalo f' e f'' são importantes para o processo de refinamento das posições dos metais pesados e para a determinação das fases. Assim sendo, a precisão dos valores destes parâmetros muitas vezes auxiliam em uma rápida convergência dos dados no cálculo e no refinamento do mapa de densidade eletrônica inicial⁶³. A determinação destes parâmetros é feita pelo programa Blu-Ice acoplado a linha de raios-X MX2 e o procedimento está descrito a seguir⁶⁴

2.11.3.1.1 Inicialmente confirmamos a presença de Selênio na amostra pelo “Excitation scan”

Inicialmente calibramos o motor que ajusta o comprimento de onda e em seguida calibramos a posição da mesa do detector. Para esta última calibração utilizamos o programa MARCCD. Após este procedimento especificamos o metal pesado que está sendo analisado, no nosso caso se trata do átomo de selênio, que contém o pico de absorção de raios-X em 12650 eV. Após a determinação da energia do pico de absorção fazemos um “Excitation scan” que irá medir a emissão de fluorescência de raios-X em função da energia incidente de raios-X, que é iniciada 500eV a menos que o pico de absorção de raios-X, com o objetivo de confirmar a presença de átomos de selênio pela amostra. Este procedimento também pode ser

utilizado para detectar a presença de outros metais pesados pela amostra, pois ele mede a fluorescência de todos os elementos presentes na amostra com o pico de absorção menor que a energia de excitação sendo utilizada. Este espectro mede a fluorescência de raios-X em uma varredura ampla de energia ao redor do pico de absorção de raios-X de forma rápida mas sem precisão.

Nesta mesma etapa ajustamos a intensidade de fluorescência sendo emitida pela amostra para 100 “counts” ajustando a quantidade de filtros, para que o espectro de fluorescência de raios-X do átomo de selênio, no próximo passo, “MAD scan”, seja adequado para o cálculo dos parâmetros de espalhamento anômalo.

2.11.3.1.2 Determinação dos parâmetros f' e f'' pelo “MAD scan”

Após a calibração dos filtros e a confirmação da presença dos átomos de selênio pela amostra procedemos com o “MAD scan”. Este modo varia a energia de raios-X de excitação e mede a emissão de fluorescência pela amostra. Esta etapa é mais demorada que a etapa anterior por conter muitos mais pontos em função da energia de raios-X incidente. A partir deste gráfico o programa autochoc⁶³, acoplado a interface Blu-Ice, calcula automaticamente os valores de espalhamento anômalo. Os valores de f' e f'' são mostrados assim como suas correspondentes energias para o pico, ponto de inflexão, e de um ponto afastado do pico de absorção de raios-X pelo selênio da amostra (remoto). O espectro de fluorescência de raios-X permite a determinação dos coeficientes de absorção do átomo, pois estes dois valores estão relacionados por uma constante⁶³.

2.11.3.1.3 Coleta de Dados MAD

Após a determinação das energias do pico, inflexão e de uma energia distante do pico de absorção, ajustamos a energia inicialmente para o pico de absorção e coletamos o primeiro conjunto de dados. Após o término desta coleta procedemos da mesma forma para a coleta dos dados das outras duas energias (inflexão e remoto).

2.11.4 Processamento dos dados cristalográficos e obtenção da estrutura tridimensional

Os dados obtidos por difração de raios-X foram processados pelo programa HKL2000⁶⁵ ou Mosflm⁶⁶. Para os conjuntos de dados que contenham homólogos com estrutura depositada no PDB foi utilizado a técnica de substituição molecular, usando o

programa PHASER^{67,68} do pacote do CCP4i para determinar as fases iniciais. No caso de conjuntos de dados que contenham metais pesados outros programas do pacote CCP4i, Shelxd⁶⁹, SHARP e auto-SHARP⁷⁰ foram utilizados para a determinação das posições iniciais dos átomos pesados e posteriormente para o cálculo das fases iniciais. A partir das fases iniciais o modelo inicial foi construído pelo programa ARP/wARP⁷¹ do pacote CCP4i. O modelo inicial foi refinado pelos programas REFMAC⁷², CNS⁷³, COOT⁷⁴ e utilizamos TLS⁷⁵ nos parâmetros de refinamento. Moléculas de água foram adicionadas automaticamente pelo programa REFMAC⁷² e manualmente pelo COOT⁷⁴. Validação da estrutura foi feita pelos programas PROCHECK⁷⁶ e CNS⁷³.

2.12 Ensaios Funcionais

2.12.1 Ensaio de 2-híbrido

2.12.1.1 Transformação da linhagem PJ69-4a com a isca de interesse

Células de levedura PJ694-a⁷⁷ foram transformadas com o plasmídeo pOBD contendo o gene de interesse (isca) segundo o método descrito por Gietz e colaboradores^{78,79}. O protocolo consiste em crescer células de levedura em placas de petri contendo YAPD sólido autoclavado (6g de extrato de levedura, 12g de peptona, 12g de glicose, 48 mg de adenina hemi-sulfato em pó ou 24mg se estiver na forma líquida, 10 g de agar e água para completar o volume para 600mL). São recolhidos 25µL de células e estas são lavadas com água autoclavada. São adicionados a célula 240 µL de 50% PEG 3350, 36 µL de 1M acetato de lítio, 50 µL de espermatozoides de salmão, 4 µL de DNA plasmidial e 30 µL de água. Esta solução é homogeneizada usando a pipeta e encubada a 42°C por 30 minutos, a cada 10 minutos as células são homogeneizadas. Posteriormente, as células são centrifugadas a 6000 r.p.m por 2 minutos, o sobrenadante é retirado e adicionado 200 µL de água e incubado a temperatura ambiente por 5 minutos. Após este período as células são homogeneizadas e plaqueadas em meio SC-W (12g de glicose, 4g de base de nitrogênio sem aminoácidos, 0,5 g de mix de aminoácidos -WHLA (300mg de isoleucina, 1500mg de valina, 200mg de argenina, 300mg de lisina, 200mg de metionina, 500mg de fenilalanina, 200mg de treonina, 300mg de tirosina, 200mg de uracila, 1000mg de ácido glutâmico, 1000mg de ácido aspártico e 4000mg de serina), 18mg de adenina, 12mg de histidina, 60mg de leucina, 10g de agar e água para

completar o volume para 600 mL e o pH é ajustado para 5.8) e encubadas a 30°C por 2-4 dias.

2.12.1.2 Teste de auto-ativação das células de levedura contendo a Isca.

As células transformadas foram submetidas a teste de auto-ativação a fim de determinar a quantidade de 3-aminotriazol necessária para os passos posteriores. Células transformadas crescidas em meio SC-W são alçadas e ressuspensas em 1mL de água autoclavada, desta solução são retiradas 5 µL e diluída em 1mL. 100 µL desta solução diluída são plaqueadas em placas de petri contendo SC-W, SC-WH (12g de glicose, 4g de base de nitrogênio sem aminoácidos, 0,5 g de mix de aminoácidos - WHLA, 18mg de adenina, 60mg de leucina, 10g de agar e água para completar o volume para 600 mL, pH 5.8), contendo diferentes concentrações de aminotriazol (0, 3mM, 5mM, 10mM, 15mM, 25mM e 35mM) e encubadas a 30°C por 5 dias. O objetivo é determinar a menor quantidade de aminotriazol que as células não conseguem crescer. O resultado deste teste para cada isca estão descritos na **Tabela 7**.

2.12.1.3 Transformação com a biblioteca genômica de Xac e confirmação dos clones positivos

A linhagem previamente transformada com a isca, crescidas em meio SC-W, é transformada com a biblioteca de Xac, clonada no vetor pOAD^{33,80,81}. Células de levedura PJ694-a⁷⁷ transformadas com a isca são crescidas em 50 mL de meio SC-W líquido durante 14 horas e recolhidas por centrifugação a 4000 r.p.m por 5 minutos e ressuspensas 150 mL de meio YAPD. A cultura de células é incubado a 30°C até atingir D.O_{600nm} de 2. As células são centrifugadas e lavadas 3 vezes com água autoclavada, e estas são ressuspensas em 3 mL de 100 mM acetato de lítio e encubadas a 30 °C por 15 minutos. As células são centrifugadas e sobre as células são adicionados 7,2 mL de 50% PEG 3350, 1,08 mL de 1M acetato de lítio, 1,5 mL de espermatozoides de salmão, 30 µL de DNA da biblioteca (fragmentos de 500-1500 pb) e 990 µL de água. Esta solução é homogenizada usando a pipeta e encubada a 30°C por 15 minutos, e posteriormente encubada a 42°C por 45 minutos, a cada 10 minutos as células são homogeneizadas. As células são centrifugadas a 6000 r.p.m por 2 minutos, o sobrenadante é retirado e adicionado 5 mL de água e incubado a temperatura ambiente por 5 minutos. Após este período as células são homogeneizadas e plaqueadas em meio SC-WHL (12g de glicose, 4g de base de nitrogênio sem

aminoácidos, 0,5 g de mix de aminoácidos -WHL, 18mg de adenina, 10g de agar e água para completar o volume para 600 mL, pH 5.8) e encubadas a 30°C por 14 dias. As células transformadas são selecionadas em meio seletivo para interação, SC-WHLA (12g de glicose, 4g de base de nitrogênio sem aminoácidos, 0,5 g de mix de aminoácidos -WHL, 10g de agar e água para completar o volume para 600 mL, pH 5.8).

2.12.1.4 Purificação do DNA plasmidial de levedura

As células crescidas em meio seletivo SC-WHL foram crescidas em 1,2 mL de SC-L (12g de glicose, 4g de base de nitrogênio sem aminoácidos, 0,5 g de mix de aminoácidos -WHL, 18mg de adenina, 12mg de histidina, 12 mg de triptofano e água para completar o volume para 600 µL, pH 5.8), a 30 °C a 200 r.p.m por 48 horas. Estas células são centrifugadas ressuspensas em Tampão Rescue (50mM Tris pH=7,5, 10mM EDTA pH =8,0, 0,3% beta-mercaptoetanol). As células foram centrifugadas e sobre as células foram adicionados 25µL de solução de lise (2mg/mL de zymoliase, 0,2mg de RNase e o volume é completado com tampão Rescue), encubadas a 37°C por 1 hora sob agitação de 150 r.p.m.. Posteriormente foram adicionados 25 µL de 10% SDS na cultura e encubada a temperatura ambiente por 5 minutos. As células foram congeladas a -80 °C por 15 minutos. Em seguida foram adicionados 100 µL de 3M Acetato de potássio pH 5,2, e encubada no gelo por 10 minutos. As células foram centrifugadas e o sobrenadante foi recolhido. Sobre este sobrenadante foi adicionado 90 µL de isopropanol e centrifugado por 15 minutos na máxima rotação da centrífuga. Descartou o sobrenadante e adicionou 100 µL de 70% etanol sob o precipitado. Centrifugou a amostra na máxima rotação da centrífuga por 15 minutos e descartou novamente o sobrenadante, deixou secar bem. A amostra foi adicionada 10 µL de água autoclavada, e então a amostra pode ser armazenada congelando-a a -20°C.

2.12.1.5 Transformação em *E. coli* KC8 e sequenciamento dos clones positivos

O DNA plasmidial purificado conforme descrito acima é utilizado para transformar células de KC8 (*hsdR*, *leuB600*, *trpC9830*, *pyrF::tn5*, *hisB463*, *lacDX74*, *strA*, *galU*,K) por eletroporação e crescidas em meio para selecionar o plástidio pOAD. 50 µL de células KC8 eletrocompetentes são misturadas a 3 µL de DNA e transferidas para cubetas de eletroporação de 0,2 mm (Gene Pulser Cuvette, cat. No. 165-2086) e

submetidas a 100 Ω , 50 μ F e 2,5 kV. Em seguida são adicionados 1mL de meio rico 2xTY e encubada a 37°C, 150 r.p.m por 1 hora. As células são centrifugadas e ressuspendidas em 100 μ L de água autoclavada e plaqueada em meio para KC8 (meio mínimo M9 com 20mg/L de histidina, 40mg/L de triptofano, 20mg/L de uracila, 1,5% de agar) com antibiótico adequado. As células plaqueadas são encubadas a 37°C por dois dias.

As colônias são inoculadas em 2 mL de 2xTY contendo antibiótico adequado a 37 °C por 24 horas e o DNA plasmidial é extraído. Para isto, as células crescidas são centrifugadas a 2700 r.p.m por 8 minutos e ressuspendidas em 240 μ L de solução GET (23 mL de solução 20% glicose, 10 mL de 500mM EDTA pH 8,0, 13 mL de 1M Tris pH 7,4 e completa o volume para 500 mL com água, filtra e armazena a solução na geladeira). As células são centrifugadas a 2700 r.p.m por 8 minutos e o precipitado é ressuspendido em 85 μ L de solução GET + RNase (20 mL de GET com 1mL de 10 mg/mL de RNase), é adicionado 60 μ L de NaOH/SDS (0,75 mL de 4M NaOH, 1,5 mL de 10% SDS e água para completar o volume para 15 mL), as células são misturadas suavemente por inversão e são encubadas a temperatura ambiente por 5 minutos e é dado um pulso de 1000 r.p.m na centrifuga. É adicionado sobre a solução 60 μ L de KOAc (147,2 g de Acetato de potássio, 7,5 mL de ácido acético glacial e água para o volume final de 500 mL, pH 6) e misturado suavemente por inversão e incubado a temperatura ambiente por 10 minutos. Após este período a solução é encubada a 90 °C por 30 minutos e em seguida no gelo por 15 minutos. As célula lisadas são centrifugadas por 15 minutos a 4000 r.p.m e o sobrenadante é aplicado sobre um filtro para a purificação de DNA plasmidial (Multiscreen HTS^{TM}, GV, Millipore, acoplado com placas de fundo V, Solumix), e centrifugado por 3 minutos a 4000 r.p.m, ao centrifugado é adicionado 100 μ L de isopropanol e centrifugado por 40 minutos a 4000 r.p.m, o precipitado é lavado com 70 % etanol e ressuspendido em 20 μ L de água autoclavada. O plasmídio purificado é então submetido à sequenciamento utilizando o oligonucleotídeo específico (5' CAGGGATGTTTAATACCACTACAATG 3') para identificação do gene (presa) fusionado ao "activation domain" de GAL4 do pOAD.

2.12.2 Ensaio de Beta-galactosidase de células de levedura

As células são crescidas em meio SC-WHLA durante a noite, a 30°C sob agitação de 200rpm. 1mL de células é centrifugada a 4°C por 10 min. a 6000 r.p.m. As células são ressuspendidas em 1mL de tampão Z (0.06M de fosfato disódio, 0.04M de

fosfato de sódio, 0.01M de cloreto de potássio, 0.001M de sulfato de magnésio, 0.05M de beta-mercaptoetanol, em pH =7.0) e a absorbância em 600nm é medida. Posteriormente é adicionado 100µL de clorofórmio e 50µL de 0.1% SDS. O meio é misturado e incubado por 5min. a 28 °C. A reação é iniciada com a adição de 25 µL de 4mg/mL de ONPG (o-nitrofenil-beta-D-galactosidase). Deixa reagir por 1hora e pára a reação pela adição de 0.5 mL de 1M de carbonato disódio. O meio é centrifugado por 5 min. a 13000 r.p.m. Mede a absorbância em 550 nm e 420 nm do sobrenadante. O cálculo da Unidade de Miller é feita segundo a equação: $\text{Miller Unit} = 1000X[(\text{OD}_{420} - 1,75X\text{O.D.}_{550})]/(60X1X\text{O.D.}_{600})$ (<http://rothlab.ucdavis.edu/protocols/beta-galactosidase-3.html>).

2.12.3 Produção de anticorpos

Foram feitos anticorpos para as proteínas RpffF, SufE, YaeQ e PilZ. Após a adaptação dos coelhos (fêmeas de cor branca com mais ou menos três meses de idade) por 2 semanas, retirou-se 5mL de sangue, para controle pré-imune, e injetou a primeira amostra (200µg de proteína, 500µL de Freud's adjuvant completo (empresa SIGMA) para um volume final de 1,5mL).

As demais três inoculações foram feitas semanalmente contendo: 200µg de proteína, 500µL de Freud's incompleto (empresa SIGMA) para um volume final de 1,5mL. Depois de uma semana da última inoculação retirou-se 10ml de sangue do coelho. O sangue foi incubado à 37°C por uma hora e as células foram centrifugadas por 15 minutos 4000 rpm, em temperatura ambiente. O sobrenadante (soro) foi dividido em alíquotas de 15µL.

2.12.4 Western Blotting

As amostras a serem analisadas são aplicadas ao gel de SDS-PAGE. Após a corrida as proteínas são transferidas para a membrana de nitrocelulose (Hybond™CEXtra, da empresa Amersham Bioscience) utilizando a cuba de transferência "Trans-Blot SD Semi-dry Transfer cell" da empresa BioRad, em tampão de transferência (5,8g de Tris, 2,9g de glicina, 0,37g de SDS, 200mL de metanol e completa o volume para 1L com água Deionizada) por 30 minutos sob 12V. A transferência é confirmada corando a membrana com corante Ponceau Red (0,2% Ponceau em 1% ácido acético). Lava a membrana com água Deionizada para retirar o corante. A membrana é encubada em tampão de TBS-TT (2,44g/L de Tris, 8g/L de

NaCl, 0,1% de Tween-20 e Triton X-100) e 10% leite em pó desnatado (Molico) sob agitação de 50 rpm por 2 horas. A membrana é lavada três vezes em tampão TBS-TT (cada lavagem é feita por 15 minutos). Após a lavagem a membrana é encubada em meio TBS-TT, 10% leite em pó desnatado (Molico) e 1:16000 de anticorpo da proteína de interesse sob agitação de 50 rpm por 12 horas. O anticorpo é bloqueado com extrato de *E.coli* BL21(DE3) da seguinte forma, adiciona 1mL de TBS-TT, 5 % leite em pó desnatado, 100µL do soro contendo os anticorpos e 500 µL de extrato de *E.coli* BL21(DE3). O extrato de *E.coli* é obtido crescendo 100 mL de 2xTY contendo células de BL21(DE3), a 37 °C, 200 r.p.m até atingir D.O_{600nm} de 2,0, as células são centrifugadas na máxima rotação da centrífuga por 5 minutos a temperatura ambiente e o precipitado é ressuspendido em 2 mL de 2xTY e efetuado a lise por sônicação. As células lisadas são centrifugadas e o sobrenadante é utilizado como extrato de *E.coli*. A membrana é lavada três vezes em tampão TBS-TT (cada lavagem é feita por 15 minutos). Após a lavagem a membrana é encubada em meio TBS-TT, 10% leite em pó desnatado e 1:30000 de proteína A peroxidase (Protein A-peroxidase *Staphylococcus aureus*/ horseradish, Sigma) sob agitação de 50 rpm por 2 horas. A membrana é lavada três vezes em tampão TBS-TT (cada lavagem é feita por 15 minutos). Após a lavagem a membrana é revelada utilizado Kit ECL Advance^{TM} Western Blotting Detection.

2.12.5 Far-Western

O ensaios de interação proteína-proteína *in vitro* por Far-Western foram feitos segundo Hall e colaboradores⁸². Amostras do lisado total de *E. coli* BL21(DE3)STAR induzido com PilB_{XAC3239}, assim como amostras controles do lisado total de *E. coli* BL21(DE3)STAR não induzido de PilB_{XAC3239} e amostra induzida de YaeQ⁸³ foram aplicadas no gel de SDS-PAGE 15% e posteriormente transferida para membrana de nitrocelulose como descrito no item 1.9.4. A membrana é encubada em tampão de TBS-TT e 10% leite em pó desnatado (Molico) sob agitação de 50 rpm por 2 horas. A membrana é lavada três vezes em tampão TBS-TT (cada lavagem é feita por 15 minutos). Após a lavagem a membrana é encubada com 1µM de PilZ_{XAC1133}, PilZ mutantes, PilZ_{XAC1133} com FimX_{XAC2398-GGDEF-EAL} e só FimX_{XAC2398-GGDEF-EAL} em tampão de TBS-TT e 10% leite em pó desnatado (Molico) sob agitação de 50 rpm por 12 horas. A membrana é lavada três vezes em tampão TBS-TT (cada lavagem é feita por 15 minutos). Após a lavagem a membrana é encubada em meio TBS-TT, 10% leite em pó desnatado (Molico) e 1:16000 de anticorpo contra PilZ (devidamente bloqueado),

para detectar PilZ, ou 1:5000 de anticorpo contra anti-cauda de His (Monoclonal Anti-polyhistidine, antibody produced in mouse, Sigma), para detectar a cauda de histidina de FimX_{XAC2398-GGDEF-EAL}, sob agitação de 50 rpm por 2 horas. A membrana é lavada três vezes em tampão TBS contendo 0,1% de Tween-20 e Triton X-100 (cada lavagem é feita por 15 minutos). Após a lavagem a membrana é encubada em meio TBS-TT, 10% leite em pó desnatado e 1:30000 de proteína A peroxidase (Protein A-peroxidase *Staphylococcus aureus*/ horseradish, Sigma) para detectar anticorpo produzido em coelho e anti IgG conjugado a proteína A (Anti-mouse IgG, whole molecule, peroxidase conjugate, Sigma) para detectar o anticorpo contra cauda de histidina, sob agitação de 50 rpm por 2 horas. A membrana é lavada três vezes em tampão TBS-TT (cada lavagem é feita por 15 minutos). Após a lavagem a membrana é revelada utilizando Kit ECL Advance™ Western Blotting Detection.

2.12.6 RMN

2.12.6.1 Ensaio de interação proteína-proteína por ressonância magnética nuclear

Proteínas específicas foram marcadas com ¹⁵N durante sua expressão heteróloga em meio mínimo contendo (¹⁵N)-sulfato de amônia. Espectros ¹H-¹⁵N HSQC destas proteínas livres e em complexo com outras proteínas não marcadas foram adquiridos. Os espectros de ¹⁵N-HSQC dos mutantes de PilZ na presença de FimX_{XAC2398}, FimX_{EAL}, FimX_{EAL-GGDEF} ou FimX_{GGDEF} não marcados foram adquiridos a 27 °C no espectrômetro de 600 MHz com criossonda, exceto PilZ_{Y22A} que foi adquirida no espectrômetro de 500 MHz. As medidas foram obtidas no espectrômetro da Bruker DRX de 500 MHz equipado com sonda de 5mm TXI do Instituto de Química, USP ou pelo espectrômetro da Varian INOVA 600 MHz equipado com criossonda de 5mm HCN PFG do LNLS. Os espectros de RMN foram processados usando o programa NMRPipe⁸⁴ e analisados usando o software CCPN Analysis⁸⁵. Perturbações no espectro foram então analisados. As concentrações das proteínas estão descritas nos Materiais e Métodos do Anexo 2.

2.12.6.2 Assinalamento da cadeia principal

O assinalamento da cadeia principal de PilZ_{XAC1133} foi feito utilizando amostras de PilZ duplamente marcada com ¹³C, ¹⁵N, em 20mM Tris pH 7,0 e 150 mM NaCl. O assinalamento da cadeia principal de PilZ_{XAC1133} foi realizado pela análise dos picos

dos espectros CBCA(CO)NH, HNCACB, HN(CO)CA e HNCA⁸⁶ adquirido a 35 °C e a 27 °C no espectrômetro da Varian INOVA 600 MHz equipado com criossonda de 5mm HCN PFG. As medidas de ¹H-¹⁵N NOE foram feitas a 27 °C no espectrômetro da Bruker DRX, operando a 500 MHz equipado com sonda de 5mm TXI. Os espectros de RMN foram processados usando NMRPipe⁸⁴ e analisados usando CCPN Analysis software⁸⁵. As ressonâncias do assinalamento da cadeia principal de PilZ_{XAC1133} foram depositados no BMRB com número de acesso 16069. O tampão e a concentração de PilZ_{XAC1133} utilizadas para estes ensaios estão descritas nos Materiais e Métodos do Anexo 2.

2.12.7 Ensaio de atividades fosfodiesterase e diguanilato ciclase por HPLC

As atividades fosfodiesterase e diguanilato ciclase foram feitas utilizando a seguinte reação: 50mM Tris pH 7, 250 mM NaCl, 10mM cloreto de magnésio, 5mM beta-mercaptoetanol, 150 µM de GTP ou 50 µM de c-diGMP, e 5 µM de proteína. Estas amostras foram encubadas a 30 °C por diferentes períodos que varia de 24 horas, 7 horas e 2 horas. Após a reação as amostras são fervidas por 20 minutos, centrifugadas na máxima rotação da centrífuga de bancada por 30 minutos e o sobrenadante é aplicado em uma coluna de purificação de PCR da Qiagen. O eluído é então aplicado na coluna de C-18, Germini, Phenomenex, de 250mm x 4,6mm acoplada a uma pré-coluna C-18, Phenomenex, de 4mm x 3mm, via o aparelho HPLC (LC-10AD_{VP} Shimadzu), equilibrada com 2% de tampão B (100 % metanol) e 98 % de tampão A (25 mM formiato de amônio, pH 7). O protocolo de separação utilizado consiste dos seguintes passos ao longo do tempo: 2% de tampão B são aplicados por 7 minutos, em seguida a concentração do tampão B é elevada para 80 % por 28 minutos e por fim a concentração do tampão B é diminuída para 2% até o final da corrida, que é por mais 5 minutos. Amostras padrões, como GMP, GDP, c-diGMP foram aplicadas para determinar o tempo de retenção destas amostras, GTP elui em 6,21 minutos e c-diGMP em 20,62 minutos. Estes valores variam em função da temperatura sendo portanto, necessário repeti-los periodicamente.

2.12.8 Inoculação de *Xac* em folhas de laranja

Amostras de diferentes cepas de *Xac* crescidas em 2xTY na D.O_{600nm} de 0.3 foram infiltradas utilizando seringas descartáveis nas folhas de laranjeiras do tipo Baia doce (*Citrus sinensis* (L.) Osbeck). As plantas foram mantidas na estufa que crescem a

28°C. Após 4 dias pode-se observar os primeiros sintomas da doença e após 10 dias tiramos fotos das folhas inoculadas. Estes experimentos foram feitos no Laboratório de Biotecnologia da EMBRAPA localizado em Cordeirópolis em São Paulo.

2.12.9 Ensaio de aderência de *Xac* em folhas de laranja

Diferentes cepas de *Xac* foram crescidas em SB, BASE, XVM2, Eiken e STON por 12 horas, sob agitação de 200 r.p.m a 28°C. Estas amostras foram centrifugadas e lavadas com água para retirar o meio de cultura residual e ressuspendidas em 10mM tampão fosfato pH 7,0. Foi adicionado 20 µl do ressuspendido, em que todas as amostras continham a mesma D.O_{600nm}, na superfície abaxial de folhas de laranjeiras do tipo Baia doce (*Citrus sinensis* (L.) Osbeck). As folhas foram incubadas em uma vasilha umidificada por 6 h a 28°C. A aderência das células na superfície das folhas foram monitoradas pela utilização do corante cristal violeta (1%) que foi adicionado sobre as folhas por 15 minutos e posterior lavagem com água. Apenas as células aderidas se mantêm coradas.

2.12.10 Ensaio de motilidade de diferentes cepas de *Xac* em meio semi-sólido.

Diferentes cepas de *Xac* foram crescidas em 2mL de 2xTY na presença de 200µg/mL de ampicilina, sob agitação de 200 r.p.m a 28 °C até D.O. de 0.3. São adicionados 5µL de células na superfície do meio semi-sólido que pode ser SB (5g/L de sacarose, 5g/L de extrato de levedura, 5g/L de peptona, 1g/L de ácido glutâmico, 0.6 % de bacto agar, pH 7.0), STON (1g/L de Triton X-100 a, 0.3g/L de extrato de levedura, 0.3 % de bacto agar, pH 7.0), Eiken (8g/L de Eiken nutriente, 6g/L de Eiken agar, pH 7.0), meio Base (0.5 % extrato de levedura, 100mM de fosfato de potássio pH 7,0, 0,6% bacto agar) e XVM2 (20mM NaCl, 10mM sulfato de amônia, 1mM cloreto de cálcio, 0,01mM sulfato de ferro, 0,16mM dihidrogenofosfato de potássio, 0,16mM hidrogenofosfato de dipotássio, 10mM frutose, 10mM sacarose, 0,03% casaminoácidos, 0.6% bacto agar e pH = 7). Posteriormente, as placas foram incubadas na estufa a 28°C por 2-4 dias e é então observada a diferença de dispersão das diferentes cepas de *Xac* na superfície da placa.

2.12.11   *Dicroísmo Circular e desnaturação térmica*

As medidas de desnaturação térmica foram feitas a $[\theta]_{205\text{nm}}$, para PilZ, e $[\theta]_{220\text{nm}}$ para FimX. As medidas dos mutantes de PilZ foram feitas no CD J-710 Circular Dichroism Spectropolarimeter, Jasco Inc. (LNLS, Campinas, Brasil), enquanto que as medidas de XAC2398^{GGDEF-EAL} foram feitas no CD J-715 Circular Dichroism Spectropolarimeter, Jasco Inc. (IQ-USP, São Paulo, Brazil). A fração da proteína não enovelada e enovelada foram calculadas a partir da curva do estado de pré transição e pós transição como descrito⁸⁷. As medidas foram feitas usando 3,5-10 μM de proteína em 10 mM Tris-Cl, pH 7.0, 10 mM NaCl e 1 mM MgCl_2 . Os ligantes c-diGMP, GTP ou ATP foram adicionados para a concentração final de 5 μM . O espectro foi adquirido na faixa de 195-260 nm e obtido a média de 8 medidas a 20°C, usando 1 mm de caminho ótico em cubas de quartzo, velocidade de 60 nm/min, tempo de resposta de 8 segundos e abertura de 1 nm.

2.12.12   *Microscopia eletrônica de transmissão*

Os ensaios de microscopia eletrônica de transmissão foram feitos utilizando diferentes cepas de *Xac* crescidas em diferentes meios semi-sólido. Foram utilizados apenas as bactérias localizadas na borda das colônias que foram transferidas para filmes de carbono e fixadas com 1% glutaraldeído. As amostras foram analisadas usando o aparelho Zeiss EM 900 transmission electron microscope no laboratório do Elliot W. Kitajima (Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, SP, Brasil).

2.12.13   *Ensaio de resistência a rifampicina para medir a taxa de mutação espontânea*

Células de diferentes cepas de *Xac* são crescidas em meio líquido 2xTY contendo 200 $\mu\text{g/mL}$ de ampicilina a 28 °C, 200 r.p.m até D.O de 0.8. 10 μL desta solução é diluída em 1 mL de água autoclavada e 100 μL são plaqueados em LB agar contendo diferentes concentrações de rifampicina (1 - 50 $\mu\text{g/mL}$ de rifampicina) . As amostras são encubadas a 30 °C, envolvidas em papel alumínio, por 4 dias.

2.12.14 **Gel shift**

Amostras de DNA de regiões promotoras dos genes XAC0136 e XAC4180 marcadas com $^{32}\alpha$ dATP (fornecidas pela aluna de doutorado Maria Cláudia Pereira Rosa) foram utilizadas para os ensaios de gel-shift. Foram preparadas quatro amostras utilizando duas concentrações de DNA (1 μ M e 10nM) e duas amostras de DNA marcadas com $^{32}\alpha$ dATP. As amostras foram preparadas adicionando 2 μ L de Tampão (50mM Tris pH 8.0, 3mM cloreto de magnésio, 10mM cloreto de potássio, 1mM DTT, 100 μ g/mL de BSA), 1,5 μ L de DNA marcado (na concentração de 13,3 μ M e 0,13 μ M), 2 μ L de 10 μ M de proteína (YaeQ ou YaeQ_{D68A} ou YaeQ_{D83A}) e 14,5 μ L de água. Amostras controles foram também preparadas da mesma forma como descrito anteriormente, mas adicionando água ao invés de proteína. Estas amostras foram aplicadas em gel de 4,5 % acrilamida (1,5 mL de TBE (10,5g/L de Tris, 5,5g/L de ácido bórico, 4 mL de 500mM EDTA pH 8), 4,5mL de 30 % acrilamida, 1,9 mL de 40% glicerol, 15 μ L de TEMED, 285 μ L de APS e 21,8 mL de água), em tampão TB (10,5g/L de Tris, 5,5g/L de ácido bórico) por 3 horas e 30 minutos a 200V. O gel é secado por 2 horas utilizando o aparelho (Model 583 Gel Dryer, Biorad) e revelado utilizando o scanner Storm (Molecular Dynamics) no modo de "Storage Phosphor mode red 635 nm/390 BP, 900V: Normal sensitivity".

As amostras de DNA simples fita foram feitas utilizando o protocolo descrito acima, mas antes de adicionar proteína a solução foi fervida por 20 minutos e colocada em seguida no gelo e posteriormente adicionado a proteína. O mesmo procedimento foi feito para obter DNA mutato por raios ultra violeta (UV), a solução de reação antes de colocar a proteína foi encubada no gelo e tratada com 1500J/cm² de raios UV por 17 minutos. Após este tratamento foi adicionado amostra de proteína e deixado por 1 hora a temperatura ambiente antes de aplicar no gel de acrilamida.

2.12.15 **Atividade Cisteína dessulfarase**

O ensaio de cisteína dessulfarase foi feito a 30 °C por 3 horas como descrito por Outten, *et al.*⁸⁸, em que a proporção de SufE/SufS foi variada de 0:1 até 20:1. De forma resumida, a reação foi feita em 160 μ L contendo 25mM de Tris pH 7,4, 100mM de NaCl, 2,5 μ M de SufE e SufS purificadas (proporção de 1:1), 10 μ M de pirodoxal 5'-fosfato, 1mM DTT e 500 μ M de cisteína (Sigma). A adição de 20 μ L de 20mM N,N-dimetil-p-fenilenediamina em meio ácido (diluída em 7.2 M HCl) e a adição de 20 μ L de 30 mM FeCl₃ (diluído em 1.2 M HCl) forma azul de metileno (C₁₆H₁₈ClN₃S) que é

medido pela absorbância no comprimento de onda de 670 nm^{89,90}. A adição de ácido ao meio faz com que os átomos de enxofre sejam reduzidos e liberados ao meio, que na presença de N,N-dimetil-p-fenilenediamina e FeCl₃ forma azul de metileno⁹⁰.

2.12.16   *Ensaio de resistência a peróxido de hidrogênio e quelante de ferro.*

O teste realizado para averiguar o nível de resistência de diferentes cepas de *Xac* a agentes oxidantes como peróxido de hidrogênio foi feito crescendo diferentes cepas de *Xac* em 2mL de meio líquido 2xTY contendo 200µg/mL de ampicilina a 28 °C, 200 r.p.m até D.O_{600nm} de 0.3. A cada 100 µL deste meio de cultura foram adicionados 1 mM de peróxido de hidrogênio e incubado em diferentes tempos a 30 °C: 0, 5, 30 minutos e 2 horas, após este período as células foram plaqueadas em meio LB contendo 200µg/mL de ampicilina e verificado a sobrevivência das células⁹¹.

O teste realizado para averiguar o nível de resistência de diferentes cepas de *Xac* ao estresse causado pela ausência de ferro foi feito utilizando diferentes cepas de *Xac* crescidas em 2mL de 2xTY até D.O_{600nm} de 0.3. Foram retirados 100 µL deste meio de cultura e plaqueados em meio LB contendo 200µg/mL de ampicilina e diferentes concentrações de 2-2'diperidil (0, 5, 20, 100 e 200µM). As amostras foram encubadas a 30 °C por 2 dias⁹¹ e analisado o nível de sobrevivência das células. O mesmo ensaio foi realizado para analisar o nível de sobrevivência das células quando crescidas em meio contendo excesso de ferro. Para isso 100 µL diferentes cepas de *Xac* crescidas em 2xTY em D.O_{600nm} de 0.3 foram plaqueados em meio LB contendo 200µg/mL de ampicilina e diferentes concentrações de FeCl₃ (100µM e 300µM de FeCl₃ diluído em 1.2M HCl)⁹¹.

CAPÍTULO 3
***Interações entre PilZ, FimX e PilB, proteínas reguladoras da
biogênese do Pilus tipo IV***

3.1 Introdução

3.1.1 Pilus tipo IV

Filamento do pilus tipo IV. O pilus tipo IV (T4P) é um filamento fino com diâmetro entre 60-80Å, longo (vários microns), flexível e bastante resistente e é encontrado na superfície de células bacterianas⁹². O T4P está envolvido em diferentes comportamentos bacterianos incluindo, por exemplo, movimento *twitching*, adesão em superfícies, transformação natural, formação de biofilme, secreção de proteínas e quimiotaxia⁹³⁻⁹⁸.

Proteínas envolvidas na biogênese do T4P. A biogênese do T4P é um processo complexo que envolve uma grande quantidade de proteínas. Em *Pseudomonas aeruginosa* (*P. aeruginosa*) foram descritas mais de 40 proteínas envolvidas no seu funcionamento, algumas no processo de expressão gênica, outras como componentes estruturais e outros que atuam como quimiosensores na regulação de processos bacterianos a partir de sinais extracelulares^{5,93-98}. Algumas destas proteínas que possuem similaridade de seqüência primária com genes de *Xac* são codificadas por genes localizados em cinco *loci* cromossômicos principais, ilustrados na **Figura 3**.

Figura 3: Pilus tipo IV em *Xac*.

Organização gênica dos genes envolvidos na produção do T4P em *Xac*. As letras em maiúscula, acima das setas, representam de forma simplificada os nomes das proteínas que estes genes, com numeração abaixo de cada seta, codificam⁹⁹⁻¹⁰¹. Por exemplo, XAC1133 codifica a proteína PilZ que está representada com a letra Z. PilZ_{XAC1133} está destacada por ser alvo do estudo.

Pilinas. Em *P. aeruginosa* o filamento do T4P é composto por um polímero helicoidal formado principalmente por PilA⁹³, a pilina majoritária, e por pilinas minoritárias, PilE, PilV, PilW, PilY1 e PilY2 que são importantes para a formação da base e da ponta do filamento¹⁰² (**Figura 4**). Algumas destas pilinas possuem similaridade de seqüência primária entre elas e a sua correta anotação se torna difícil. Como em *P. aeruginosa*, em *Xac* a pilina majoritária, *pilA*, está localizada no *loci* onde se encontram os genes *pilR*, *pilS*, *pilB*, *pilC* e *pilD*, fazendo com que a sua identificação seja mais evidente⁹³.

Xac possui dois genes que codificam a PilA, XAC3240 e XAC3241 (**Figura 3**), que têm alta similaridade de seqüência primária entre eles sugerindo que estes vieram de uma duplicação gênica e, portanto, são parálogos e devem possuir funções similares¹⁰³. Interessantemente, XAC3241 possui duas lisinas depois da metionina inicial, sugerindo que, pela regra do “N-end rule”, XAC3241 possa ser rapidamente degradada no citosol¹⁰⁴.

Funcionamento do T4P. Tanto PilA quanto PilE possuem uma seqüência consenso, localizada na região N-terminal, composta por 23 resíduos, no qual, PilD, a prepilina peptidase, é responsável pelo reconhecimento, clivagem e metilação do resíduo fenilalanina da seqüência G^AFxxxE^{105;106}. Após o processamento, as pilinas são polimerizadas por PilB via consumo de ATP com o auxílio de PilC, proteína localizada na membrana interna¹⁰⁷. O filamento do T4P sendo formado atravessa a membrana externa através de poros formados por hexâmeros de PilQ (Q)¹⁰⁸. Este filamento é posteriormente despolimerizado por PilT, via consumo de ATP¹⁰⁹.

PilB, assim como PilT, formam estruturas hexaméricas associadas na face citosólica da membrana interna bacteriana, e são responsáveis pelo funcionamento dos ciclos de polimerização e despolimerização do filamento do T4P via consumo de ATP, respectivamente, (**Figura 4**)¹¹⁰. Estes ciclos compõem o funcionamento geral do T4P em vários processos biológicos, como por exemplo, no processo de autotransformação, em que o filamento do T4P interage com DNA do meio externo e o internaliza quando o filamento é despolimerizado⁹³⁻⁹⁸. Outro exemplo é no processo de movimentação *twitching*, em que as adesinas, pilinas minoritárias, localizadas na ponta do T4P, aderem na superfície (substratos sólidos ou outras células) e quando há a despolimerização do filamento a bactéria anda em direção ao topo do filamento⁹³⁻⁹⁸.

A biogênese do T4P em *P. aeruginosa* é modulada por FimX (entre outros reguladores) em resposta a fatores ambientais que possivelmente são sentidos de

PilA no periplasma³. No entanto, até hoje não se sabe o papel direto de PilZ_{PA2960} no processo de biogênese do pilus tipo IV.

Diferentemente do que é observado em *P. aeruginosa*, células de *N. meningitidis* nocautes ao gene *pilZ* possui pilus na membrana externa, mas não tem a capacidade de formar agregados¹¹³. Células de *Xanthomonas campestris* pv *campestris* nocautes ao gene *pilZ* tem uma pequena redução na dispersão das células em meio semi-sólido contendo Eiken agar quando comparadas a células tipo selvagem¹¹⁴. Estas diferenças no fenótipo de células nocautes de *pilZ* sugerem que a função de PilZ na produção de T4P possa ser modulada por vários fatores ambientais que desencadeiam diferentes respostas dependendo do organismo.

Alguns membros da família de PilZ interagem com c-diGMP. Com o sequenciamento de centenas de genomas bacterianos foram identificados mais de 1500 proteínas contendo domínios ortólogos à PilZ, sendo que a maioria deles possuem baixa similaridade de seqüência primária. Por este motivo, a família de PilZ foi, posteriormente, dividida em 28 grupos ortólogos e, mais importante ainda, membros de alguns destes grupos ortólogos demonstraram a capacidade de ligar ao segundo mensageiro bacteriano *bis*(3'→5')GMP cíclico (c-diGMP)¹¹⁵⁻¹¹⁸. Isso gerou um grande interesse científico para estudar a correlação entre c-diGMP e sua interação com PilZ, já que existem poucos receptores de c-diGMP conhecidos^{119,117}.

A família de PilZ é caracterizada por ser um domínio com aproximadamente 120 aminoácidos e contem 10 resíduos conservados, sendo que os cinco primeiros, RxxxR e D/NxSxxG, localizados no N-terminal, estão diretamente relacionados com a capacidade de membros desta família de interagir com c-diGMP. Mutações nestes resíduos abolem a capacidade da proteína, como VC0042 de *Vibrio cholerae* e PA4608 de *P. aeruginosa*, de interagirem com c-diGMP^{119; 117}. Os domínios PilZ podem ser encontrados sozinhos ou associados com outros domínios regulatórios tais como: REC, MCP, domínios que ligam AMP, PKinase, domínios que ligam ao fator sigma-54, PAS, EAL, HD-GYP e GGDEF, sendo que os últimos três domínios, GGDEF e EAL/HD-GYP, estão envolvidos na síntese e degradação de c-diGMP, respectivamente^{115,120}.

3.1.3 Sinalização por c-diGMP.

A maioria das espécies bacterianas podem viver como unidades independentes ou como membros agregados de uma colônia. Ambos os estilos de vida oferecem

vantagens e desvantagens, dependendo das circunstâncias ambientais. Por exemplo, para o fitopatógeno *Xanthomonas campestris* pv *campestris* (Xcc), a formação de colônias é importante para a sua sobrevivência na superfície da planta, enquanto que o crescimento solitário é necessário para o desenvolvimento da doença no hospedeiro¹²¹. Esta decisão de vida comunitária vs solitária está muitas vezes correlacionada a outros processos fisiológicos importantes, como, por exemplo, motilidade, adesão a superfícies, produção de fatores de virulência, secreção de exopolissacarídeos e enzimas, transformação natural, e *quorum sensing*¹²²⁻¹²⁴.

Nos últimos anos emergiu a percepção que aparentemente há um mecanismo universal de transdução de sinais dentro das eubactérias para controlar estes fenótipos correlacionados. No centro deste mecanismo existem vias de transdução de sinal mediadas pela molécula c-diGMP^{41,115,125-130} ilustrada na **Figura 5**. Somente para citar alguns exemplos, já foi demonstrado que o c-diGMP pode desempenhar papéis centrais na i) produção de polissacarídeos extracelulares em *Salmonella*, *P. aeruginosa*, *Yersinia*, *Xanthomonas* e *Vibrio cholerae*, ii) motilidade em *Salmonella*, *P. aeruginosa*, *E. coli*, *Caulobacter crescentus* e *Xanthomonas* e iii) virulência em *Salmonella*, *Vibrio cholerae* e *Xanthomonas campestris*.

O c-diGMP é sintetizado a partir de duas moléculas de GTP por uma classe de enzimas que contém domínios GGDEF e é hidrolisado à pGpG ou a duas moléculas de GMP por proteínas que possuem domínios EAL ou HD-GYP, respectivamente^{41,123,131,132}. Os nomes dos domínios foram originados dos motivos de aminoácidos conservados presentes no sítio ativo dessas enzimas^{39,41,127,130}. Domínios GGDEF, EAL e HD-GYP foram encontrados literalmente em milhares de proteínas bacterianas: por exemplo, o banco de dados Pfam contém > 10 mil seqüências GGDEF e quase 6 mil seqüências EAL. Além disso, estão sendo caracterizados domínios protéicos e até domínios de RNA (“riboswitches”) que ligam c-diGMP e, por isso, podem atuar como receptores^{4,115,133}. Os receptores protéicos de c-diGMP podem ser derivados de domínios GGDEF ou EAL não-catalíticos (chamados domínios “degenerados”) ou podem pertencer a outras famílias protéicas, das quais uma é a superfamília de domínios PilZ, ubíquo em eubactérias, como descrito acima.

Figura 5: Estrutura do segundo mensageiro bacteriano c-di-GMP¹³⁴.

O c-diGMP, *bis*(3'→5')GMP cíclico, é composto por duas moléculas de GMP em que o 3'-OH de cada ribose faz uma ponte fosfodiéster com o fosfato 5' do açúcar do outro nucleotídeo.

3.2 Objetivos

Este trabalho teve como objetivo o estudo estrutural e funcional de PilZ_{XAC1133}, ortóloga ao PilZ_{PA2690} de *P. aeruginosa*. Foram utilizadas diferentes técnicas de biologia estrutural que incluem cristalografia de raios-X e RMN. Além dos estudos estruturais, ensaios funcionais foram realizados, a fim de compreender com mais detalhe a correlação da estrutura de PilZ com a sua função biológica. Para isto foram empregados várias técnicas, tais como: mutação sítio dirigida, produção de nocautes gênicos, ensaios de 2-híbrido e outros ensaios de interação proteína-proteína.

3.3 Resultados

3.3.1 Determinação da Estrutura de PilZ_{XAC1133} pela técnica MAD.

Purificação e cristalização: PilZ_{XAC1133} cristalizou em mais de dez condições diferentes. No entanto, na maioria das condições de cristalização os cristais de PilZ_{XAC1133} eram poli cristais¹³⁵. Em apenas uma condição (vide Materiais e Métodos do manuscrito em **Anexo 1**)¹³⁶ PilZ_{XAC1133} formou cristais que forneceu dados para resolver as fases e determinar a sua estrutura tridimensional.

Tentativas para resolver as fases de cristais nativos de PilZ_{XAC1133} pela técnica de substituição molecular não tiveram êxito quando utilizamos como modelo as estruturas de PilZ disponíveis nos bancos de dados. Por isso, resolvemos adotar a técnica MAD (multi-wavelength anomalous dispersion) para determinar a estrutura de PilZ_{XAC1133}. Para isso, produzimos PilZ_{XAC1133} em meio mínimo contendo selênio-metionina (SelMet) com o objetivo de obter todas as cinco metioninas da proteína de PilZ_{XAC1133} na forma de SelMet. Esta incorporação é essencial para determinar as fases do cristal da proteína utilizando a técnica MAD.

PilZ_{XAC1133} contendo selênio metionina foi purificada e cristalizada. A partir do espectro de emissão de fluorescência de raios-X dos átomos de selênios do cristal de PilZ_{XAC1133} foram obtidos os valores de comprimento de onda para a coleta dos conjuntos de dados do “peak”, “inflection” e “remote” (**Figura 6**). Os conjuntos de dados foram coletados na linha MX2 no LNLS em Campinas e foram processados, integrados e escalonados utilizando o programa HKL2000⁶⁵. Neste processamento, as reflexões dos pares de Friedel foram tratadas como reflexões independentes. Os três conjuntos de dados foram juntados e escalonados pelos programas, CAD e Scaleit da interface gráfica ccp4i⁶⁷. A estatística da coleta e do processamento dos dados estão descritos no manuscrito de cristalização de PilZ_{XAC1133}, no **Anexo 1**¹³⁶.

Figura 6: Espectro de excitação de raios-X em que se mede a fluorescência de raios-X do cristal de $\text{PilZ}_{\text{XAC1133}}$ contendo selênio-metinina.

a: Espectro de fluorescência de raios-X dos átomos de selênio do cristal de $\text{PilZ}_{\text{XAC1133}}$ (dados obtidos pelo programa Blu-Ice^{137}); **b:** Curvas dos valores dos coeficientes de espalhamento anômalo f' e f'' calculadas a partir dos valores de emissão de fluorescência dos átomos de selênio do cristal de $\text{PilZ}_{\text{XAC1133}}$, indicados na **Tabela 1** do **Anexo 1** (calculado pelo programa Chooch^{63}).

Anexo 1 – Expressão, cristalização e análise cristalográfica preliminar de PilZ_{XAC1133} de *Xanthomonas axonopodis* pv *citri*

- Expression, crystallization and preliminary crystallographic analysis of PilZ_{XAC1133} from *Xanthomonas axonopodis* pv. *citri*

crystallization communications

Acta Crystallographica Section F
Structural Biology
and Crystallization
Communications

ISSN 1744-3091

Cristiane R. Guzzo and Chuck S. Farah*

Departamento de Bioquímica, Instituto de
Química, Universidade de São Paulo,
05508-000 São Paulo-SP, Brazil

Correspondence e-mail: chsfarah@iq.usp.br

Received 17 November 2008
Accepted 16 February 2009

Expression, crystallization and preliminary crystallographic analysis of PilZ_{XAC1133} from *Xanthomonas axonopodis* pv. *citri*

Proteins containing PilZ domains are widespread in Gram-negative bacteria and have recently been shown to be involved in the control of biofilm formation, adherence, aggregation, virulence-factor production and motility. Furthermore, some PilZ domains have recently been shown to bind the second messenger bis(3'→5')cyclic diGMP. Here, the cloning, expression, purification and crystallization of PilZ_{XAC1133}, a protein consisting of a single PilZ domain from *Xanthomonas axonopodis* pv. *citri*, is reported. The closest PilZ_{XAC1133} homologues in *Pseudomonas aeruginosa* and *Neisseria meningitidis* control type IV pilus function. Recombinant PilZ_{XAC1133} containing selenomethionine was crystallized in space group *P*6₁. The unit-cell parameters were *a* = 62.125, *b* = 62.125, *c* = 83.543 Å. These crystals diffracted to 1.85 Å resolution and a MAD data set was collected at a synchrotron source. The calculated Matthews coefficient suggested the presence of two PilZ_{XAC1133} molecules in the asymmetric unit.

1. Introduction

Bacteria from the genus *Xanthomonas* infect many plant hosts, including crops of economic importance, and *Xanthomonas axonopodis* pv. *citri* (Xac) causes citrus canker in a wide variety of citrus species (Brunings & Gabriel, 2003). The sequencing of the Xac genome (Da Silva *et al.*, 2002) revealed the existence of a large number of proteins containing GGDEF, HD-GYP, EAL and PilZ domains (Andrade *et al.*, 2006). These domains have been implicated in the production (GGDEF), degradation (HD-GYP and EAL) and binding (PilZ) of the second messenger bis(3'→5')cyclic diGMP (c-diGMP; Paul *et al.*, 2004; Simm *et al.*, 2004; Ryan *et al.*, 2006; Pratt *et al.*, 2007). C-diGMP signalling is now recognized as extremely important for a variety of complex bacterial phenomena including biofilm formation, virulence, motility, pilus expression and photosynthesis (Römling *et al.*, 2005). The only high-resolution structures of PilZ domains available are the NMR structure of the *Pseudomonas aeruginosa* protein Pa4608 (Ramelot *et al.*, 2007) and crystal structures of the PilZ domain-containing protein VCA0042 from *Vibrio cholerae* in the presence and absence of c-diGMP (Benach *et al.*, 2007).

PilZ domains typically contain around 120 residues and make up a superfamily that is highly diverse at the level of primary structure and that can be divided into many paralogous groups (Amikam & Galperin, 2006; Benach *et al.*, 2007). The Xac genome codes for four proteins with PilZ domains, one of which is XAC1133, or PilZ_{XAC1133}, which is coded by the *pilZ* gene (Da Silva *et al.*, 2002). The closest PilZ_{XAC1133} homologues in *P. aeruginosa* and *Neisseria meningitidis* control type IV pilus (T4P) function, although perhaps in different ways. In *P. aeruginosa*, PilZ is necessary for T4P biogenesis and TP4-dependent twitching motility (Alm *et al.*, 1996). On the other hand, a *pilZ* knockout in *N. meningitidis*, although piliated, is not able to produce T4P-dependent bacterial aggregates (Carbannelle *et al.*, 2005).

In this report, we describe the cloning and expression of recombinant PilZ_{XAC1133} containing selenomethionine and the collection of

© 2009 International Union of Crystallography
All rights reserved

crystallization communications

a multiwavelength anomalous dispersion (MAD) data set that is being used to determine its three-dimensional structure.

2. Cloning and expression of PilZ_{XAC1133}

PilZ_{XAC1133} was amplified by PCR from Xac genomic DNA (Da Silva *et al.*, 2002) using the primers 5'-TTCTCCATGGACCATATGAGT-GCAATGAAT-3' and 5'-CGGGATCCCTCGAGTTACATCGTAT-GCGTCG-3'. The PCR product codes for the full-length protein (117 amino acids; 12.3 kDa). It was digested with *Nde*I and *Bam*HI and subcloned into the pET-3a expression vector (Studier *et al.*, 1990) previously digested with the same endonucleases. PilZ_{XAC1133} was expressed in *Escherichia coli* strain BL21 (DE3) (Studier *et al.*, 1990). Selenomethionine-containing protein was produced by growing a 500 ml culture in M9 medium to an OD_{600 nm} of 0.8 at 310 K, at which point 100 mg l⁻¹ lysine, 100 mg l⁻¹ phenylalanine, 100 mg l⁻¹ threonine, 50 mg l⁻¹ isoleucine, 50 mg l⁻¹ valine and 60 mg l⁻¹ selenomethionine were added (Berne *et al.*, 1999). After 15 min, 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added and the cells were grown for 4 h before sedimentation and storage at 193 K.

3. Protein purification

Cells from a 1 l culture were resuspended in 25 ml 50 mM Tris-HCl pH 8.0, 25% sucrose, 1 mM EDTA and lysed using a French press. The soluble fraction was applied onto an SP-Sepharose Fast Flow (FF) HiLoad 16/10 column (Amersham Pharmacia) previously equilibrated with 50 mM Tris-HCl pH 8.0 and 1 mM EDTA. Bound proteins were eluted using a 0–1 M NaCl gradient over 12 column volumes. The fractions containing the purest samples of PilZ_{XAC1133} were pooled. In spite of the fact that the purification only involved one chromatographic step, the elevated pI value of the protein (theoretical value of 9.05) allowed us to obtain an adequate number of fractions that had very little contaminant visible in an overloaded Coomassie-stained SDS polyacrylamide gel. The purity, as estimated by visual inspection of the gel, was above 95%. The protein solution was then dialyzed against 5 mM Tris-HCl pH 7.0 and concentrated to 5–10 g l⁻¹ using Centricon (Millipore) concentrators with a 3 kDa membrane cutoff.

4. Crystallization

Initial crystallization conditions were screened by the sparse-matrix sampling approach using Crystal Screen and Index Screen (Hampton Research) matrices. Initially, crystals were obtained under several conditions of vapour diffusion using the sitting-drop technique at 291 K. Optimization of the crystallization conditions was then achieved by modification of the concentration of the precipitating reagent, the buffer pH and the temperature. The best crystals were obtained by mixing equal volumes (1.5 µl) of a 7.2 mg ml⁻¹ protein solution in 5 mM Tris-HCl pH 7.0 with reservoir solution consisting of 24% (w/v) PEG 4000, 0.1 M Tris-HCl pH 8.0 and 0.2 M MgCl₂. Crystals were initially grown at 281 K for one month, which was followed by transfer to 291 K to obtain mature-sized crystals (Fig. 1). The crystal was frozen immediately before data collection in a stream of nitrogen at 100 K. No cryoprotectant was used.

5. Data collection

Data were collected on the W01B-MX2 beamline of the Laboratório Nacional de Luz Síncrotron, Campinas, São Paulo using a MAR

Table 1

Crystal parameters and data-reduction statistics.

Data sets were measured from a single crystal. Values in parentheses are for the highest resolution shell. Each member of a Friedel pair was counted as a separate reflection.

	Peak	Peak	Inflection	Remote
Space group	<i>P</i> 6 ₁	<i>P</i> 6 ₁ 22	<i>P</i> 6 ₁ 22	<i>P</i> 6 ₁ 22
Unit-cell parameters				
<i>a</i> (Å)	62.125	62.125	62.168	62.202
<i>b</i> (Å)	62.125	62.125	62.168	62.202
<i>c</i> (Å)	83.543	83.542	83.605	83.658
Resolution range (Å)	40.00–1.85	40.00–1.86	40.00–1.95	40.00–1.95
	(1.92–1.85)	(1.93–1.86)	(2.02–1.95)	(2.02–1.95)
No. of observed reflections	328014	323942	211604	273081
No. of unique reflections	30576	15168	13219	13262
<i>I</i> (σ(<i>I</i>))	24.7 (1.74)	34.01 (2.60)	33.63 (7.46)	25.6 (5.3)
Multiplicity	10.7 (5.9)	21.4 (11.9)	16.0 (15.9)	20.6 (20.4)
Completeness (%)	99.0 (92.1)	99.7 (97.8)	99.8 (100.0)	99.9 (100.0)
<i>R</i> _{merge} † (%)	8.6 (60.1)	8.7 (55.8)	8.6 (43.1)	11.6 (62.5)
No. of images	360	360	248	319
Oscillation angle (°)	1	1	1	1
Wavelength (Å)	0.97829	0.97829	0.97818	1.03448
<i>f</i> / <i>f</i> ''	6.43/–7.81	6.43/–7.81	3.56/–10.39	3.60/–2.80

$$\dagger R_{\text{merge}} = \sum_{\text{all}} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{\text{all}} \sum_i I_i(hkl).$$

Mosaic 225 CCD detector. Crystals were flash-frozen and maintained at 100 K in a stream of cold nitrogen gas during measurement. MAD data sets were collected using a single crystal at the three wavelengths 0.97829, 0.97818 and 1.03448 Å, corresponding to peak, inflection and remote points of the fluorescence spectrum of the PilZ_{XAC1133} crystal, respectively. The *f*' and *f*'' anomalous scattering factors shown in Table 1 were estimated from the fluorescence spectrum using the *CHOOCH* software (Evans & Pettifer, 2001). Diffraction intensities for the data sets were integrated and scaled using the programs *DENZO* and *SCALEPACK*, respectively (Otwinowski & Minor, 1997). The Friedel mates were scaled separately during data processing to calculate the anomalous signal.

The crystal diffracted to 1.85 Å resolution in space group *P*6₁. We note that the MAD data set was initially processed in space group *P*6₁22 (Table 1) and these data were used to obtain initial estimates of phases and to initiate model building (to be published elsewhere). However, refinement convergence could not be achieved without reducing the space-group symmetry to *P*6₁ (to be published elsewhere). For this reason, Table 1 shows statistics for the processing of the MAD data set in *P*6₁22 for all three wavelengths and the statistics for the peak data set after reprocessing in *P*6₁. There are two

Figure 1
Crystal of PilZ_{XAC1133}, with approximate dimensions of 0.15 × 0.1 × 0.10 mm.

crystallization communications

PilZ_{XAC1133} monomers per P6₁ asymmetric unit (Matthews coefficient $V_M = 1.9 \text{ \AA}^3 \text{ Da}^{-1}$) and the estimated solvent content is 33.8%.

We thank Lucas Sanfelici, Walan Grizolli, Beatriz Guimarães and João Alexandre R. G. Barbosa of the Laboratório Nacional de Luz Síncrotron for technical help and useful discussions. This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, Grant 2005/59243-3) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

References

- Alm, R. A., Boder, A. J., Free, P. D. & Mattick, J. S. (1996). *J. Bacteriol.* **178**, 46–53.
- Amikam, D. & Galperin, M. Y. (2006). *Bioinformatics*, **22**, 3–6.
- Andrade, M. O., Alegria, M. C., Guzzo, C. R., Docena, C., Rosa, M. C., Ramos, C. H. & Farah, C. S. (2006). *Mol. Microbiol.* **62**, 537–551.
- Brunings, A. M. & Gabriel, D. W. (2003). *Mol. Plant Pathol.* **4**, 141–157.
- Benach, J., Swaminathan, S. S., Tamayo, R., Handelman, S. K., Folta-Stogniew, E., Ramos, J. E., Forouhar, F., Neely, H., Seetharaman, J., Camilli, A. & Hunt, J. F. (2007). *EMBO J.* **26**, 5153–5166.
- Berne, P. F., Doublé, S. & Carter, C. W. Jr (1999). *Crystallization of Nucleic Acids and Proteins: A Practical Approach*, 2nd ed., edited by A. Ducruix & R. Giegé, pp. 45–73. Oxford University Press.
- Carbonnelle, E., Hélaine, S., Prouvensier, L., Nassif, X. & Pelicic, V. (2005). *Mol. Microbiol.* **55**, 54–64.
- Da Silva, A. C. R. *et al.* (2002). *Nature (London)*, **417**, 459–463.
- Evans, G. & Pettifer, R. F. (2001). *J. Appl. Cryst.* **34**, 82–86.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Paul, R., Weiser, S., Amiot, N. C., Chan, C., Schirmer, T., Giese, B. & Jenal, U. (2004). *Genes Dev.* **18**, 715–727.
- Pratt, J. T., Tamayo, R., Tischler, A. D. & Camilli, A. (2007). *J. Biol. Chem.* **282**, 12860–12870.
- Ramelot, T. A., Yee, A., Cort, J. R., Semesi, A., Arrowsmith, C. H. & Kennedy, M. A. (2007). *Proteins*, **66**, 266–271.
- Römling, U., Gomelsky, M. & Galperin, M. Y. (2005). *Mol. Microbiol.* **57**, 629–639.
- Ryan, R. P., Fouhy, Y., Lucey, J. F., Crossman, L. C., Spiro, S., He, Y. W., Zhang, L. H., Heeb, S., Cámara, M., Williams, P. & Dow, J. M. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 6712–6717.
- Simm, R., Morr, M., Kader, A., Nimtz, M. & Römling, U. (2004). *Mol. Microbiol.* **53**, 1123–1134.
- Studier, F. W., Rosenberg, A. H., Dunn, J. I. & Dubendorff, J. W. (1990). *Methods Enzymol.* **185**, 60–89.

Determinação dos sítios de Selênio no cristal de PilZ_{XAC1133}-SelMet: Os sítios de selênio foram determinados pelo programa Shelxd¹³⁸ utilizando o conjunto de dados processados do “peak” (**Figura 7**). Foram encontrados 8 sítios de selênio pelo programa Shelxd (solução com coeficiente de correlação de 49,04) e estes sítios foram utilizados como sítios iniciais no programa auto-SHARP, em que os três conjuntos de dados foram utilizados.

Estes sítios foram refinados e as fases iniciais do mapa de densidade eletrônica de PilZ_{XAC1133} foram determinadas pelo programa auto-SHARP. Os dados estatísticos do poder de faseamento do conjunto de dados do “peak”, assim como os valores estatísticos do primeiro mapa de densidade eletrônica de PilZ_{XAC1133} estão mostrados na **Tabela 11**. A partir deste mapa, o programa ARP/wARP⁷¹ construiu três fragmentos de PilZ_{XAC1133}, que correspondem aos aminoácidos de 10-58, 61-71 e 79-105. Neste momento, como descrito no **Anexo 1**, estávamos usando o grupo espacial P6₁22 que possui apenas uma molécula de PilZ_{XAC1133} na unidade assimétrica.

Figura 7: Determinação dos sítios de Selênio do cristal de PilZ_{XAC1133}.

Foram calculados 1000 soluções para a posição dos átomos de selênio no cristal de PilZ_{XAC1133} pelo programa Shelxd. No entanto, apenas as soluções de alto Coeficiente de Correlação (C.C) são corretas.

Tabela 11: Dados estatísticos do primeiro mapa de densidade eletrônica.

Dados estatísticos do poder de faseamento ("Phasing Power" – PP) tanto para reflexões centricas (PP_cen) quanto para reflexões acentricas (PP_acen) dos conjuntos de dados do peak.

Peak	
PP_acen	1.562
PP_cen	0
Mapa de Densidade Eletônica	
FOMacen	0.29693
FOMcent	0.16697
Solvent(%)	40.8
Rfactor(%)	21.6
Overall Correlation (%)	83.14

Refinamento da Estrutura de PilZ: O modelo foi refinado pelo programa Refmac5 e manualmente através da interface gráfica do COOT⁷⁴. Inicialmente determinamos a estrutura de PilZ_{XAC1133} no grupo espacial P6₁22, onde há uma molécula de PilZ_{XAC1133} na unidade assimétrica. No entanto, os valores de R_{free} e R_{factor}, valores estatísticos do erro do modelo, não estavam abaixando. Para tentar resolver este problema, mudamos o grupo espacial de PilZ_{XAC1133} de P6₁22 para P6₁, grupo espacial de menor simetria, onde há duas moléculas de PilZ_{XAC1133} na unidade assimétrica.

Para encontrar a posição da segunda molécula de PilZ_{XAC1133} na unidade assimétrica foi utilizado o programa PHASER utilizado como modelo a própria estrutura de PilZ_{XAC1133}. A estrutura foi refinada utilizando TLS (Translation, Libration e Screw)⁷⁵ e os valores de R_{free} e R_{factor} abaixaram para 0.190 e 0.243, respectivamente. A estrutura de PilZ_{XAC1133} foi depositada no PDB (*Protein Data Bank*) com PDBID: 3CNR. O Ramachadran Plot da estrutura de PilZ_{XAC1133} depositada está mostrado na **Figura 8**. Os dados estatísticos do refinamento da estrutura de PilZ_{XAC1133} estão descritos no manuscrito em **Anexo 2**, a seguir.

Análise da superfície enterrada entre as moléculas de PilZ localizadas na unidade assimétrica: A maior superfície enterrada entre as duas moléculas de PilZ_{XAC1133} no cristal é de 395.3 Å². Esta pequena área de superfície enterrada é devida, provavelmente, apenas ao empacotamento cristalino e não a uma interação biologicamente relevante¹³⁹. Os resultados da eluição de PilZ_{XAC1133} na coluna de Gel

filtração Superdex 75, (**Figura 8**), foi condizente com está observação. O volume de eluição de PilZ_{XAC1133} foi de 73 mL. A fórmula que correlaciona a massa da proteína com o volume de eluição desta coluna é dada pela equação: $\log MM = -1.3014.V/45,7 + 3,2624$. Desta forma, a massa molecular (MM) de PilZ_{XAC1133} tanto na concentração de 7.2 g/L quanto na concentração de 3.6g/L é de 15.3 KDa. A massa molecular de PilZ_{XAC1133} teórica de 12.43 KDa, mostrando que PilZ_{XAC1133} eluiu na forma monomérica.

Figura 8: Ramachandran plot da estrutura de PiIZ_{XAC1133}.

A maioria dos resíduos da estrutura de PiIZ_{XAC1133}, 95,6%, estão localizados em regiões favoráveis e apenas 4,4% estão localizados em regiões adicionalmente favoráveis.

Figura 9: Eluição de PilZ_{XAC1133} na coluna de Gel Filtração Superdex 75.

Cromatogramas da eluição de PilZ_{XAC1133} após aplicação da amostra na coluna de Gel Filtração 75 utilizando duas concentrações de proteínas: 7.1g/L (**a**) e 3.5 g/L (**b**).

3.3.2 Análise estrutural e funcional de PilZ e seu papel no funcionamento do T4P.

Estrutura de PilZ: A estrutura de PilZ_{XAC1133} é composta por uma folha beta composta por 5 fitas ($\beta 1$, $\beta 3$ - $\beta 6$) que se curvam formando um barril quase fechado com uma fenda que ocorre entre as fitas $\beta 1$ e $\beta 3$. Associado a esta fenda está localizada uma mini-fita beta, $\beta 2$, que contém apenas dois resíduos e que está posicionada de forma antiparalela à fita $\beta 3$. Além do mais, a estrutura de PilZ_{XAC1133} contém duas hélices, $\alpha 1$, localizada entre as fitas $\beta 1$ e $\beta 2$, e a hélice $\alpha 2$ localizada no C-terminal (figura 1 do Anexo 2).

PilZ_{XAC1133} não contém o motivo estrutural importante para a interação ao c-diGMP. A comparação estrutural de PilZ_{XAC1133} com as estruturas dos dois outros domínios de PilZ depositadas no PDB (PilZ_{VCA0042} de *Vibrio cholerae*, apo (PDBID: 1YLN) e em complexo com c-diGMP (PDBID: 2RDE)¹¹⁶ e PilZ_{PA4608} de *P. aeruginosa* (PDBID: 1YWU)¹¹⁸), revelaram mudanças estruturais significativas no sítio de interação ao c-diGMP (figura 1 do Anexo 2). A consequência destas mudanças na estrutura secundária e terciária é que o sítio de interação ao c-diGMP não existe na estrutura de PilZ_{XAC1133}. Isto explica a observação da incapacidade de PilZ_{XAC1133} (figura 3 suplementar, do Anexo 2), assim como PA2960 de interagirem com c-diGMP^{6,119}. Durante o término deste trabalho, Li e colaboradores mostraram que a estrutura de PilZ_{XC1028} de *Xanthomonas campestris*, ortólogo ao PilZ_{XAC1133}, apresenta as mesmas mudanças estruturais significativas no sítio de interação ao c-diGMP¹⁴⁰.

Aminoácidos conservados no grupo ortólogo PilZ/PA2960: O grupo ortólogo PilZ/PA2960 possui cinco motivos de sequência conservados que designamos **MI-MIV** e que correspondem aos seguintes resíduos na proteína de PilZ_{XAC1133}: L₂₁Y₂₂xxY₂₅M₂₆P₂₇F₂₈xxxG₃₂G₃₃(I/L)₃₄F₃₅(V/I)₃₆XT₃₈ (**MI**), G₆₅xV₆₇xW₆₉xxP₇₂ (**MII**), G₈₂(I/V)₈₃G₈₄(I/V)₈₅ (**MIII**), (R/K)₉₇xxI₁₀₀E₁₀₁xxL₁₀₄ (**MIV**) e S₁₁₀xxxT₁₁₄XT₁₁₆ (**MV**) (Figura 3 do Anexo 2).

Estes motivos se localizam em uma única face da estrutura de PilZ_{XAC1133} (Figura 1D suplementar, do Anexo 2). Alguns destes resíduos, Y₂₂, Y₂₅, P₂₇, F₂₈, F₃₅, W₆₉, P₇₂ e L₁₀₄, são hidrofóbicos ou aromáticos e estão expostos na superfície da estrutura de PilZ_{XAC1133}. Esta observação sugere que estes resíduos possam pertencer à interface de interação de PilZ_{XAC1133} com outras proteínas.

Alguns mutantes de PilZ_{XAC1133} foram produzidos e purificados (**Figura 10 b**), com base na análise descrita acima, PilZ_{Y22A}, PilZ_{F25A}, PilZ_{F28A}, PilZ_{W69A}, PilZ_{R97A}, PilZ_{E101A}, PilZ_{L104A} e PilZ_{Δ107} (deleção dos dez últimos aminoácidos, de 107-117, que contem o motivo V).

PilZ interage com o domínio EAL de FimX e com PilB: Ensaios de 2-híbrido utilizando como presa PilZ_{XAC1133}, devidamente clonada no vetor pOBD, contra uma biblioteca genômica de *Xac*, devidamente clonada no vetor pOAD, identificou duas presas: o domínio EAL de FimX_{XAC2398} e PilB_{XAC3239}. Além do mais, quando utilizamos FimX_{GGDEF-EA(267-689)} como isca contra a biblioteca genômica de *Xac*, uma das presas foi PilZ_{XAC1133}.

Estes resultados foram confirmados *in vitro*, por ensaios de Far-Western e RMN (**Figuras 5 e 7 e a Figura suplementar 5 do Anexo 2**). A **Figura 10 a** mostra as proteínas puras utilizadas para estes ensaios *in vitro*: FimX_{XAC2398}, domínio EAL de FimX_{XAC2398} (FimX_{EAL(436-689)}), domínio GGDEF de FimX_{XAC2398} (FimX_{GGDEF(267-432)}), fragmento de FimX_{XAC2398} contendo os domínios GGDEF e EAL (FimX_{his-GGDEF-EAL(267-689)}) e PilB_{XAC3239}.

Estas interações inéditas de interação proteína-proteína de PilZ com dois componentes da maquinaria de biogênese de pilus tipo IV, FimX e PilB, é a primeira evidência física que conecta PilZ fisicamente a maquinaria de biogênese de pilus tipo IV.

FimX de *P. aeruginosa* regula a produção do T4P de forma dependente a fatores ambientais^{5,111}. O domínio RR de FimX, que não possui o aspartato conservado para ser fosforilado, está envolvido na localização polar de FimX, provavelmente pela interação com outra proteína, todavia desconhecida, localizada na membrana^{5,111}. Os demais domínios, PAS, GGDEF e EAL em que os dois últimos domínios são degenerados, mas o domínio EAL ainda possui a capacidade de ligar ao segundo mensageiro bacteriano, c-diGMP. Estas observações sugerem que estes domínios possam ser responsáveis pela interação com moléculas citossólicas produzidas por estímulos ambientais, que irão modular a função de FimX na maquinaria de biogênese do T4P^{5,111}.

Não observamos nenhuma atividade diguanilato ciclase (DGC-A) nem fosfodiesterase (PDE-A) para os domínios EAL, GGDEF e EAL-GGDEF de FimX_{XAC2398}, mesmo na presença de PilZ_{XAC1133} e/ou GTP e/ou c-diGMP (**Figura 11**).

Nestes ensaios, os reagentes e os produtos foram analisados utilizando HPLC, vide Materiais e Métodos. Para determinar se FimX é capaz de interagir ao c-diGMP, medimos a estabilidade térmica de seus fragmentos na presença e na ausência do dinucleotídeo e observamos um aumento de estabilidade somente para o fragmento de FimX_{GGDEF-EAL} contendo os domínios GGDEF e EAL, mas não para os domínios isolados (**Figura 6 do Anexo 2**). A interação também foi detectada por RMN através da observação da atenuação da intensidade das ressonâncias específicas do c-diGMP na presença de FimX_{XAC2398} (**Figura suplementar 8 do Anexo 2**). Logo, concluímos que apesar de PilZ_{XAC1133} não interagir com c-diGMP, o fragmento de FimX_{XAC2398}, contendo os domínios GGDEF e EAL, interage.

Figura 10: Géis de eletroforese de SDS-PAGE das proteínas purificadas.

a) Amostras purificadas de (1) FimX_{XAC2398} inteira (peso molecular 75,8 KDa), (2) FimX_{his-GGDEF-EAL(267-689)} (peso molecular de 46,2 KDa), (3) FimX_{EAL(436-689)} (peso molecular 27,6 KDa), (4) FimX_{GGDEF(267-432)} (peso molecular 17,6 KDa), (5) marcador de peso molecular e (6-7) PilB_{XAC3239} (peso molecular 62,5 KDa). **b)** Amostras purificadas de (1) PilZ (peso molecular 12,4 KDa) e de seus mutantes: (2) PilZ_{Y22A} (peso molecular 12,3 KDa), (3) PilZ_{F25A} (peso molecular 12,3 KDa), (4) PilZ_{F28A} (peso molecular 12,3 KDa), (5) PilZ_{W69A} (peso molecular 12,3 KDa), (6) PilZ_{R97A} (peso molecular 12,4 KDa), (7) PilZ_{E101A} (peso molecular 12,4 KDa), (9) PilZ_{L104A} (peso molecular 12,4 KDa) e (10) PilZ_{Δ107} (peso molecular 11,2 KDa), (8) marcador de peso molecular.

EAL com cdiGMP

EAL com cdiGMP e PilZ

c

GGDEF com GTP

GGDEF com cdiGMP

minutos

GGDEF-EAL com GTP

GGDEF-EAL com cdiGMP

minutos

GGDEF-EAL com cdiGMP e PilZ

d

A.U.

Figura 11: Ensaio de atividade fosfodiesterase e diguanilato ciclase.

Ensaio de atividade fosfodiesterase e diguanilato ciclase de (a) FimX_{XAC2398} e seus domínios (b) EAL, (c) GGDEF e (d) GGDEF-EAL, na ausência e na presença de PilZ_{XAC1133}. Os cromatogramas mostram o tempo de retenção dos nucleotídeos presentes após a reação na coluna de C-18 acoplada a uma pré-coluna C-18 via HPLC. Amostras padrões mostraram que GTP elui com aproximadamente 6-7 minutos e o c-diGMP em 19-20 minutos. Detalhes das condições de reação e do método utilizado para separar os nucleotídeos na coluna C-18 estão descritos nos Matérias e Métodos item 2.12.7.

O C-terminal de PilZ interage com PilB mas não com FimX e a Y22 de PilZ é importante para a interação com ambas as proteínas. Ensaio de 2-híbridos utilizando PilZ_{XAC1133}, PilZ_{Y22A} e PilZ_{Δ107} como presas e PilB_{XAC3239} e FimX_{XAC2398} como iscas mostraram que o C-terminal de PilZ_{XAC1133}, que contém o motivo MV, está envolvido na interação com PilB_{XAC3239}, mas não com FimX_{XAC2398} (**Figura 4, 7 a e Figura 5 suplementar do Anexo 2**). Estes resultados sugerem que o C-terminal de PilZ_{XAC1133} que contém 10 resíduos desordenados, são importantes para a interação com PilB_{XAC3239}, mas não fazem parte da interface de interação com FimX_{XAC2398}. O mutante PilZ_{Y22A} apresentou uma diminuição na sua interação com PilB_{XAC3239} e também com FimX_{XAC2398} (**Figura 4 e 7 e Figura 5 suplementar do Anexo 2**).

PilZ pode interagir com FimX e com PilB simultaneamente. Ensaio de Far-Western mostram que PilZ_{XAC1133} interage com FimX_{XAC2398} e com PilB_{XAC3239} ao mesmo tempo. Não detectamos interações diretas entre FimX_{XAC2398} e PilB_{XAC3239} (**Figura 7 c do Anexo 2**).

Anexo 2 – Estrutura de PilZ e sua Interação com PilB e com o Domínio EAL de FimX: Implicações no Controle da Biogênese do Pilus Tipo IV.

**- PilZ protein Structure and interactions with PilB and FimX EAL Domain:
Implications for control of type IV pilus biogenesis.**

Fig. 1. PilZ_{XAC1133} structure and comparison with the PilZ domain from VCA0042. (a) Ribbon representation of PilZ_{XAC1133}. β1, α1 and β2 are colored green while β3–β6 are colored cyan and α2 is colored gray. (b) Ribbon representation of the PilZ domain from VCA0042 (PDB 2RDE, residues 136–247)²² in which β1 and β2 are colored green, while β3–β6 are colored cyan and α1 is colored gray. A stick model of c-diGMP bound to VCA0042 is shown in orange. (c and d), Topology diagrams of PilZ_{XAC1133} (c) and the VCA0042 PilZ domain (d). The β-strands are shown as triangles and the helices are shown as rectangles. The coloring scheme is the same as that used in a and b. (e) The alignment of the PilZ_{XAC1133} sequence with that of the PilZ domains from PA4608, VCA0042 and PP4397. Secondary structure elements are labeled below the alignment for PilZ_{XAC1133} and above the alignment for PA4608, VCA0042 and PP4397, and are highlighted in green for α1 and β1 in PilZ_{XAC1133}, cyan for the remaining β-strands, and gray for the C-terminal helix. The PilZ consensus sequence from Pfam (PF07238) is shown in which the 10 well conserved residues in the PilZ superfamily (see Supplementary Data Fig. 1A) are highlighted in yellow. The corresponding residues in PA4609, VCA0042, PP4397 and PilZ_{XAC1133} are shown in red. Panels a and b were generated using PyMOL [http://www.pymol.org].

lism regulators,¹⁰ chemosensory systems,¹ and proteins involved in c-diGMP signaling.^{7,11}

The filament of the TP4 pilus is composed of a helical polymer of mostly PilA pilin subunits.¹ Cycles of polymerization, attachment and depolymerization mediate several pilus-dependent phenomena, including twitching motility and natural transformation.^{12,13} Prepilin monomers are processed in the inner mem-

brane by the prepilin peptidase PilD¹⁴ before being incorporated into the base of the growing polymer in a process that requires the inner membrane integral protein PilC and the ATPase activity of PilB.¹⁵ The growing pilus passes through the outer membrane through pores formed by PilQ.¹⁶ Pilus retraction involves depolymerization at the pilus base and requires the ATPase activity of PilT.¹⁷

PilZ from *P. aeruginosa* (PA2960) was identified in 1996 as necessary for T4P biogenesis and twitching motility.¹⁸ The *pilZ* knockout is defective in the secretion of PilA polymers to the bacterial surface, though intracellular PilA pools are not affected.¹⁸ On the other hand, a knockout of the PilZ ortholog in *Neisseria meningitidis* is pilated but is not able to produce bacterial aggregates,¹⁹ and a PilZ knockout in *Xanthomonas campestris* pv *campestris* has slightly reduced T4P-dependent motility on semi-solid Eiken agar.²⁰ These observations suggest that the role of PilZ in T4P biogenesis and function could be modulated by unidentified environmental factors in different ways in different species.

To date, no PilZ-interacting protein has been identified and so the precise mechanism by which it controls T4P function is not known. However, over the past decade bacterial genome sequencing projects have identified a highly diverse superfamily of over 1500 PilZ homology domains, most of which have very little sequence similarity with the original PilZ protein. The low level of sequence identity among PilZ homology domains led to their division into several PilZ orthologous groups.^{21,22} Strong interest in these PilZ homology domains has arisen due to the observation that some can bind c-diGMP.^{11,22,23,24,25} Interestingly, a recent analysis of the eight PilZ paralogs in *P. aeruginosa* showed that the original PilZ protein (PA2960) was conspicuous in its inability to bind c-diGMP *in vitro*.²⁴ Also, only two of five *Vibrio cholerae* proteins containing PilZ homology domains, PlzB and PlzD, bind c-diGMP *in vitro*.¹¹ Knockouts in *plzB*, *plzC* and *plzD* resulted in variable changes in motility, biofilm formation and intestinal colonization.¹¹ PilZ domains can be found alone, in tandem or with other domains most often associated with signal transduction such as REC, PAS, EAL, GGDEF, and HD-GYP domains,²¹ the last three of which are also involved in c-diGMP signaling.

Sequence conservation in the superfamily of PilZ homology domains (~120 amino acids) is generally restricted to 10 residues (Fig. 1e; Supplementary Data Fig. 1A),²¹ the first five of which are found in two N-terminal motifs RxxxR and D/NxSxxG that are involved directly in interactions with c-diGMP.²² Here, we have determined the crystal structure of XAC1133 from the phytopathogen *Xanthomonas axonopodis* pv *citri* (PilZ_{XAC1133}), the cause of citrus canker disease. PilZ_{XAC1133}, a single-domain protein, belongs to the orthologous group defined by the original PilZ protein PA2960 that lacks the above signature motifs necessary for c-diGMP binding. Its structure, and that of the ortholog XC1028 from *X. campestris* pv *campestris*²⁶ reported during the preparation of this manuscript (PDB ID 3DSG), presents significant differences from those of three other distantly related PilZ homologs determined to date from *P. aeruginosa*,²⁷ *Pseudomonas putida* (unpublished results) and from *V. cholerae*²² that explain why PilZ homologs most closely related to PA2960 do not interact with c-diGMP.²⁴ We show also that PilZ_{XAC1133} interacts

with the PilB ATPase and the EAL domain of XAC2398, a homolog of the T4P regulatory protein FimX from *P. aeruginosa*.^{7,28} Both of these interactions are mediated *via* a conserved patch of hydrophobic and aromatic residues on the PilZ surface, while the interaction with PilB also involves conserved residues within a disordered C-terminal tail. These are the first protein-protein interactions demonstrated for any PilZ homology domain and are the first evidence for a direct physical connection between PilZ, FimX and the TP4 machinery.

Results

PilZ_{XAC1133} structure

The crystal structure of selenomethionine-labeled PilZ_{XAC1133} was determined to 1.9 Å resolution. Initial phases were estimated using a multiple wavelength anomalous dispersion (MAD) dataset, whose crystal parameters and data-reduction statistics have been reported.²⁹ MAD phasing was carried out as described in Materials and Methods and the refinement statistics are presented in Table 1. The PilZ_{XAC1133} structure (Fig. 1a) is composed of a central five-stranded β-sheet (β1, β3–β6) that curves upon itself but does not form a closed barrel, leaving a gap between β1 and β3. Associated with

Table 1. PilZ_{XAC1133} refinement statistics

Space group	P6 ₁
Unit cell dimensions	
<i>a</i> (Å)	62.125
<i>b</i> (Å)	62.125
<i>c</i> (Å)	83.543
Resolution (Å)	35.0 – 1.90 (1.95 – 1.90)
Molecules/A.U.	2
Residues/chain	117
Water molecules	119
Reflections used for refinement	13687
Completeness (%)	99.77 (98.4)
<i>R</i> -factor ^a	0.190 (0.228)
<i>R</i> _{free} ^a	0.244 (0.295)
Figure of merit	0.82
Average <i>B</i> -factors ^b	
All protein atoms (Å ²)	41.2
Protein main chain atoms (Å ²)	38.5
Protein side chain atoms (Å ²)	44.1
Water molecules (Å ²)	47.5
RMSD, from ideal	
Bond lengths (Å)	0.016
Bond angles (°)	1.547
Ramachandran plot	
Most favored (%)	95.7
Additionally allowed (%)	4.3
Generously allowed (%)	0.0
Disallowed (%)	0.0

The data collection statistics are published elsewhere.²⁹

^a *R*-factor = $\sum |F_o - F_c| / \sum F_o$ for 5% of the data excluded from *R*_{free}.

^b Mean *B*-values are based on the full *B*-factors, which include the contribution from the TLS parameters as calculated with TLSANL from the CCP4 package.³⁰

this open barrel is a small two residue β -strand ($\beta 2$) that binds in an antiparallel fashion with the N-terminus of $\beta 3$. Furthermore, two α -helices are found at one end of the PilZ_{XAC1133} barrel structure: $\alpha 1$ is located between $\beta 1$ and $\beta 2$ and $\alpha 2$ at the C-terminus.

As discussed below, the PilZ_{XAC1133} structure presents some surprising differences when compared to the structures of other PilZ homology domains. These differences have important implications regarding PilZ function. Therefore, in order to obtain evidence that these differences are maintained in solution, we produced $^{15}\text{N}/^{13}\text{C}$ -labeled PilZ_{XAC1133} and assigned its backbone resonances using multidimensional NMR. Analysis of chemical shifts with TALOS³⁰ show that the secondary structure elements present in solution correspond very well to those observed in the crystal structure (Fig. 2a).

There are two polypeptide chains in the asymmetric unit. Their structures are very similar, with an RMSD of 0.82 Å for all atoms and 0.06 Å for main chain atoms.

The first eight and the last 10 amino acid residues in chain A, and the first seven and the last 11 in chain B are absent. Also lacking density in both monomers is the sequence GAQG_{NR} (residues 74–79), which corresponds to a loop between $\beta 4$ and $\beta 5$. In order to investigate the correlation between the absence of electron density and the degree of flexibility of the protein backbone in solution, we measured backbone ^{15}N NMR relaxation rates and heteronuclear nuclear Overhauser enhancement (NOE) of PilZ_{XAC1133}. Low or negative $\{^1\text{H}\}-^{15}\text{N}$ heteronuclear NOE values at 27 °C were observed for amide protons in regions for which no electron density was observed (residues 1–7, 74–78, and 107–117) as well as for G90, indicating motions of backbone HN bond vectors on fast timescales (Fig. 2b).

The largest interface between monomer units in the crystal (corresponding to the asymmetric unit) buries only 397 Å² of surface per monomer and so seems to be a result of crystal packing. Consistent with this observation, PilZ_{XAC1133} (13 kDa) elutes as a monomer in gel-filtration chromatography (data

Fig. 2. NMR analysis of PilZ_{XAC1133} secondary structure and dynamics. (a) Values of ϕ (black) and ψ (red) dihedral angles predicted by TALOS³⁰ and based on $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$, ^{13}CO and ^{15}N chemical shift data (open triangles), and observed in the X-ray structure (filled circles), along the PilZ_{XAC1133} amino acid sequence. The secondary structure elements observed on the PilZ_{XAC1133} crystal structure were analyzed using PROCHECK⁸³ and are shown at the bottom (X-ray ss). Yellow arrows represent β -strands, and red cylinders represent α -helices. (b) Top: Distribution of main chain B-factor values for the backbone atoms in the PilZ_{XAC1133} crystal structure along the amino acid sequence. Mean B-values are based on the full B-factors, which include the contribution from the TLS parameters as calculated with TLSANL from the CCP4 package. Bottom: Heteronuclear $\{^1\text{H}\}-^{15}\text{N}$ NOE values for PilZ_{XAC1133} backbone ^{15}N spins as a function of the amino acid sequence.

not shown). On the other hand, the correlation time estimated by the ratio of ^{15}N R_2/R_1 NMR relaxation rates of rigid HN bond vectors, is 9.3 ns (data not shown). This value is compatible with an 18 kDa protein, which suggests some aggregation in the NMR sample where the protein was present at 0.3 mM.

Structural comparison between PilZ_{XAC1133} with other PilZ domains

The structures of three PilZ domains are available in the Protein Data Bank (PDB): the NMR structure of *P. aeruginosa* protein PA4608 (PDBID 1YWU),²⁷ the crystal structure of *P. putida* PP4397 (2GJG; Joint Center for Structural Genomics, unpublished results) and the crystal structures of *V. cholerae* protein VCA0042 determined in the apo conformation (1YLN; M. Zhou *et al.* unpublished results) and in complex with c-diGMP (2RDE).²² Although PA4608, PP4397 and VCA0042 share <20% sequence identity (Fig. 1E), they have similar structural topologies consisting of a closed β -barrel composed of six antiparallel sheets followed by an α -helix (Supplementary Data Fig. 2). Comparison of the PilZ_{XAC1133} structure with that of PA4608, PP4397 and VCA0042 PilZ domains revealed significant structural differences in protein topology involving the first two secondary structure elements; namely, $\beta 1$ - $\alpha 1$ in PilZ_{XAC1133} and $\beta 1$ - $\beta 2$ in PA4608, PP4397 and VCA0042. Furthermore, $\beta 1$ in PilZ_{XAC1133} hydrogen bonds with $\beta 4$ in a parallel manner, while $\beta 1$ in PA4608, PP4397 and VCA0042 associates with both $\beta 4$ and $\beta 2$ in an anti-parallel manner (Fig. 1a-d; Supplementary Data Fig. 2).

PA4608 and VCA0042 bind c-diGMP.^{11,27,22} Ligand binding is mediated by the RxxxR motif, located in an unstructured region preceding $\beta 1$, and the D/NxSxxG motif, located in a loop between $\beta 2$ and $\beta 3$ (Fig. 1e; Supplementary Data Fig. 1A).^{21,22} Mutations in these residues have been shown to abolish binding to c-diGMP,^{11,23,24,25} and the region containing these motifs corresponds precisely to that in which the PilZ_{XAC1133} structure differs from VCA0042 most significantly (Fig. 1a-d). These differences show clearly that the c-diGMP binding site localized in the N-termini and in the loop between $\beta 2$ and $\beta 3$ from VCA0042 does not exist in the PilZ_{XAC1133} structure (Fig. 1; Supplementary Data Fig. 2). As a consequence, PilZ_{XAC1133} should not interact with c-diGMP. To test this hypothesis, the ^1H - ^{15}N heteronuclear single quantum coherence (HSQC) spectra of ^{15}N labeled PilZ_{XAC1133} were obtained in the absence and in the presence of c-diGMP at a 1:1 molar ratio and no chemical shift perturbation was observed (Supplementary Data Fig. 3A). Moreover, the addition of c-diGMP does not increase the thermal stability of PilZ_{XAC1133} as measured in circular dichroism experiments (Supplementary Data Fig. 3B).

PilZ_{XAC1133} is a member of a PilZ ortholog group with distinct sequence signatures that map to one face of the 3D structure

There are more than 1500 sequences for eubacterial PilZ homology domains annotated in the Pfam database.³¹ These domains are highly diverse, with only a handful of well-conserved residues as described above (Supplementary Data Fig. 1A). Benach *et al.* showed that this superfamily can be divided into 28 families that also cluster together on a basis that reflects their context within the architectures of multi-domain proteins.²² Figure 3a shows that the group of homologs most similar to PA2690 and PilZ_{XAC1133} (which we refer to as the PilZ PA2960/XAC1133 orthologous group) does not possess the RxxxR or D/NxSxxG motifs involved in c-diGMP binding though it does possess one conserved basic residue preceding the first beta β strand (not modeled in the PilZ_{XAC1133} structure due to lack of electron density) and G33 corresponds structurally to the last residue of the D/NxSxxG motif at the N-terminus of $\beta 3$ (Fig. 1e; Supplementary Data Fig. 1b).

The sequence conservation profile of the PilZ PA2960/XAC1133 orthologous group, as defined by the approximately 100 PA2960/XAC1133 orthologs in the public databases (Fig. 3a; Supplementary Data Fig. 1B), presents five highly conserved sequence motifs (MI – MV) that correspond to the following residues in PilZ_{XAC1133}: L₂₁Y₂₂xxY₂₅M₂₆-P₂₇F₂₈xxxG₃₂G₃₃(I/L)₃₄F₃₅(V/I)₃₆xT₃₈ (MI), G₆₅x-V₆₇xW₆₉xxP₇₂ (MII), G₈₂(I/V)₈₃G₈₄(I/V)₈₅ (MIII), (R/K)₉₇xxI₁₀₀E₁₀₁xxL₁₀₄ (MIV) and S₁₁₀xxxT₁₁₄-xT₁₁₆ (MV). The first four motifs all cluster to one face of the PilZ_{XAC1133} structure (Fig. 3b; Supplementary Data Fig. 1D) and display several highly conserved aromatic or hydrophobic residues exposed at the protein surface: Y₂₂, Y₂₅, P₂₇, F₂₈, F₃₅, W₆₉, P₇₂ and L₁₀₄ (Fig. 3c). The last motif (MV) is located in an unstructured C-terminal region following $\alpha 2$, consistent with the ^{15}N NMR heteronuclear NOE data presented above (Fig. 2b).

Perturbations in stability and structure caused by PilZ_{XAC1133} mutants

In order to probe the relationships between conserved residues and PilZ_{XAC1133} function, we produced a set of PilZ_{XAC1133} mutants in which some of the conserved aromatic or hydrophobic residues were mutated to alanine (PilZ_{Y22A}, PilZ_{Y25A}, PilZ_{F28A}, PilZ_{W69A}, PilZ_{L104A}). Two highly conserved charged residues in the C-terminal helix, $\alpha 2$, were also mutated (PilZ_{R97A}, PilZ_{E101A}). Finally, a deletion of the last 10 amino acids (PilZ _{Δ 107-117}) removed all of the C-terminal residues that were not modeled due to lack of electron density in the PilZ_{XAC1133} (including conserved motif MV, S₁₁₀xxxT₁₁₄xT₁₁₆). The conformational stability of these mutants was evaluated by following circular dichroism during thermal denaturation of the purified recombinant

Fig. 3. Sequence conservation profile within the PA2960/XAC1133 orthologous group and its mapping onto the PilZ_{XAC1133} structure. (a) Multiple sequence representation of the PilZ PA2960/XAC1133 orthologous group. Approximately 100 sequences most similar to PilZ_{XAC1133} (e-values $<10^{-5}$) were aligned and the height of the residue symbol indicates the degree of conservation. The PilZ_{XAC1133} sequence numbering is shown below the alignment. Secondary structure elements are indicated below using red ovals (α -helix) and yellow arrows (β -strands). Boxes indicate conserved motifs MI-MV. The figure was generated using the WebLogo server (<http://weblogo.berkeley.edu/>). (b) Cartoon and semi-transparent surface depiction of the PilZ_{XAC1133} structure in which conserved motifs MI-MIV are colored blue, yellow, pink and orange, respectively (MV is found within the unstructured C-terminus and was not modeled in the PilZ_{XAC1133} structure). (c) Conserved hydrophobic and aromatic side-chains on the surface of PilZ_{XAC1133} are shown in red. Panels b and c were generated using PyMOL [<http://www.pymol.org/>].

proteins. The Y22A, F28A, W69A, E101A, and Δ 107-117 mutations resulted in little or no reduction in thermal stability (Supplementary Data Fig. 4a), while the Y25A, R97A, and L104A mutations decreased the structural stability of PilZ_{XAC1133} significantly (Supplementary Data Fig. 4b).

We analyzed the ^{15}N - ^1H HSQC spectra of these mutants in order to assess how each mutation affected the PilZ_{XAC1133} structure (Supplementary Data Fig. 5). Extreme line-broadening was observed in the spectra of the PilZ_{Y25A} (data not shown) and PilZ_{L104A} (Supplementary Data Fig. 5). This is consistent with the reduced stability of mutants Y25A and L104A, and may be due to disruption of a network of hydrophobic interactions involving these two residues that could be important for stabilizing the structure (Supplementary Data Fig. 6A). The HSQC spectra of the remaining mutants showed sharp lines indicative

of well-folded proteins. In all cases, chemical shift perturbations could be observed for specific resonances. (The residues whose ^{15}N - ^1H resonances presented the most significant perturbations in PilZ_{Y22A}, PilZ_{F28A}, PilZ_{W69A}, PilZ_{R97A}, PilZ _{Δ 107-117} are mapped to the PilZ structure in Supplementary Data Figure 6B-F and discussed in the figure legend.)

PilZ_{XAC1133} protein-protein interactions with XAC2398, a FimX homolog, suggest a role in c-diGMP signaling

We used PilZ_{XAC1133} as a bait in two-hybrid assays against a *Xac* total genomic DNA prey library.³² One of the interactions detected was with the uncharacterized protein XAC2398. This hypothetical protein has 689 residues and is composed of four domains: a response regulator receiver domain

(REC, residues 11–120), a PAS domain (residues 144–254), a GGDEF domain (residues 268–425), and an EAL domain (residues 441–685). XAC2398 is a homolog of the *P. aeruginosa* FimX protein, which is involved in the control of T4P biogenesis and twitching motility (see below).⁷ We refer to XAC2398 as FimX_{XAC2398}. The smallest fragment of FimX_{XAC2398} that interacted with PilZ_{XAC1133} corresponds to residues 419–689, which begins at the extreme C-terminus of the GGDEF domain and includes the whole EAL domain. Furthermore, when we used an FimX_{XAC2398} fragment containing only the GGDEF and EAL domains (FimX_{GGDEF-EAL}, residues 270–689) as a bait in two-hybrid assays against the *Xac* prey library, one of the preys identified corresponded to almost full-length PilZ_{XAC1133} (beginning at residue 3). To confirm the interaction between PilZ_{XAC1133} and the EAL domain of FimX_{XAC2398}, specific two-hybrid assays were performed using the following FimX_{XAC2398} fragments as baits against a PilZ_{XAC1133} prey: FimX_{GGDEF} (residues 270–432), FimX_{EAL} (residues 436–689) and FimX_{GGDEF-EAL}. All transformed yeast colonies containing both PilZ_{XAC1133} preys and FimX_{GGDEF-EAL} or XAC2398_{EAL} baits grew in selective medium for the interaction (SC-WLHA), while none of the yeast colonies containing the PilZ_{XAC1133} prey and XAC2398_{GGDEF} bait grew in selective medium (Table 2; Supplementary Data Fig. 7A–C). This result was confirmed in two-hybrid assays using the β -galactosidase activity as a reporter for these interactions (Fig. 4).

We used NMR spectroscopy to analyze the PilZ_{XAC1133}-FimX_{XAC2398} interaction *in vitro*. The ¹H-¹⁵N HSQC spectrum of ¹⁵N-labeled PilZ_{XAC1133} (Fig. 5a) was compared to that obtained in the presence of full-length unlabeled FimX_{XAC2398}, FimX_{GGDEF-EAL}, FimX_{GGDEF}, and FimX_{EAL} (Fig. 5b–e). Significant perturbations in the ¹H-¹⁵N HSQC spectrum, including severe peak-broadening, were observed in all cases except for FimX_{GGDEF}. Only very small perturbations

were observed in the HSQC spectrum of PilZ_{XAC1133} upon addition of FimX_{GGDEF}, even at the elevated concentrations used in these experiments (0.225 mM). Therefore, both two-hybrid and NMR results suggest that PilZ_{XAC1133} binds to FimX_{XAC2398} principally *via* interactions with its EAL domain.

FimX_{GGDEF-EAL} and the FimX_{GGDEF-EAL}-PilZ complex bind c-diGMP

Some GGDEF and EAL domains can bind c-diGMP.^{33,34,35,36} However, several GGDEF and EAL domains appear to be enzymatically inactive due to variations in otherwise highly conserved residues required for catalysis.^{28,35,36,37,38} We measured the thermal stability of FimX_{GGDEF-EAL} in the absence and in the presence of the dinucleotide in order to determine whether FimX_{XAC2398} could bind c-diGMP. We found that c-diGMP increases the stability of FimX_{GGDEF-EAL} by 4 °C (Fig. 6a), while GTP and ATP did not increase the stability of FimX_{GGDEF-EAL} (Fig. 6c). Furthermore, no c-diGMP-induced increases in stability were observed for isolated EAL or GGDEF domains, nor for a 1:1 mixture of both FimX_{GGDEF} and FimX_{EAL} (data not shown). Therefore, it is not clear from our data to which domain c-diGMP binds, and the possibility exists that it may require both domains for binding. Ligand binding was observed in NMR experiments in which specifically the H8 and H1' c-diGMP resonances were attenuated significantly upon the addition of full-length FimX_{XAC2398} (Supplementary Data Fig. 8). Finally, we observed that c-diGMP induces a 3°C increase in stability of the FimX_{GGDEF-EAL}-PilZ_{XAC1133} complex, indicating that PilZ_{XAC1133} does not inhibit c-diGMP binding to XAC2398_{GGDEF-EAL} under the assay conditions used (Fig. 6b). The above results show that while PilZ does not bind c-diGMP on its own, it does interact with FimX_{XAC2398} protein domains that bind the cyclic dinucleotide.

Table 2. Directed two-hybrid assays to evaluate interactions between FimX fragments and PilZ mutants and between PilB and PilZ mutants

Two hybrid bait-prey pair	Independent clones tested (SC-WL selection)	Positive interactions (SC-WHLA selection)	SC-WHLA <i>versus</i> SC-WL (%)
FimX _{GGDEF-EAL} + PilZ _{XAC1133}	40	40	100
FimX _{GGDEF-EAL} + PilZ _{Y22A}	23	2	8.7
FimX _{GGDEF-EAL} + PilZ _{Δ107-117}	48	48	100
FimX _{GGDEF} + PilZ _{XAC1133}	36	0	0
FimX _{GGDEF} + PilZ _{Y22A}	35	0	0
FimX _{GGDEF} + PilZ _{Δ107-117}	36	0	0
FimX _{EAL} + PilZ _{XAC1133}	40	40	100
FimX _{EAL} + PilZ _{Y22A}	44	7	15.9
FimX _{EAL} + PilZ _{Δ107-117}	48	48	100
PilB _{XAC3239} + PilZ _{XAC1133}	33	33	100
PilB _{XAC3239} + PilZ _{Y22A}	37	3	8.1
PilB _{XAC3239} + PilZ _{Δ107-117}	44	0	0

PJ694-a yeast cells (*MATa trp1-901 leu2-3, 112 ura3-52 his3-200 gal4Δgal80 ΔLYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ*) were co-transformed with bait-prey pairs cloned in the pOBD and pOAD two-hybrid vectors, respectively (indicated in the first column). Co-transformants were selected by growth in SC medium lacking tryptophan and leucine (SC-WL). These clones were tested individually for growth in SC plates lacking tryptophan, leucine, histidine and adenine (SC-WHLA). PilZ_{XAC1133} interacts with FimX_{GGDEF-EAL}, FimX_{EAL} and PilB_{XAC3239} but not with FimX_{GGDEF}. The mutant PilZ_{Δ107-117} maintains interactions with FimX_{GGDEF-EAL} and FimX_{EAL} but does not interact with PilB_{XAC3239}. The mutant PilZ_{Y22A} has reduced capacity to interact with FimX_{XAC2398} and PilB_{XAC3239}.

Fig. 4. β -Galactosidase activity measured in two-hybrid assays to test interactions of PilZ_{XAC1133} with FimX_{XAC2398} and PilB_{XAC3239}. Colonies of Pj694-a yeast cells transformed with PilZ_{XAC1133}, PilZ_{Y22A}, or PilZ_{Δ107-117} preys and/or the FimX_{GGDEF-EAL}, FimX_{GGDEF}, FimX_{EAL}, or PilB_{XAC3239} baits were selected on SC-L, SC-WL or SC-WLHA plates. Individual colonies were used to inoculate liquid cultures and β -galactosidase activity was measured as described in Materials and Methods. Colonies containing the PilZ_{XAC1133} or PilZ_{Δ107-117} preys and FimX_{GGDEF-EAL} or FimX_{EAL} baits present significantly greater activity than the no bait control (PilZ alone) or cells containing the PilZ_{XAC1133} prey and FimX_{GGDEF} bait grown in SC-WL media. Cells containing the PilZ_{XAC1133} prey and FimX_{GGDEF} bait did not grow in SC-WLHA medium. These results are consistent with the conclusion that PilZ_{XAC1133} interacts with the FimX_{XAC2398} EAL domain and that PilZ_{XAC1133} residues 107–117 are not required for interaction. Similarly, PilZ_{XAC1133}, but not PilZ_{Δ107-117}, interacts with PilB_{XAC3239}. All cells carrying the PilZ_{Y22A} prey did not grow in SC-WLHA and showed only background levels of β -galactosidase when grown in SC-WL medium, consistent with our observations that this mutant does not interact with FimX_{XAC2398} or PilB_{XAC3239}. The β -galactosidase activity is within the range normally observed using Pj694-a yeast cells.⁶⁷

PilZ_{XAC1133} protein–protein interactions with PilB provide the first direct physical link to T4P biogenesis

Using PilZ_{XAC1133} as a bait in two-hybrid assays against the *Xac* prey library, we detected two preys derived from XAC3239, a PilB ortholog (PilB_{XAC3239}). These preys correspond to almost full-length PilB_{XAC3239} fragments, the smallest of which begins at residue 59 of this 578 residue protein. PilB proteins are hexameric ATPases associated with the base of T4P at the cytosolic face of the plasma membrane where they are involved in the control of T4P polymerization/extension.^{39,40} To confirm the interaction between PilZ_{XAC1133} and PilB_{XAC3239}, two-hybrid assays were performed using PilB_{XAC3239} as bait against the PilZ_{XAC1133}

prey. All yeast cells transformed with both bait and prey grew in selective medium for the interaction (Table 2; Supplementary Data Fig. 7D) and displayed significant β -galactosidase activity (Fig. 4). The PilB_{XAC3239} interaction with PilZ_{XAC1133} was observed in far-Western assays using immobilized PilB_{XAC3239} and anti-PilZ_{XAC1133} antibodies (Fig. 7A; Supplementary Data Fig. 9).

A ternary PilB_{XAC3239}–PilZ_{XAC1133}–FimX_{XAC2398} complex

The PilB_{XAC3239}–PilZ_{XAC1133} interaction detected in the far-Western assay was stable in the presence of FimX_{GGDEF-EAL} in the presence and in the absence of c-diGMP (Fig. 7b). In order to determine whether PilZ_{XAC1133}, PilB_{XAC3239} and FimX_{GGDEF-EAL} could form a stable ternary complex, we performed far-Western assays in which PilB_{XAC3239} immobilized on the membrane was incubated with PilZ_{XAC1133} and FimX_{GGDEF-EAL}. Using an antibody against the polyhistidine tag in recombinant FimX_{GGDEF-EAL}, we observed that FimX_{GGDEF-EAL} colocalizes with PilB_{XAC3239} only in the presence of PilZ_{XAC1133} (Fig. 7c). This observation implies that PilZ_{XAC1133} can bind to PilB_{XAC3239} and FimX_{XAC2398} simultaneously. This ternary complex is observed in the presence and in the absence of up to 2 μ M c-diGMP (Fig. 7c).

PilZ residue Y22, but not residues 107–117, is involved in interactions with FimX_{XAC2398}

Since the most highly conserved residues in the PilZ PA2690/XAC1133 orthologous group map to either one face of the PilZ structure (motifs MI-MIV) or to an unstructured C-terminus that includes motif MV, we used the PilZ_{Y22A} and PilZ_{Δ107-117} mutants as prey in two-hybrid assays against the FimX_{GGDEF-EAL}, FimX_{GGDEF} and FimX_{EAL} baits. The results presented in Table 2 and Fig. 4 (and Supplementary Data Fig. 7A–C) show that the Y22A mutation abolishes interactions with FimX_{GGDEF-EAL} and FimX_{EAL}, while the Δ 107–117 deletion did not effect this interaction. FimX_{GGDEF} did not interact with either of these mutants in these assays. We used the purified ¹⁵N labeled PilZ_{XAC1133} mutants and FimX_{XAC2398} to assay for interactions under NMR conditions. (Supplementary Data Fig. 5 compares the ¹H-¹⁵N HSQC spectra for the ¹⁵N-labeled PilZ_{Y22A}, PilZ_{F28A}, PilZ_{W69A}, PilZ_{R97A}, PilZ_{E101A}, PilZ_{L104A} and PilZ_{Δ107-117} in the absence and in the presence of unlabelled full-length FimX_{XAC2398}). In all cases, including PilZ_{Y22A}, extensive peak broadening was observed that resulted in the disappearance of the majority of the PilZ_{XAC1133} peaks. Therefore, the Y22A mutation probably reduces the affinity for FimX_{XAC2398} to levels that cannot be detected in the two-hybrid assays but are still observed at the high concentrations of protein in the NMR sample (0.1 – 0.2 mM). Furthermore, in all of these experiments, almost all the main-chain ¹H-¹⁵N peaks observable in the ¹H-¹⁵N-HSQC spectra of

Fig. 5. PilZ_{XAC1133} interacts with the EAL domain of FimX_{XAC2398}. (a) ^1H - ^{15}N HSQC spectrum of PilZ_{XAC1133} acquired at 27 °C. The box at the bottom right shows an amplification of the region displaying the greatest spectral overlap. The assigned resonances are shown. (b-e) ^1H - ^{15}N HSQC spectra of 1:1 mixtures of ^{15}N -labeled PilZ_{XAC1133} and unlabeled full-length FimX_{XAC2398} (b) FimX_{EAL-GGDEF} (c) FimX_{GGDEF} (d) and FimX_{EAL} (e). The resonances of the disordered C-terminal tail of PilZ_{XAC1133} are still observable in the ^{15}N -HSQC spectrum of the complex with full-length FimX_{XAC2398} (b), indicating that the C-terminal backbone experiences a different motional regime. The spectra were acquired at 500 MHz ^1H frequency, except for the spectrum in a, which was acquired at 600 MHz. The sample conditions were as described in Materials and Methods.

the ^{15}N -PilZ_{XAC1133}-FimX_{XAC2398} complexes are derived from the disordered C-terminus (residues 110, 111, 112, 114 and 117 labeled in Fig. 5b; Supplementary Data Fig. 5). This is consistent with the conclusion that the last 10 PilZ_{XAC1133} residues are not involved in interaction with FimX_{XAC2398}.

PilZ residues 107–117 are involved in interactions with PilB_{XAC3239}

When PilZ_{Y22A} and PilZ_{Δ107-117} were used as prey in two-hybrid assays, the no interaction with PilB_{XAC3239} was observed (Table 2, Fig. 4; Supplementary Data

Fig. 6. Conformational stability of PilZ_{XAC1133} and FimX_{XAC2398}. (a–c) Thermal denaturation of 3.4 μ M FimX_{GGDEF-EAL} in the absence (a and c) or in the presence (b) of 3.5 μ M PilZ_{XAC1133}. The conditions were: 10 mM Tris–HCl, pH 7.0, 10 mM NaCl, 1 mM MgCl₂. Where indicated, 5 μ M c-diGMP (a and b) or ATP or GTP (c) were added. In all experiments, the denaturation was accompanied by measuring $[\theta]_{220\text{nm}}$ and the signal was converted to fraction of unfolded protein (Yu) as described in Materials and Methods.

Fig. 7D), suggesting that both the conserved hydrophobic surface and disordered C-terminal residues might have a role in interactions with PilB. The PilB_{XAC3239} interactions with PilZ_{XAC1133} and its mutants were analyzed by far-Western assays using immobilized PilB_{XAC3239} and anti-PilZ_{XAC1133} antibodies (Fig. 7a). No interaction was observed between PilB_{XAC3239} and PilZ_{Y22A}, or PilZ_{W69A} and PilZ_{R87A}, while very weak interactions were observed for PilZ_{Y28A}, PilZ_{L104A} and PilZ_{Δ107–117} (Fig. 7a). The PilB_{XAC3239} interactions observed for PilZ_{Y25A} and PilZ_{E101A} were significant, but less than that observed for wild type PilZ_{XAC1133} (Fig. 7a).

***pilZ*_{XAC1133}, *fimX*_{XAC2398} and *pilB*_{XAC3239} mutants present increased motility on semisolid surfaces**

*Xac pilZ*_{XAC1133} and *fimX*_{XAC2398} null mutants were able to induce citrus canker symptoms in a manner similar to that observed for wild type *Xac* (Supplementary Data Fig. 10). The *pilZ* gene was originally identified in *P. aeruginosa* as necessary for twitching motility.¹⁸ We have not been able to establish reproducible subsurface stab twitching assays for *Xac*. However, we were able to do motility assays on semi-solid agar surfaces. We found that the *pilZ*_{XAC1133} knockout has increased motility on semi-solid agar surfaces in several nutrient media tested, including SB, XVM2 and STON (Fig. 8; Supplementary Data Fig. 11). Complementation with a plasmid carrying the wild type *pilZ*_{XAC1133} gene under control of its own promoter or under the control of a constitutive *lac* promoter led to full or partial recovery of the wild type phenotype (Fig. 8; Supplementary Data Fig. 11). A similar increase in motility was observed for the *fimX*_{XAC2398} knockout strain in all three media, and complementation by the introduction of the *fimX*_{XAC2398} gene *in trans* resulted in recovery of the wild type phenotype (Supplementary Data Fig. 11). Furthermore, the *pilB*_{XAC3239} knockout strain had increased motility in SB medium (Fig. 8c). The addition of NaCl (300 mM final concentration) or using base medium containing 100 mM potassium phosphate⁷ severely decreased motility of the wild type and *pilZ*_{XAC1133} knockout strains (data not shown). In Eiken agar, the *pilZ*_{XAC1133} knockout had motility very similar to that observed for the wild type *Xac* strain, while the *fimX*_{XAC2398} knockout had increased motility (data not shown). Complementation of the *pilZ*_{XAC1133} null strain with a plasmid carrying the *pilZ*_{Δ107–117} gene did not recover the wild type phenotype in SB medium (Fig. 8a). As the deletion of residues 107–117 weakens the PilZ–PilB interaction (above), this result is consistent with a role for PilZ in the control of PilB function. Finally, in these assays, a knockout in the *fliC* gene that codes for the principal flagellar subunit did not show any motility change in SB and STON media, and had increased motility in Eiken agar (data not shown), indicating that the observed motility is not dependent on flagella. Bacterial spreading on semisolid surfaces in the absence of T4P and flagella, called sliding motility, has been observed before in *P. aeruginosa* (see Discussion).⁴¹

Discussion

While the central role of c-diGMP signaling in bacteria is now well recognized, few specific c-diGMP targets are known. Other than the active sites of the GGDEF, EAL and HDGYP domains, the only other known binding sites for c-diGMP are c-diGMP riboswitches,⁴⁰ the allosteric site (I site) of the DGC PleD from *C. crescentus*,³⁴ PleD,⁴² and

Fig. 7. PilZ_{XAC1133} and PilZ_{XAC1133}-FimX_{XAC2398} interacts with PilB_{XAC3239}. (a) Far-Western assays used to detect interactions between PilB_{XAC3239} and wild-type (WT) or mutant PilZ_{XAC1133}. Total cell lysates of *E. coli* cells expressing PilB_{XAC3239} (+) or an unrelated protein YacQ (-) were separated by SDS-PAGE, transferred to nitrocellulose membranes and incubated with PilZ_{XAC1133} or its mutants. Bound PilZ was detected with PilZ_{XAC1133} antiserum and protein A peroxidase. Y22A, F28A, W69A, R97A, L104A and $\Delta_{107-117}$ mutations resulted in severe reductions in binding. (b) PilZ_{XAC1133} can bind to PilB_{XAC3239} in the presence of FimX_{XAC2398}. Far-Western assays were performed as described for a, except that FimX_{GGDEF-EAL} and c-diGMP were added to the incubations where indicated. The amount of protein or c-diGMP present in the incubations was 1 μ M (+), 2 μ M (++) or 4 μ M (++++). (c) Far-Western assays were used as described for b, except that bound FimX_{XAC2398} was detected using anti-serum against its polyhistidine tag. Protein and c-diGMP concentrations are represented as in b. Arrows show detection of a signal for PilZ in a and b, and for FimX in c.

FleQ⁴³ from *P. aeruginosa* and PilZ-homology domains from a number of bacteria.^{11,22,23,24,25} However, the only PilZ domain to which c-diGMP binding has been shown to regulate specific protein function is that found in subunit A of cellulose synthase.⁴⁴ All other pilZ phenotypes have involved stand-alone PilZ domains or PilZ domains contained within proteins of unknown molecular function.²¹ Some PilZ homology domains have been shown to bind c-diGMP, while others have failed to do so in *in vitro* binding assays.^{11,24,26}

The structure of PilZ_{XAC1133} reveals an unexpected structural diversity in the topology of the PilZ fold,

which abolishes the c-diGMP binding site in the PilZ PA2960/XAC1133 orthologous group. Highly conserved residues within this group map to one surface of the domain structure and include a significant number of exposed hydrophobic side chains that potentially mediate interactions with other proteins, including FimX_{XAC2398} and PilB_{XAC3239} identified in this work. This hypothesis is supported by the reduced interactions observed using PilZ_{XAC1133} mutants (specifically PilZ_{Y22A} for FimX_{XAC2398} and PilZ_{Y22A}, PilZ_{Y28A}, PilZ_{W69A} and PilZ_{L104A} for PilB_{XAC3239}). Furthermore, two-hybrid and far-Western analysis indicate that the well-conserved

Fig. 8. Motility of *Xac* and its mutants on semisolid surfaces after two days of growth in SB medium containing 0.6% (w/v) agar. (a) wt, Wild type *Xac*; $\Delta fimX$, $\Delta fimX_{XAC2398}$ knockout; $\Delta pilZ$, $\Delta pilZ_{XAC1133}$ knockout; $\Delta pilZ + pilZ_{pMR20}$ and $\Delta pilZ + pilZ_{\Delta 107-117 pMR20}$, $\Delta pilZ_{XAC1133}$ knockout strain complemented with the $pMR20_{pilZ}$ and $pMR20_{pilZ_{\Delta 107-117}}$ constructs, respectively, that express $pilZ_{XAC1133}$ or $pilZ_{\Delta 107-117}$ under the control of a constitutive *lac* promoter. (b) wt, Wild type *Xac*; $\Delta pilZ$, $\Delta pilZ_{XAC1133}$ knockout; $\Delta pilZ + pilZ_{pUFR047}$, $\Delta pilZ_{XAC1133}$ knockout complemented with the $pUFR047_{pilZ}$ construct that expresses $pilZ_{XAC1133}$ under the control of its own promoter. (c) wt, Wild type *Xac*; $\Delta pilB$, $\Delta pilB_{XAC3239}$ knockout.

though conformationally flexible C-terminus of $PilZ_{XAC1133}$ contributes to its interactions with $PilB_{XAC3239}$. The observation that the Y22A mutation reduces $PilZ_{XAC1133}$ binding to both $FimX_{XAC2398}$ and $PilB_{XAC3239}$ suggests that this residue might have a regulatory role in these interactions: *via* a direct interaction with one of the two partner proteins, Y22 could trigger a conformational change in $PilZ$ that would then allow it to bind the other partner in the ternary complex. One observation consistent with this idea is that the Y22A mutation results in significant chemical perturbations in several residues, some of which are located in the unstructured C-terminal region (residues 107, 114, and 116) identified to be important for interactions with $PilB$. Furthermore, deletion of residues 107–117 perturbs the chemical shift of Y22 as well as other residues in the $\beta 1$ - $\alpha 1$ loop and within the N-terminal half of $\alpha 1$ (Supplementary Data Figs. 5, 6B and 6F).

The $PilZ_{XAC1133}$ – $FimX_{XAC2398}$ interaction is mediated by the $FimX_{XAC2398}$ EAL domain. $FimX_{XAC2398}$ is a homolog of the *P. aeruginosa* *FimX* protein.^{7,28} These two proteins share 30% sequence identity

over their entire length.⁷ There are at the moment 42 *FimX* homologs with the same domain topology in the public databases and all lack an aspartate critical for phosphorylation of their REC domains by cognate histidine kinases (data not shown).⁷ Interestingly, all these homologs are found in organisms whose genomes code for a member of the $PilZ$ PA2690/ $XAC1133$ orthologous group (data not shown).

The sequence requirement for substrate binding and catalysis by EAL and GGDEF domains has been investigated intensively over the past few years. Rao *et al.* have proposed a catalytic mechanism for EAL domains in which they identify seven residues essential for PDE activity: a general base catalyst, four residues that coordinate an essential magnesium ion and two residues that have a structural role.³⁸ While the *P. aeruginosa* *FimX* protein has both nonconservative substitutions in several of these residues, there are conflicting reports regarding its PDE activity.^{28,38} In $FimX_{XAC2398}$, the residue involved in general base catalysis is conserved (E653), but all four Mg^{2+} -coordinating residues are modified (Q477, R534, Q566, and E596, which are

Fig. 9. PilB-PilZ-FimX interactions in the context of T4P. The filament of TP4 is a polymer of pilin subunits (mostly PilA). After being processed by the prepilin peptidase, pilin subunits associated with the inner membrane are incorporated into the base of the pilus filament in a process that requires the inner membrane protein PilC and the ATPase activity of PilB, which forms a hexameric ring associated with the cytoplasmic face of the inner membrane.^{4,57} The growing pilin polymer passes through pores formed by PilQ in the outer membrane.¹⁶ Pili are depolymerized in a process that requires the hexameric PilT ATPase.¹⁷ PilZ interacts with PilB and with XAC2398, a FimX homolog (FimX_{XAC2398}). FimX itself may be tethered to the inner membrane via interactions between its REC domain with an unknown protein (?).^{7,28} These interactions point to a means by which specific signals that impinge on FimX_{XAC2398} could directly, or indirectly via

PilZ, regulate PilB function or contribute to the control of which ATPase – PilB or PilT – is positioned to control T4P polymerization or retraction.

normally E, N, E, and D residues, respectively, in active EAL domains).³⁸ This extensive modification of the EAL motif in FimX_{XAC2398} could explain our failure to detect PDE activity in purified recombinant FimX_{XAC2398} or its fragments under all conditions tested, including in the presence of GTP and/or PilZ_{XAC1133} (data not shown). In the case of GGDEF domains, two classes of c-diGMP binding sites have been identified: the catalytic (A) site and the inhibitory (I) site. These sites are defined by characteristic GG(D/E)EF and RXXD motifs, respectively, which are separated by a five residue linker.⁴⁵ These two motifs are severely substituted in both FimX_{XAC2398} (GEHSF and DXXV) and FimX from *P. aeruginosa* (GDSIF and GXXA), and so it is unlikely that c-diGMP binds to the GGDEF domains of these proteins. GTP binding by noncatalytic GGDEF domains has been reported,³⁷ and has been proposed for the GGDEF domain of *P. aeruginosa* FimX.²⁸ The above considerations are consistent with our inability to detect diguanylate cyclase or PDE activities for FimX_{XAC2398} and suggest that c-diGMP binding probably occurs via the EAL domain, even though we detected binding only when using the FimX_{GGDEF-EAL} fragment. The noncatalytic binding of c-diGMP by EAL domains has been shown to have physiological relevance in some cases. For example, the degenerate and inactive EAL domain of LapD from *Pseudomonas fluorescens* acts as a sensor of cytoplasmic c-diGMP levels in the control of localization of the LapA adhesin on the bacterial surface.³⁶

While no interaction between PilZ and EAL domains has been reported, recent studies have shown that alginate biosynthesis in *P. aeruginosa* is controlled by two membrane-associated proteins; Alg44, which binds c-diGMP via a PilZ domain,^{24,46} and by MucR, which contains C-terminal GGDEF and EAL domains.⁴⁷ It has been proposed that MucR regulates alginate biosynthesis through generation of a localized c-diGMP pool in the vicinity of Alg44.⁴⁷ Another nonexclusive possibility, suggested by our results, is that Alg44 and MucR could bind to one another directly via interactions involving their respective PilZ and EAL domains. Furthermore, this interaction could be modulated by the binding of c-diGMP to either or both proteins. This raises the possibility that PilZ-EAL domain interactions are not restricted to the FimX-PilZ-T4P system described here but, rather, could be a more general phenomenon. This is suggested also by the observation that PilZ and EAL domains are located close to each other at the C-termini of some proteins.²¹

The direct physical interactions of PilZ_{XAC1133} with FimX_{XAC2398} and PilB_{XAC3239} and the phenotypes of the *pilZ*_{XAC1133}, *fimX*_{XAC2398} and *pilB*_{XAC3239} knockout strains in surface motility assays is consistent with the little we know regarding the functions of their most closely related homologs in *P. aeruginosa*. The transport of PilA monomers to the bacterial surface is greatly reduced in *P. aeruginosa pilZ*, *fimX* and *pilB* mutants, while the cellular pools of PilA subunits are not modified significantly.^{18,7,48,49} Although we have

been unable to establish a reproducible twitching motility assay for *Xac*, we can observe its spreading across semisolid medium. This spreading is not swarming since it is not reduced in the *fliC_{XAC1975}* mutant strain. (Note that while bacterial swarming on wet surfaces requires flagella,⁵⁰ flagellum-independent migration on semi-solid surfaces has been observed in *Vibrio cholerae* and *Escherichia coli*⁵¹). Some reports have provided evidence that T4P are required for swarming motility in *P. aeruginosa*,^{52,53} while others have reported swarming in their absence.^{41,54,55} Under certain nutrient conditions, *P. aeruginosa* motility on swarm plates is increased in a *pilA* mutant that lacks T4P.⁵⁵ Furthermore, Murray and Kazmierczak have shown recently that *P. aeruginosa* exhibits sliding motility on a semisolid agar surface in the absence of both flagella and T4P.⁴¹ Sliding is a passive bacterial transport mechanism that is highly correlated with the secretion of surface-wetting substances, such as rhamnolipids, lipopeptides, lipopolysaccharides and glycolipids, that reduce surface tension and friction between cell and surface; in this way, the expansive forces within a growing culture can be harnessed as a means of transport.^{50,56} This motility is inhibited by surface structures such as T4P or flagella, possibly by increasing interaction between the bacterial cell and the substrate.⁴¹ The *pilZ_{XAC1133}*, *fimX_{XAC2398}*, *pilB_{XAC3239}*, and *fliC_{XAC1975}* phenotypes are therefore consistent with increased *Xac* sliding motility on semisolid agar in mutants lacking T4P or flagella.

Here, we provide the first demonstration of a physical interaction between PilZ and a T4P component (PilB), and provide evidence that the interaction involves a conserved motif at the disordered PilZ C-terminus. PilB is an ATPase that forms a hexameric ring structure associated with the cytoplasmic face of the inner membrane and has a central role in T4P biogenesis. The mechanism by which pilus assembly occurs at the inner membrane is unknown but a model based on the genetic and biochemical evidence as well as the available structural data on T4P components suggests that PilB and the inner membrane protein PilC work together to transfer pilin subunits from an inner membrane pool to the base of the growing pilus polymer, and insertion of each pilin monomer is associated with the hydrolysis of one ATP molecule: ATP hydrolysis induces a conformational change in PilB that is transmitted *via* PilC to drive the growing filament outward towards the periplasm.^{4,57} The homologous PilT ATPase presumably controls pilus retraction/depolymerization *via* a similar process but whose ATP-induced conformational change results in subunits being extracted from the base of the filament.^{4,57,58}

It is not clear whether the PilZ–PilB interaction controls PilB ATPase activity directly or whether PilZ acts as an adapter molecule to which other regulatory molecules may be assembled to interact transiently with PilB or other T4P subunits. The interaction between PilZ_{XAC1133} and the FimX_{XAC2398} points to a possible means by which specific

environmental signals can be transmitted to PilZ_{XAC1133} and, directly or indirectly, to PilB and the rest of the T4P apparatus. In *P. aeruginosa*, FimX localizes to the cell pole in a manner that depends on both EAL domain PDE activity and the N-terminal REC domain,²⁸ and the REC domain on its own has been shown to localize to the cell pole, presumably *via* interactions with unidentified proteins.⁷ The FimX_{XAC2398}–PilZ_{XAC1133}–PilB_{XAC3239} interaction observed in this study implies that FimX_{XAC2398} may be tethered to the membrane *via* both its N-terminal REC domain and its C-terminal EAL domain. Figure 9 presents a model of the T4P that has incorporated key features regarding what we know about PilB_{XAC3239}, FimX_{XAC2398} and PilZ_{PA2960/XAC1133} orthologous group function as well as possible mechanisms by which pilus biogenesis and function may be controlled *via* the interactions identified in this study. These possible mechanisms include: (i) regulation of the kinetics of the PilB ATPase cycle *via* direct interactions with PilZ or by PilB–FimX interactions that are too weak to be observed in the absence of PilZ; (ii) modulation of PilZ–PilB or FimX–PilB interactions by c-diGMP binding to FimX and/or the binding of other signaling molecules to other FimX domains; and (iii) the possibility that FimX–PilZ–PilB interactions may contribute to the control of which hexameric ATPase, PilB or PilT, is positioned to control T4P polymerization or retraction.

It has been proposed that c-diGMP signaling pathways operating in parallel avoid cross-talk (or c-diGMP “noise”) by sequestering the signaling modules of a pathway into specific microenvironments.^{45,59,60} This “functional sequestration” is realized *via* the formation of multidomain complexes that are held together within the same polypeptide chain or *via* noncovalent interactions.⁵⁹ Hengge has proposed that this local confinement has led to the evolution of signaling systems in which c-diGMP metabolism has been discarded.⁵⁹ The FimX–PilZ–PilB interactions described here are the first system that involves a full set of “degenerate” GGDEF, EAL and PilZ domains in which signaling to the final target, in this case the PilB ATPase, is mediated principally *via* protein–protein interactions. Whether these interactions are modulated by c-diGMP binding to FimX is still unclear.

Materials and Methods

Cloning

Full-length PilZ_{XAC1133} was cloned in expression plasmid pET3a and purified, as described.²⁹ Site-directed mutagenesis of the *pilZ_{XAC1133}* gene in the pET-3a vector was used to produce the Y22A, Y25A, F28A, W69A, R97A, E101A, L104 and Δ107–117 mutants *via* a single-step PCR protocol using the QuickChange Site-Directed Mutagenesis kit (STRATAGENE). Also, the mutant PilZ preys (Y22A and Δ107–117) in the two-hybrid vector pOAD were produced using the same primers and protocol. (The primers used for each mutation are listed in Supplementary Data Table 1D.) Coding sequences for

FimX_{XAC2398}, FimX_{GGDEF-EAL}, FimX_{EAL}, FimX_{GGDEF}, and PilB_{XAC3239} were amplified by PCR from *Xac* genomic DNA⁶¹ using specific primers and cloned into the expression vectors pET3a and/or pET28a,⁶² or into the two-hybrid bait vector pOBD.⁶³ (The primer sequences are listed in Supplementary Data Table 1A.) FimX_{XAC2398} and PilB_{XAC3239} correspond to the full-length proteins, while FimX_{GGDEF-EAL}, FimX_{GGDEF}, and FimX_{EAL} correspond to FimX_{XAC2398} residues 267–689, 267–432, and 436–689, respectively. All clones were confirmed by DNA sequencing.

Construction of *Xac* knockout strains

In-frame deletions in the *pilZ*_{XAC1133}, *xac2398* and *fliC*_{XAC1975} genes were obtained by a two-step allelic exchange procedure. Briefly, approximately 1 kb corresponding to upstream and downstream regions of the two genes were amplified by PCR (the primers are listed in Supplementary Data Table 1B.) Each pair of fragments was ligated to produce a deleted version of the gene and then cloned into the pNPTS138 suicide vector (Alley Dickon, unpublished results) in the restriction sites given in Supplementary Data Table 1B. These plasmids were introduced into *Xac* by electroporation, and sucrose-sensitive and kanamycin- and ampicillin-resistant colonies were selected. Liquid cultures of these colonies were grown without selection, plated and individual colonies were selected for simultaneous kanamycin sensitivity and sucrose resistance. Deletions were confirmed by PCR. In order to complement the *pilZ*_{XAC1133} and *xac2398* knockouts, fragments coding for *pilZ*_{XAC1133} and *xac2398* plus approximately 1 kb of upstream sequences were amplified using the primers described in Supplementary Data Table 1C and cloned into pUFR047.⁶⁴ The *pilZ*_{XAC1133} knockout was also complemented with the pMR20 vector⁶⁵ containing the coding sequences for PilZ_{XAC1133} and PilZ_{Δ107-117}. These sequences were amplified using the oligonucleotides given in Supplementary Data Table 1C and the respective pET3a-derived expression vectors as templates. The amplified fragments were cloned into the BamHI and EcoRI sites of pMR20. The pUFR047_ *pilZ*, pUFR047_ *fimX*, pMR20_ *pilZ* and pMR20_ *pilZ*_{Δ107-117} constructs were used to transform the *Xac* *pilZ*_{XAC1133} and *fimX*_{XAC2398} mutants to produce strains *pilZ*_{XAC1133}+pUFR047_ *pilZ*, *fimX*_{XAC2398}+pUFR047_ *fimX*, *pilZ*_{XAC1133}+pMR20_ *pilZ* and *pilZ*_{XAC1133}+pMR20_ *pilZ*_{Δ107-117}. In order to produce the *pilB*_{XAC3239} knockout strain, a 1000 bp fragment corresponding to an internal segment of the *pilB*_{XAC3239} gene was amplified using the oligonucleotides given in Supplementary Data Table 1B then cloned to TOPO TA cloning vector (Invitrogen), digested with EcoRI and ligated into pNPTS138. *Xac* cells were transformed using this construct (pNPTS_ *pilB*) and initially selected for growth in 100 µg/ml of kanamycin and sensitivity to 5% (w/v) sucrose. These colonies were grown in the presence of kanamycin without sucrose and kanamycin- and sucrose-resistant clones were selected. Clones were confirmed by PCR to have two truncated copies of the *pilB*_{XAC3239} gene.

Protein expression and purification

Conditions for protein expression (host *E. coli* strain and temperature) are summarized in Supplementary Data Table 1A. Proteins were produced by growing 1 l

cultures of transformed *E. coli* cells in 2xTY culture supplemented with ampicillin (100 mg/ml) or kanamycin (100 mg/ml) to an A_{600 nm} of 0.8, at which time isopropyl-D-thiogalactoside (IPTG) was added (final concentration 0.4 mM) and the cells were grown for 4 h before sedimentation and storage at –80°C. Cells were resuspended in 25 ml of 50 mM Tris-HCl (pH 8.0), 25% (w/v) sucrose, 1 mM EDTA, and lysed using a French press. EDTA was omitted for FimX_{GGDEF-EAL}. To purify FimX_{XAC2398}, the soluble fraction of the cell lysate was loaded onto a Q-Sepharose Fast Flow Hiload 16/10 column (Amersham Pharmacia) that was equilibrated with 50 mM Tris-HCl (pH 6.3), 14 mM β-mercaptoethanol, 1 mM EDTA. Bound proteins were eluted using a 0–1M NaCl gradient over 12 column volumes. Fractions containing FimX_{XAC2398} were loaded onto a Superdex 200 Fast Flow Hiload 26/60 column (Amersham Pharmacia) that was equilibrated with 50 mM Tris-HCl (pH 8.0), 14 mM β-mercaptoethanol, 250 mM NaCl. To purify FimX_{GGDEF-EAL}, which contains an N-terminal polyhistidine tag, the soluble fraction of the cell lysate was loaded onto a HisTrap HP column (GE Healthcare) that was equilibrated with 50 mM Tris-HCl (pH 8.0), 1 mM β-mercaptoethanol, 200 mM NaCl and eluted with a 0–1 M imidazole gradient. Fractions containing FimX_{GGDEF-EAL} were then separated by passage through a Superdex 200 column that was equilibrated with 5 mM Tris-HCl (pH 7.0), 10 mM NaCl. To purify FimX_{GGDEF}, the soluble fraction of the bacterial lysate was loaded onto a Q-Sepharose Fast Flow column that was equilibrated with 50 mM Tris-HCl (pH 8.0), 14 mM β-mercaptoethanol, 1 mM EDTA. Bound proteins were eluted using a 0–1 M NaCl gradient over 12 column volumes. Fractions containing FimX_{GGDEF} were pooled, concentrated and separated by passage through a Superdex 75 Fast Flow column that was equilibrated with 50 mM Tris-HCl (pH 8.0), 14 mM β-mercaptoethanol, 250 mM NaCl. To purify FimX_{EAL}, the soluble fraction of the bacterial cell lysate was loaded onto a Q-Sepharose Fast Flow Hiload column that was equilibrated with 50 mM Tris-HCl (pH 7.0), 14 mM β-mercaptoethanol, 1 mM EDTA. Bound proteins were eluted using a 0–0.5 M NaCl gradient over 12 column volumes. Fractions containing XAC2398EAL were pooled, concentrated, and loaded onto a Superdex 75 Fast Flow column that was equilibrated with 5 mM Tris-HCl (pH 7.0), 1 mM β-mercaptoethanol, 10 mM NaCl. After the final chromatographic separations, fractions containing FimX_{XAC2398}, FimX_{GGDEF-EAL}, FimX_{GGDEF} or FimX_{EAL} were dialyzed separately against 5 mM Tris-HCl (pH 7.0) and concentrated to approximately 5 g/l using Centricon (Millipore) concentrators with 10 kDa cut-off.

Two-hybrid assays

The construction and analysis of the *Xac* genomic DNA library in the plasmid vector pOAD⁶³ was as described.^{32,66} *Saccharomyces cerevisiae* strain PJ694-a (*MATa trp1-901 leu2-3, 112 ura3-52 his3-200 gal4Δ gal80Δ LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ*)⁶⁷ was grown in SC medium (0.66% (w/v) nitrogen base without amino acids, 2% (w/v) glucose, 0.008% (w/v) adenine, 0.083% (w/v) amino acids mixture, pH 5.6). When necessary, SC medium was prepared lacking one or more of the components adenine, histidine, tryptophan, and leucine. Solid medium also contained 1.6% (w/v) Bacto-agar and 5 mM 3-aminotriazole. β-

galactosidase activity was measured by selecting cells from SC-WL or SC-WHLA plates to inoculate SC-WL or SC-WHLA liquid cultures that were grown overnight at 30 °C, collected by centrifugation, resuspended in 1 ml of YPD medium (20 g/l Tryptone, 10 g/l yeast extract, 20 g/l glucose, pH 5.8) and grown for another 4 h. Cells were collected, resuspended in the same volume of Z-buffer (0.06 M Na₂HPO₄, 0.04 M NaH₂PO₄, 0.01 M KCl, 0.001 M MgSO₄, 0.05 M β-mercaptoethanol, pH 7.0) and the $A_{600\text{nm}}$ of the resuspended cells was measured. After that, 100 μl of chloroform and 50 μl of 0.1% (w/v) SDS were added, vortex mixed and incubated for 5 min at 28 °C. The reaction was initiated by adding 0.2 ml of 4 mg/mL *o*-nitrophenyl-β-d-galactoside (ONPG), vortex mixed, incubated at 28 °C and time measured until the reaction was terminated by the addition of 0.5 ml of 1 M Na₂CO₃. The solution was centrifuged for 5 min at 13000g and the $A_{420\text{nm}}$ and at $A_{550\text{nm}}$ were recorded. The Miller units were calculated as:

$$\text{Miller units} = 1000 [(A_{420\text{nm}} - 1.75 A_{550\text{nm}})] / (TVA_{600\text{nm}})$$

where T is reaction time (in minutes), and V is volume (in ml).

Structure determination and refinement

The production and crystallization of selenomethionine-labeled PilZ_{XAC1133} was as described.²⁹ Briefly, crystallization was carried out in sitting-drop plates by mixing equal volumes (1.5 μl) of 7.2 mg/ml protein and well solution (24% (w/v) PEG 4000, 0.1 M Tris-HCl (pH 8.0), 0.2 M MgCl₂) and growing at 281 K. A MAD dataset from a single crystal that diffracted to 1.9 Å was collected as described.²⁹ No σ cutoff was used during the refinement. Selenium sites were found using SHELXD,⁶⁸ and their positions were subsequently refined and phases calculated using SHARP.⁶⁹ Phases were refined by density modification using SOLOMON⁷⁰ and DM.⁷¹ The figure of merit for acentric and centric reflections was 0.64 and 0.27, respectively, for data up to 2.42 Å and 0.34 and 0.18 for the data up to 1.9 Å. The resulting electron density map was used to construct a preliminary polyaniline model with ARP/wARP⁷² that included 82 of the 117 residues distributed over three peptide segments. Interpretation of the electron density maps and construction of the missing residues was done with COOT.⁷³ Initially, the space group was incorrectly assigned as $P6_122$ and the 2-fold axes were close enough to real symmetry axes to allow MAD phasing and model building, but not close enough for refinement convergence. The data were reprocessed in space group $P6_1$ and the second PilZ_{XAC1133} molecule in the asymmetric unit was obtained using Phaser.⁷⁴ Structural refinement of the PilZ_{XAC1133} model was done using CCP4,⁷⁵ REFMAC,⁷⁶ CNS,⁷⁷ and COOT.⁷³ TLS⁷⁸ was used in the final cycles of refinement with each chain in the asymmetric unit divided into 15 independent groups. Water molecules were added with ARP/wARP, REFMAC, and directly by COOT. The first eight residues in chain A, the first seven residues in chain B, the last 12 residues for chain A, the last 13 residues for chain B and an internal loop (residues 73–79) of both chains were not modeled due to the absence of electron density. The final R -factor, R_{free} and figure of merit values are 0.190, 0.244 and 0.82, respectively. The refinement data statistics are given in Table 1. Analysis of the final PilZ_{XAC1133} model with PROCHECK6 showed that 95.7% of the residues were in most favored regions and 4.3% were in additionally allowed regions in the Ramachandran plot. The

coordinates of the final PilZ_{XAC1133} model have been deposited in the Protein Data Bank (see below).

NMR spectroscopy

NMR samples of ¹³C,¹⁵N-labeled PilZ_{XAC1133} or ¹⁵N-labeled PilZ_{XAC1133} mutants, consisted of 0.5 mM protein (unless stated otherwise) in 20 mM Tris-HCl (pH 7.0), 0.150 M NaCl. The sequential assignment of backbone (¹H, ¹⁵N, ¹³C^α, ¹³C^β) resonances of PilZ_{XAC1133} was done by the analysis of CBCA(CO)NH, HNCACB, HN(CO)CA and HNCA spectra⁷⁹ acquired at 35 °C and at 27 °C in a Varian INOVA 600 MHz spectrometer equipped with a 5 mm HCN PFG cold probe. Measurements of heteronuclear [¹H]-¹⁵N NOE were carried out in a Bruker DRX spectrometer, operating at 500 MHz, and equipped with a 5 mm TXI probe at 27 °C, using saturation time of 3 s and an inter-scan delay of 1 s. Saturated and control spectra were acquired in interleaved fashion. ¹⁵N HSQC experiments of ¹⁵N-labeled PilZ_{XAC1133} and PilZ mutants in the presence of unlabeled FimX_{XAC2398}, FimX_{EAL}, FimX_{EAL-GGDEF} or FimX_{GGDEF} were acquired at 27 °C and at 600 MHz, except for the Y22A mutant that was collected at 500 MHz. NMR spectra were processed using NMRPipe,⁸⁰ and analyzed using the CCPN Analysis software.⁸¹

Far Western protein–protein interaction assays

Purified PilZ_{XAC1133} was used to obtain polyclonal antiserum by immunizing rabbits. Lysates of *E. coli* cultures expressing recombinant PilB_{XAC3239} protein (or Xac YaeQ protein as a negative control) were separated by SDS-PAGE and transferred to nitrocellulose membranes. The membranes were probed with purified recombinant PilZ_{XAC1133} or PilZ_{XAC1133} mutants (1 μM) for 12 h at room temperature. Where indicated, XAC2398_{GGDEF-EAL} and/or c-diGMP was also added at concentrations of 1–4 μM. Membranes were then washed and incubated with anti-PilZ_{XAC1133} or anti-His tag antiserum (to detect XAC2398-GGDEF-EAL), followed by protein A-peroxidase conjugate and detection of peroxidase activity.

Motility assays

Motility was measured on the surface of semisolid medium. The following media were used: SB (5 g/l sucrose, 5 g/l yeast extract, 5 g/l Peptone, and 1 g/l glutamic acid, pH 7.0, 0.6% Bacto-agar), XVM2 (20 mM NaCl, 10 mM (NH₄)₂SO₄, 1 mM CaCl₂, 0.01 mM FeSO₄, 5 mM MgSO₄, 0.16 mM KH₂PO₄, 0.32 mM K₂HPO₄, 10 mM fructose, 10 mM sucrose, 0.03% (v/v) casein acid hydrolysate (Casaminoacids), pH 6.7, 0.6% Bacto-agar), STON (1 g/l Tryptone, 0.3% yeast extract, 0.3% Bacto-agar), and EIKEN (8 g/l Eiken nutrient broth, 0.6% (w/v) Eiken agar). Molten medium was poured into plates and allowed to solidify and dry for 1 h. *Xac* cells were grown in 2xTY to mid-log phase at 28 °C, diluted to $A_{600\text{nm}}$ of 0.3 and 5 μl aliquots were used to inoculate the semi-solid surface of the medium followed by incubation at 28 °C for two to four days.

Protein stability studies

Thermal unfolding was followed by measuring $[\theta]_{205\text{nm}}$ or $[\theta]_{220\text{nm}}$ as the temperature was increased (0.2 °C/min). The fraction of unfolded protein was calculated by using

the pre-transition and post-transition baselines to calculate the fractional change in signal at each temperature point as described.⁸² Protein concentrations were as indicated in the figures and buffer conditions were 10 mM Tris-HCl, pH 7.0, 10 mM NaCl and 1 mM MgCl₂. Where indicated, 5 μ M c-diGMP, GTP or ATP was added.

Protein Data Bank accession codes

The atomic coordinates and structure factors of PilZ_{XAC1133} have been deposited in the Protein Data Bank (PDB) with accession code 3CNR. The resonance assignments for PilZ_{XAC1133} have been deposited in the (Biological Magnetic Resonance Data Bank) with accession number 16069.

Acknowledgements

We thank the Brazilian Synchrotron Light Laboratory (LNLS) and the Central Analítica of the Institute of Chemistry of the University of São Paulo for providing access to the MX2 beamline and Varian INOVA 600 MHz and the Bruker DRX 500 MHz spectrometers. We thank M. Uemi, A. C. Zeri, M. Sorça, R. Rabelo-Schefer, R. de Oliveira, F. Manso Prado, L. Navarro, A. do Amaral, P. Cappelletti, C. Docena, I. de Souza Costa, R. Nagem, B. Guimaraes and L. Sanfelici for technical help and discussions. This work was funded by FAPESP and CNPq, Brazil.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2009.07.065

References

- Mattick, J. S. (2002). Type IV pili and twitching motility. *Annu. Rev. Microbiol.* **56**, 289–314.
- Craig, L., Pique, M. E. & Tainer, J. A. (2004). Type IV pilus structure and bacterial pathogenicity. *Nature Rev. Microbiol.* **2**, 363–378.
- Han, X., Kennan, R. M., Parker, D., Davies, J. K. & Rood, J. I. (2007). Type IV fimbrial biogenesis is required for protease secretion and natural transformation in *Dichelobacter nodosus*. *J. Bacteriol.* **189**, 5022–5033.
- Craig, L. & Li, J. (2008). Type IV pili: paradoxes in form and function. *Curr. Opin. Struct. Biol.* **18**, 267–277.
- Miller, R. M., Tomaras, A. P., Barker, A. P., Voelker, D. R., Chan, E. D., Vasil, A. I. & Vasil, M. L. (2008). *Pseudomonas aeruginosa* twitching motility-mediated chemotaxis towards phospholipids and fatty acids: specificity and metabolic requirements. *J. Bacteriol.* **190**, 4038–4049.
- Pellic, V. (2008). Type IV pili: e pluribus unum? *Mol. Microbiol.* **68**, 827–837.
- Huang, B., Whitchurch, C. B. & Mattick, J. S. (2003). FimX, a multidomain protein connecting environmental signals to twitching motility in *Pseudomonas aeruginosa*. *J. Bacteriol.* **185**, 7068–7076.
- Hobbs, M., Collie, E. S., Free, P. D., Livingston, S. P. & Mattick, J. S. (1993). PilS and PilR, a two-component transcriptional regulatory system controlling expression of type 4 fimbriae in *Pseudomonas aeruginosa*. *Mol. Microbiol.* **7**, 669–682.
- da Silva Neto, J. F., Koide, T., Abe, C. M., Gomes, S. L. & Marques, M. V. (2008). Role of sigma54 in the regulation of genes involved in type I and type IV pili biogenesis in *Xylella fastidiosa*. *Arch. Microbiol.* **189**, 249–261.
- O'Toole, G. A., Gibbs, K. A., Hager, P. W., Phibbs, P. V., Jr & Kolter, R. (2000). The global carbon metabolism regulator Crc is a component of a signal transduction pathway required for biofilm development by *Pseudomonas aeruginosa*. *J. Bacteriol.* **182**, 425–431.
- Pratt, J. T., Tamayo, R., Tischler, A. D. & Camilli, A. (2007). PilZ domain proteins bind cyclic diguanylate and regulate diverse processes in *Vibrio cholerae*. *J. Biol. Chem.* **282**, 12860–12870.
- Skerker, J. M. & Berg, H. C. (2001). Direct observation of extension and retraction of type IV pili. *Proc. Natl Acad. Sci. USA*, **98**, 6901–6904.
- Aas, F. E., Wolfgang, M., Frye, S., Dunham, S., Lovold, C. & Koomey, M. (2002). Competence for natural transformation in *Neisseria gonorrhoeae*: components of DNA binding and uptake linked to type IV pilus expression. *Mol. Microbiol.* **46**, 749–760.
- Lory, S. & Strom, M. S. (1997). Structure-function relationship of type-IV prepilin peptidase of *Pseudomonas aeruginosa* – a review. *Gene*, **192**, 117–121.
- Chiang, P., Habash, M. & Burrows, L. L. (2005). Disparate subcellular localization patterns of *Pseudomonas aeruginosa* Type IV pilus ATPases involved in twitching motility. *J. Bacteriol.* **187**, 829–839.
- Bitter, W., Koster, M., Latijnhouwers, M., de Cock, H. & Tommassen, J. (1998). Formation of oligomeric rings by XcpQ and PilQ, which are involved in protein transport across the outer membrane of *Pseudomonas aeruginosa*. *Mol. Microbiol.* **27**, 209–219.
- Merz, A. J., So, M. & Sheetz, M. P. (2000). Pilus retraction powers bacterial twitching motility. *Nature*, **407**, 98–102.
- Alm, R. A., Boder, A. J., Free, P. D. & Mattick, J. S. (1996). Identification of a novel gene, pilZ, essential for type 4 fimbrial biogenesis in *Pseudomonas aeruginosa*. *J. Bacteriol.* **178**, 46–53.
- Carbannelle, E., Helaine, S., Prouvensier, L., Nassif, X. & Pellic, V. (2005). Type IV pilus biogenesis in *Neisseria meningitidis*: PilW is involved in a step occurring after pilus assembly, essential for fibre stability and function. *Mol. Microbiol.* **55**, 54–64.
- McCarthy, Y., Ryan, R. P., O'Donovan, K., He, Y. Q., Jiang, B. L., Feng, J. X. *et al.* (2008). The role of PilZ domain proteins in the virulence of *Xanthomonas campestris* pv. *campestris*. *Mol. Plant Pathol.* **9**, 819–824.
- Amikam, D. & Galperin, M. Y. (2006). PilZ domain is part of the bacterial c-di-GMP binding protein. *Bioinformatics*, **22**, 3–6.
- Benach, J., Swaminathan, S. S., Tamayo, R., Handel, S. K., Foltz-Stogniew, E., Ramos, J. E. *et al.* (2007). The structural basis of cyclic diguanylate signal transduction by PilZ domains. *EMBO J.* **26**, 5153–5166.
- Ryjenkov, D. A., Simm, R., Romling, U. & Gomelsky, M. (2006). The PilZ domain is a receptor for the second messenger c-di-GMP: the PilZ domain protein YcgR

- controls motility in enterobacteria. *J. Biol. Chem.* **281**, 30310–30314.
24. Merighi, M., Lee, V. T., Hyodo, M., Hayakawa, Y. & Lory, S. (2007). The second messenger bis-(3'-5')-cyclic-GMP and its PilZ domain-containing receptor Alg44 are required for alginate biosynthesis in *Pseudomonas aeruginosa*. *Mol. Microbiol.* **65**, 876–895.
  25. Christen, M., Christen, B., Allan, M. G., Folcher, M., Jenö, P., Grzesiek, S. & Jenal, U. (2007). DgrA is a member of a new family of cyclic diguanosine monophosphate receptors and controls flagellar motor function in *Caulobacter crescentus*. *Proc. Natl Acad. Sci. USA*, **104**, 4112–4117.
  26. Li, T. N., Chin, K. H., Liu, J. H., Wang, A. H. & Chou, S. H. (2009). XC1028 from *Xanthomonas campestris* adopts a PilZ domain-like structure without a c-di-GMP switch. *Proteins: Struct. Funct. Genet.* **75**, 282–288.
  27. Ramelot, T. A., Yee, A., Cort, J. R., Semesi, A., Arrowsmith, C. H. & Kennedy, M. A. (2007). NMR structure and binding studies confirm that PA4608 from *Pseudomonas aeruginosa* is a PilZ domain and a c-di-GMP binding protein. *Proteins: Struct. Funct. Genet.* **66**, 266–271.
  28. Kazmierczak, B. I., Lebron, M. B. & Murray, T. S. (2006). Analysis of FimX, a phosphodiesterase that governs twitching motility in *Pseudomonas aeruginosa*. *Mol. Microbiol.* **60**, 1026–1043.
  29. Guzzo, C. R. & Farah, C. S. (2009). Expression, crystallization and preliminary crystallographic analysis of PilZ(XAC1133) from *Xanthomonas axonopodis* pv. citri. *Acta Crystallogr. F*, **65**, 304–306.
  30. Cornilescu, G., Delaglio, F. & Bax, A. (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR*, **13**, 289–302.
  31. Finn, R. D., Tate, J., Mistry, J., Coghill, P. C., Sammut, S. J., Hotz, H. R. et al. (2008). The Pfam protein families database. *Nucleic Acids Res.* **36**, D281–D288.
  32. Alegria, M. C., Docena, C., Khater, L., Ramos, C. H., da Silva, A. C. & Farah, C. S. (2004). New protein-protein interactions identified for the regulatory and structural components and substrates of the type III Secretion system of the phytopathogen *Xanthomonas axonopodis* Pathovar citri. *J. Bacteriol.* **186**, 6186–6197.
  33. Bobrov, A. G., Kirillina, O. & Perry, R. D. (2005). The phosphodiesterase activity of the HmsP EAL domain is required for negative regulation of biofilm formation in *Yersinia pestis*. *FEMS Microbiol. Lett.* **247**, 123–130.
  34. Chan, C., Paul, R., Samoray, D., Amiot, N. C., Giese, B., Jenal, U. & Schirmer, T. (2004). Structural basis of activity and allosteric control of diguanylate cyclase. *Proc. Natl Acad. Sci. USA*, **101**, 17084–17089.
  35. Schmidt, A. J., Ryjenkov, D. A. & Gomelsky, M. (2005). The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: enzymatically active and inactive EAL domains. *J. Bacteriol.* **187**, 4774–4781.
  36. Newell, P. D., Monds, R. D. & O'Toole, G. A. (2009). LapD is a bis-(3',5')-cyclic dimeric GMP-binding protein that regulates surface attachment by *Pseudomonas fluorescens* Pf0-1. *Proc. Natl Acad. Sci. USA*, **106**, 3461–3466.
  37. Christen, M., Christen, B., Folcher, M., Schauerte, A. & Jenal, U. (2005). Identification and characterization of a cyclic di-GMP-specific phosphodiesterase and its allosteric control by GTP. *J. Biol. Chem.* **280**, 30829–30837.
  38. Rao, F., Yang, Y., Qi, Y. & Liang, Z. X. (2008). Catalytic mechanism of cyclic di-GMP-specific phosphodiesterase: a study of the EAL domain-containing RocR from *Pseudomonas aeruginosa*. *J. Bacteriol.* **190**, 3622–3631.
  39. Jakovljevic, V., Leonardy, S., Hoppert, M. & Sogaard-Andersen, L. (2008). PilB and PilT are ATPases acting antagonistically in type IV pilus function in *Myxococcus xanthus*. *J. Bacteriol.* **190**, 2411–2421.
  40. Sudarsan, N., Lee, E. R., Weinberg, Z., Moy, R. H., Kim, J. N., Link, K. H. & Breaker, R. R. (2008). Riboswitches in eubacteria sense the second messenger cyclic di-GMP. *Science*, **321**, 411–413.
  41. Murray, T. S. & Kazmierczak, B. I. (2008). *Pseudomonas aeruginosa* exhibits sliding motility in the absence of type IV pili and flagella. *J. Bacteriol.* **190**, 2700–2708.
  42. Lee, V. T., Matewish, J. M., Kessler, J. L., Hyodo, M., Hayakawa, Y. & Lory, S. (2007). A cyclic-di-GMP receptor required for bacterial exopolysaccharide production. *Mol. Microbiol.* **65**, 1474–1484.
  43. Hickman, J. W. & Harwood, C. S. (2008). Identification of FleQ from *Pseudomonas aeruginosa* as a c-di-GMP-responsive transcription factor. *Mol. Microbiol.* **69**, 376–389.
  44. Ross, P., Mayer, R., Weinhouse, H., Amikam, D., Huggir, Y., Benziman, M. et al. (1990). The cyclic diguanylic acid regulatory system of cellulose synthesis in *Acetobacter xylinum*. Chemical synthesis and biological activity of cyclic nucleotide dimer, trimer, and phosphothioate derivatives. *J. Biol. Chem.* **265**, 18933–18943.
  45. Jenal, U. & Malone, J. (2006). Mechanisms of cyclic-di-GMP signaling in bacteria. *Annu. Rev. Genet.* **40**, 385–407.
  46. Oglesby, L. L., Jain, S. & Ohman, D. E. (2008). Membrane topology and roles of *Pseudomonas aeruginosa* Alg8 and Alg44 in alginate polymerization. *Microbiology*, **154**, 1605–1615.
  47. Hay, I. D., Remminghorst, U. & Rehm, B. H. (2009). MucR, a novel membrane-associated regulator of alginate biosynthesis in *Pseudomonas aeruginosa*. *Appl. Environ. Microbiol.* **75**, 1110–1120.
  48. Turner, L. R., Lara, J. C., Nunn, D. N. & Lory, S. (1993). Mutations in the consensus ATP-binding sites of XcpR and PilB eliminate extracellular protein secretion and pilus biogenesis in *Pseudomonas aeruginosa*. *J. Bacteriol.* **175**, 4962–4969.
  49. Nunn, D., Bergman, S. & Lory, S. (1990). Products of three accessory genes, pilB, pilC, and pilD, are required for biogenesis of *Pseudomonas aeruginosa* pili. *J. Bacteriol.* **172**, 2911–2919.
  50. Harshey, R. M. (2003). Bacterial motility on a surface: many ways to a common goal. *Annu. Rev. Microbiol.* **57**, 249–273.
  51. Brown, II & Hase, C. C. (2001). Flagellum-independent surface migration of *Vibrio cholerae* and *Escherichia coli*. *J. Bacteriol.* **183**, 3784–3790.
  52. Kohler, T., Curty, L. K., Barja, F., van Delden, C. & Pechere, J. C. (2000). Swarming of *Pseudomonas aeruginosa* is dependent on cell-to-cell signaling and requires flagella and pili. *J. Bacteriol.* **182**, 5990–5996.
  53. Overhage, J., Lewenza, S., Marr, A. K. & Hancock, R. E. (2007). Identification of genes involved in swarming motility using a *Pseudomonas aeruginosa* PAO1 mini-Tn5-lux mutant library. *J. Bacteriol.* **189**, 2164–2169.
  54. Rashid, M. H. & Kornberg, A. (2000). Inorganic polyphosphate is needed for swimming, swarming,

- and twitching motilities of *Pseudomonas aeruginosa*. *Proc. Natl Acad. Sci. USA*, **97**, 4885–4890.
55. Shrout, J. D., Chopp, D. L., Just, C. L., Hentzer, M., Givskov, M. & Parsek, M. R. (2006). The impact of quorum sensing and swarming motility on *Pseudomonas aeruginosa* biofilm formation is nutritionally conditional. *Mol. Microbiol.* **62**, 1264–1277.
  56. Henrichsen, J. (1972). Bacterial surface translocation: a survey and a classification. *Bacteriol. Rev.* **36**, 478–503.
  57. Craig, L., Volkmann, N., Arvai, A. S., Pique, M. E., Yeager, M., Egelman, E. H. & Tainer, J. A. (2006). Type IV pilus structure by cryo-electron microscopy and crystallography: implications for pilus assembly and functions. *Mol. Cell*, **23**, 651–662.
  58. Savvides, S. N. (2007). Secretion superfamily ATPases swing big. *Structure*, **15**, 255–257.
  59. Hengge, R. (2009). Principles of c-di-GMP signalling in bacteria. *Nature Rev. Microbiol.* **7**, 263–273.
  60. Ryan, R. P., Fouhy, Y., Lucey, J. F. & Dow, J. M. (2006). Cyclic di-GMP signaling in bacteria: recent advances and new puzzles. *J. Bacteriol.* **188**, 8327–8334.
  61. da Silva, A. C., Ferro, J. A., Reinach, F. C., Farah, C. S., Furlan, L. R., Quaggio, R. B. *et al.* (2002). Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature*, **417**, 459–463.
  62. Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60–89.
  63. Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R. *et al.* (2000). A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature*, **403**, 623–627.
  64. De Feyter, R., Yang, Y. & Gabriel, D. W. (1993). Gene-for-genes interactions between cotton R genes and *Xanthomonas campestris* pv. *malvacearum* avr genes. *Mol. Plant Microbe Interact.* **6**, 225–237.
  65. Roberts, R. C., Mohr, C. D. & Shapiro, L. (1996). Developmental programs in bacteria. *Curr. Top. Dev. Biol.* **34**, 207–257.
  66. Alegria, M. C., Souza, D. P., Andrade, M. O., Docena, C., Khater, L., Ramos, C. H. *et al.* (2005). Identification of new protein-protein interactions involving the products of the chromosome- and plasmid-encoded type IV secretion loci of the phytopathogen *Xanthomonas axonopodis* pv. *citri*. *J. Bacteriol.* **187**, 2315–2325.
  67. James, P., Halladay, J. & Craig, E. A. (1996). Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics*, **144**, 1425–1436.
  68. Schneider, T. R. & Sheldrick, G. M. (2002). Substructure solution with SHELXD. *Acta Crystallogr. D*, **58**, 1772–1779.
  69. La Fortelle, E. D. & Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* **276**, 472–494.
  70. Abrahams, J. P. & Leslie, A. G. (1996). Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr. D*, **52**, 30–42.
  71. Cowtan, K. (1994). 'dm': an automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter Protein Crystallogr.* **31**, 34–38.
  72. Perrakis, A., Morris, R. & Lamzin, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.* **6**, 458–463.
  73. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D*, **60**, 2126–2132.
  74. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674.
  75. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D*, **50**, 760–763.
  76. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D*, **53**, 240–255.
  77. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D*, **54**, 905–921.
  78. Painter, J. & Merritt, E. A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D*, **62**, 439–450.
  79. Sattler, M., S. J., Griesinger C. (1999). Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Progr. Nuclear Magn. Reson. Spectrosc.* **34**, 93–158.
  80. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
  81. Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M. *et al.* (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins: Struct. Funct. Genet.* **59**, 687–696.
  82. Pace, C. N. & Sholtz, J. M. (1997). Determining the conformational stability of a protein. In (Creighton, T. E., ed.), pp. 299–321, 2nd edit IRL Press, Oxford.
  83. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & M. T. J. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.

Supplemental Electronic Information for:

**PILZ PROTEIN STRUCTURE AND INTERACTIONS WITH
PILB AND THE FIMX EAL DOMAIN: IMPLICATIONS FOR
CONTROL OF TYPE IV PILUS BIOGENESIS**

Cristiane R. Guzzo¹, Roberto K. Salinas¹, Maxuel O.

Andrade¹

and Chuck S. Farah^{1*}

From Departamento de Bioquímica, Instituto de Química,
Universidade de São Paulo, CEP 05508-000, SP, Brazil.

Running head: PilZ structure and interactions

Address correspondence to: Chuck S. Farah, Inst. Química,
Universidade de São Paulo, Av. Prof. Lineu Prestes 748,

São Paulo, SP, 05508-000, Brazil

(+55-11) 3091-8519, Fax (+55-11) 3915-5579; E-mail:

chsfarah@iq.usp.br

This Supplemental Information consists of:

- Supplemental Figures 1-8
- Supplemental Table 1

Supplemental Figures

Supplementary Figure 1.

Supplemental Figure 1. Sequence conservation in PilZ homology domains. There are now over 1500 PilZ sequences annotated in the Pfam database¹. **A)** A representative sample of these (the 191 sequences in the Pfam seed alignment) was used to produce a sequence conservation profile. Only 10 positions (circled) show a significant level of conservation. The first five are involved in c-diGMP binding (R,R,D,S,G). Essentially the same profile is obtained when aligning all 1574 members of the Pfam PilZ family (data not shown). **B)** A simple Blast search² of the public sequence databases using the PilZ_{XAC1133} sequence identifies approximately 100 closely related sequences with e-values < 10⁻⁵ and identities greater than 33%. The resultant alignment³ and sequence conservation profile of the PA2960/XAC1133 orthologous group is shown. Note that the sequence profile at the N-terminal region of the protein is modified and the c-diGMP binding motif has been lost. **C)** As the main structural differences observed in the PilZ_{XAC1133} structure correspond to the first two secondary structure elements, up to the highly conserved glycine at position 33, we used the first 33 PilZ_{XAC1133} residues to perform an iterative Psi Blast⁴ search of the public protein sequence database. After 9 iterations, 76 proteins were identified with e-values below 10⁻⁴ whose sequence conservation profile is shown here and is, as expected, very similar to the profile shown in (B). Numbering in (B) corresponds to that of the PilZ_{XAC1133} sequence while the numbering schemes in (A) and (C) are arbitrary and derived directly from the alignments. **D)** Highly conserved residues map to one face of the PilZ_{XAC1133} structure. Cartoon and semi-transparent surface depiction of the PilZ_{XAC1133} structure in which conserved motifs MI-MIV are colored red. Parts A-C of the figure were generated using the WebLogo server (<http://weblogo.berkeley.edu/>)

Supplemental Figure 2. Comparison of PilZ_{XAC1133} topology with that of other PilZ domains. Cartoon diagrams of **A)** PilZ_{XAC1133} in which secondary structure elements have been labeled, **B)** PA4608 from *P. aeruginosa* (1YWU), **C)** the PilZ domain (residues 138-247) of VCA0042 from *V. cholera* and **D)** the PilZ domain (residues 131-247) from PP4397 from *P. putida* (2GJG).

Supplemental Fig. 3. **A)** Comparison of the ^1H - ^{15}N HSQC spectra of PilZ_{XAC1133} in the absence (red) and presence (blue) of c-diGMP at a molar ratio of 1:1. The spectra were acquired at 600 MHz as data matrices of 512x128 complex points. The protein and c-diGMP concentrations were 225 μM . Conditions: 5mM Tris pH 7.0, 10mM NaCl and 10mM MgCl_2 . **B)** Conformational stability of PilZ_{XAC1133} in the absent and present of c-diGMP. Conditions were: 10 μM protein in 10 mM Tris-HCl, pH 7.0, 10 mM NaCl and 1 mM MgCl_2 . Where indicated, 5 μM c-diGMP was added. The denaturation was accompanied by measuring $[\theta]_{205\text{nm}}$ and the signal was converted to fraction unfolded protein (Yu) as described in Methods.

Supplemental Figure 4. Conformational stability of PilZ_{XAC1133} and FimX_{XAC2398}. **A and B,** Thermal denaturation of PilZ_{XAC1133} (PilZ) and its mutants F28A, W69A, E101A, Δ 107-117 (**A**) and Y25A, R97A, L104A (**B**).

PilZ_{XAC1133}/PilZ_{Y22A} **PilZ_{Y22A} + FimX_{XAC2398}** **PilZ_{XAC1133}/PilZ_{F28A}** **PilZ_{F28A} + FimX_{XAC2398}**

PilZ_{XAC1133}/PilZ_{W69A} **PilZ_{W69A} + FimX_{XAC2398}** **PilZ_{XAC1133}/PilZ_{R97A}** **PilZ_{R97A} + FimX_{XAC2398}**

PilZ_{XAC1133}/PilZ_{E101A} **PilZ_{E101A} + FimX_{XAC2398}** **PilZ_{XAC1133}/PilZ_{L104A}** **PilZ_{L104A} + FimX_{XAC2398}**

Supplemental Figure 5. Perturbations in the PilZ_{XAC1133} ¹⁵N HSQC spectrum caused by specific mutations and by binding of FimX_{XAC2398}. The ¹H-¹⁵N HSQC spectrum of wild type PilZ_{XAC1133} (red) was superimposed on the spectra of PilZ_{XAC1133} mutants (blue). In addition, the ¹⁵N HSQC spectra of the PilZ_{XAC1133} mutants (0.1 – 0.2 mM) acquired in the absence (blue) and in the presence (black) of FimX_{XAC2398} at a 1:1 molar ratio are compared. All spectra were referenced for the position of the water signal using NMRPipe⁷. Specific resonances in the wild-type spectra whose chemical shifts changed in the Y22A, F28A, W69A, R97A, and Δ107-117 mutant spectra are labeled. While the spectrum for PilZ_{E101A} presented sharp lines, a large number of the resonances were modified with respect to the wild-type protein. The ¹H-¹⁵N HSQC spectrum of PilZ_{L104A}, presented extreme line broadening. The same was observed for PilZ_{Y25A} (data not shown).

Supplemental Figure 6. Mapping of perturbations in the $\text{PilZ}_{\text{XAC1133}}$ ^{15}N HSQC spectrum caused by specific mutations in the $\text{PilZ}_{\text{XAC1133}}$ structure. In the discussion that follows, residues whose ^{15}N - ^1H HSQC cross peaks were significantly perturbed by the mutations through visual comparison with the wild-type HSQC spectrum (see **Supplemental Figure 5**). The mutated residues are colored blue and the significantly perturbed residues are white or red as indicated. **A)** The Y25A and L104A substitutions both resulted in extreme line-broadening in the ^{15}N - ^1H HSQC spectra (part (a) and data not shown). This could be due to disruption of a network of hydrophobic interactions involving these two residues and Y22 that could be important for stabilizing the structure. This is consistent with reduced stabilities of mutants Y25A and L104A (**Supplemental Fig. 4B**). **B)** The Y22A substitution affects cross-peaks of nearby residues within $\alpha 1$ S23, A24, Y25 (white) and residues located in the C-terminal A105 (red), L107, D111, T114 and T116 (no electron density). Furthermore, the resonances of three lysine residues, K16, K40 and K66 (red) presented a significant reduction in peak intensity. **C)** Perturbations caused by the F28A mutant are observed in the resonances of residues Y25, K30, Q86 V29 and G32 (white) in its immediate vicinity as well as in residues within the flexible $\beta 5$ - $\beta 6$ loop (T71, A73, A75 and Q76) and in the unstructured C-terminal region (G106 and K112). Some residues in these latter two regions are not observed in the crystal structure and are represented by the red surfaces for residues 71, 73, 80 and 106. **D)** The W69A substitution affects cross-peaks of nearby residues F35, T70, T71, A80 and Q86 (white). W69 and I68 are located within a β -bulge. **E)** Perturbations caused by the R97A mutation in residues V29, K30, G32, D89, E94, A95, N98, I100, E101 could be explained by the three rotamers observed for the R97 side-chain in the crystal structure. The two rotamers in chain A are in cyan while the rotamers in chain B are in dark blue. One rotamer, observed in both monomers of the asymmetric unit, H-bonds with the E101 side chain and is also within H-bond distance of the main-chain oxygen atoms of Y25 and P27. In monomer A, the other R97 rotamer H-bonds with the N98 and E101 side chains. In monomer B, the second rotamer swings around to make new H-bonds with the E94 side chain and the main-chain oxygens of G31, P88, D89 and G90. The D89 side-chain, which may adopt two conformations, can H-bond with the K30 side-chain, which is in a loop with V29 and G31. **F)** Deletion of residues 107-117, lacking density in the crystal, perturbs the amide cross peaks of residues 105 (white) and 106 (blue) in $\alpha 2$ as well as residues in the $\beta 1$ - $\alpha 1$ loop and within the N-terminal half of $\alpha 1$ (residues 16-18, 20-22). Figures were produced with Pymol⁶.

Supplemental Figure 7. Two-hybrid assays show that $\text{PilZ}_{\text{XAC1133}}$ interacts with the FimX_{EAL} domain and with PilB . Representative two-hybrid assays in which PJ694-a yeast cells (*MATa trp1-901 leu2-3, 112 ura3-52 his3-200 gal4Δ gal80Δ LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ*)⁹ transformed with $\text{PilZ}_{\text{XAC1133}}$, $\text{PilZ}_{\text{Y22A}}$, or $\text{PilZ}_{\Delta 107-117}$ preys and/or the **(A)** $\text{FimX}_{\text{GGDEF-EAL}}$ (G6), **(B)** $\text{FimX}_{\text{GGDEF}}$ (GG), **(C)** FimX_{EAL} (EAL) or **(D)** $\text{PilB}_{\text{XAC3239}}$ baits (PilB). These cells were grown on semisolid YAPD medium (left), SC medium lacking tryptophan and leucine (SC-WL, center) or SC medium lacking tryptophan, leucine, histidine and adenine (SC-WLHA, right).

Supplemental Fig. 8. The NMR signal of the H8 (A) and H1' (B) nuclei of the guanine group from c-diGMP show a large attenuation and a slight downfield shift upon the addition of full-length FimX_{XAC2398}, indicating binding to the protein. The ¹H 1D spectra of c-diGMP (red), c-diGMP:FimX_{XAC2398} (1:1) (blue), and FimX_{XAC2398} (green) were acquired at 500 MHz, with water pre-saturation during the recycle delay of 2s, 8192 complex points, and in 90% D₂O. Conditions were 5 mM Tris, pH 7.0, 25 mM NaCl, 100 μM c-diGMP and/or 100 μM FimX_{XAC2398}. Analysis and processing of the data were made using TopSpin1.3 (Bruker BioSpin).

Overlay PilZ _{XAC1133}	+	
Immobilized	+	+
PilB _{XAC3239-His}		
Antibody	Anti PilZ	Anti His

Supplemental Figure 9: Control Western and Far-Western assays to demonstrate interactions between PilB_{XAC3239} and PilZ_{XAC1133}. Total cell lysates of *E. coli* cells expressing PilB fused with a polyhistidine tag, PilB_{XAC3239-His} (pilB cloned in pET28a), were separated by SDS-PAGE, transferred to nitrocellulose membranes. Left panel: The membrane was incubated with PilZ_{XAC1133}. Bound PilZ was detected with PilZ_{XAC1133} antiserum and protein-A peroxidase. Right panel:

PilB_{XAC3239-His} was detected using anti-serum against its polyhistidine tag (His) and protein-A peroxidase.

Supplemental Figure 10. PilZ_{XAC1133} is not required for development of citrus canker symptoms. Baia sweet orange (*Citrus sinensis* (L.) Osbeck) were grown under greenhouse conditions and then maintained at 28°C with daylight. Plants were inoculated with *Xac* cultures (OD₆₀₀ = 0.3 in 2xTy medium) by infiltration into the abaxial surface of attached leaves by pressure using a needleless syringe. Photos were obtained 10 days after inoculation. **A)** *Xac* strains: wild type (wt), pilZ_{XAC1133} (ΔpilZ), pilZ_{XAC1133}+pUFR047_pilZ (ΔpilZ + pilZ) and **B)** *Xac* strains: wild type (wt), fimX_{XAC2398} (ΔfimX), fimX_{XAC2398}+pUFR047_fimX (ΔfimX + fimX).

Supplemental Figure 11. Motility of *Xac* and the FimX_{XAC2398} mutant on semisolid surfaces after 2 or 4 days of growth. The composition of SB (A), STON (B), XVM2 (C) and STON (D) of the media are described in the Methods section and the amount of agar is indicated in parenthesis. Wild-type (wt), pilZ_{XAC1133} (ΔpilZ), pilZ_{XAC1133}+pUFR047_pilZ (ΔpilZ + pilZ), fimX_{XAC2398} (ΔfimX), fimX_{XAC2398}+pUFR047_fimX (ΔfimX + fimX) *Xac* strains were used.

Supplemental Table

Supplemental Table 1: Oligonucleotide sequences used in this study

A) Oligonucleotide sequences used to amplify a gene or gene fragment for cloning in expression vectors pETa or pET28a and in the two-hybrid assay vector pOBD. The restriction endonucleases used to digest the specific vectors and the restriction endonucleases used to digest the amplified fragments are shown in parenthesis. *E. coli* strains and temperatures used for protein expression are indicated in the last column.

XAC gene number	Protein or Domain	Primer Sequence	pET28a	pET3a	pOBD	<i>E. coli</i> strain BL21(DE3)
2398	FimX	5' CTAGGGAGCCACCATGGAAGGCAAGATC 3'	NcoI/BamHI (NcoI/BamHI)	----- -----	----- -----	STAR 28°C
		5' GCGGATCCCTCGAGCTTTAGCCGAAGTCCG 3'				
2398	FimX _{GGDEF-EAL} (residues 267-689)	5' TAT ACCATGGTCCATATGTTGCGCCAGCGCGAC 3'	NdeI/BamHI (NdeI/BamHI)	----- -----	NcoI/Sal (NcoI/XhoI)	RP 32°C
		5' GCGGATCCCTCGAGCTTTAGCCGAAGTCCG 3'				
2398	FimX _{GGDEF} (residues 267-432)	5' TATACCATGGTCCATATGTTGCGCCAGCGCGAC 3'	----- -----	NdeI/BamHI (NdeI/BamHI)	NcoI/Sal (NcoI/XhoI)	RIL 37°C
		5' GCGGATCCCTCGAGTCAAGCGGATCGTAGATGC 3'				
2398	FimX _{EAL} (residues 436-689)	5' GATCCCATGGTGCATATGCGGGCCGAGGAAGAAGC 3'	----- -----	NdeI/BamHI (NdeI/BamHI)	NcoI/Sal (NcoI/XhoI)	RIL 37°C
		5' GCGGATCCCTCGAGCTTTAGCCGAAGTCCG 3'				
3239	PilB	5' TTTTCCATGGTTCATATGAATAGTGTCTGACCGCAAAATTGG 3'	NdeI/HindIII (NdeI/HindIII)	NdeI/HindIII (NdeI/HindIII)	NcoI/Sal (NcoI/XhoI)	STAR 37°C
		5' TTTCTCGAGAAGCTTTAGTCTTCGTACACGATTGATCTCGG 3'				
1133	PilZ	5' TTC TCC ATG GACCATATG AGT GCA ATGAAT 3'	----- -----	NdeI/BamHI* (NdeI/BamHI)*	NcoI/Sal (NcoI/XhoI)	DE3* 37°C*
		5'CG GGATCCCTCGAGTTACATCGTATGCGTCG 3'				

* Cloning described in reference¹⁰.

B) Oligonucleotide sequences used to construct $\Delta pilZ_{XAC1133}$, $\Delta fimX_{XAC2398}$, $\Delta pilB_{XAC3239}$ and $\Delta fliC_{XAC1975}$ knockouts. For the $\Delta pilZ_{XAC1133}$, $\Delta fimX_{XAC2398}$, and $\Delta fliC_{XAC1975}$ knockouts, two 1 kb fragments corresponding to the 5' (fragment 1) and 3' (fragment 2) regions of each gene were amplified using the indicated oligonucleotides. These fragments were joined by ligation and then inserted into the pNPTS138 suicide vector. The restriction endonucleases used to digest the vector and the amplified fragments (parenthesis) are indicated in the last column. For the $\Delta pilB_{XAC3239}$ knockout the amplified fragment was initially cloned into the TOPO TA vector and then the EcoRI fragment was cloned into pNPTS138.

Knockout		Primer Sequence	pNPTS138
$\Delta pilZ_{XAC1133}$	Fragment 1	5' CACGATATCGACGAGCCTATGCCATGACCGC 3'	EcoRV/BamHI (EcoRV/BglII)
		5' GAAGATCTTGCACTCATCGAGTTCCCTAGAACGTGGC 3'	
	Fragment 2	5' GAAGATCTTCGGACAAGCCGACGCATACGATGTAAGC 3'	EcoRV/BglII (BglII/BamHI)
		5' TTGGATCCGCCCCGTGCTCAGCTTGCCG 3'	
$\Delta fimX_{XAC2398}$	Fragment 1	5' GAGGAAGCTTCGAGGTCCAGACCAGCCAGGG 3'	HindIII (HindIII/NdeI)
		5' CGT CGA CGA TCA TAT GCA GGC GGA GGG TGA GAT CTT TG 3'	
	Fragment 2	5' CATTCTGACATATGTCCGAATTCTGTCGCCGATGCGC 3'	HindIII (HindIII/NdeI)
		5' CGC CAT GTT CAA CCG CAT TTT CCT AAG CTT CCT GC 3'	
$\Delta fliC_{XAC1975}$	Fragment 1	5' ACCGAAGCTTGCTTCTGGATTTCAGCCGCGATG 3'	HindIII (HindIII/NcoI)
		5' GTT GAT TAC CTG TGC CAT GGT TTG ATT TCT CCT AAG 3'	
	Fragment 2	5' GGTTCCATGGAGAAGCTGTCTGAGCCTGCTGCAG 3'	HindIII (HindIII/NcoI)
		5' GAG GTC GCT GAC CGT GTT GGA AGC TTG GAC GTG 3'	
$\Delta pilB_{XAC3239}$	-	5' GGTCTGTTCTGATAGGGGTGAGCAATCCG 3'	EcoRI
		5'GACGAGGTGATGTTGTAGGGCGCGATGCG 3'	

C) Oligonucleotide sequences used to construct the plasmids to complement the *pilZ*_{XAC1133} and *fimX*_{XAC2398} knockouts. The restriction endonucleases used to digest the pUFR047 vector¹¹ or pMR20 vector¹² are shown in the last column. Restriction endonucleases used to digest the amplified fragment are shown in parenthesis.

Complement	Construct Name	Primer Sequences	Restriction enzymes
PilZ _{XAC1133}	pUFR047_ <i>pilZ</i>	5' CACGGATCCGACGGAGCCTATGCCATGACCGC 3' 5' CG GGATCCCTCGAGTTACATCGTATGCGTCG 3'	BamHI/ SalI (BamHI/ XhoI)
PilZ _{XAC1133} and PilZ _{Δ107-117}	pMR20_ <i>pilZ</i> and pMR20_ <i>pilZ</i> _{Δ107-117}	5' GAAGGATCCATACATATGAGTGCAATGAATG 3' 5' CGGAATTCCTTACATCGTATGCGTCG 3'	BamHI/EcoRI (BamHI/EcoRI)
FimX _{XAC2398}	pUFR047_ <i>fimX</i>	5' GAGGAAGCTTCGAGGTCCAGACCAGCCAGGG 3' 5' GCGGATCCCTCGAGCTTTCAGCCGAACCTCG 3'	HindIII/SalI (HindIII/XhoI)

D) Primers used to produce site-directed mutants within the *pilZ*_{XAC1133} gene cloned in the pET-3a vector (for heterologous protein expression) and the pOAD vector (for yeast two-hybrid assays).

PilZ Mutants	Primer Sequences
PilZ _{Y22A}	5' TGAAAGACAAGCCAGCGCTCGCCAGCGCCTACATGCCGTTTCGT 3' 5' ACGAACGGCATGTAGGCGCTGGCGAGCGCTGGCTTGTCTTTCA 3'
PilZ _{Y25A}	5' AGCCAGCGCTCTACAGCGCCGCCATGCCGTTCTGTAAGGCGGG 3' 5' CCGCCTTTACGAACGGCATGGCGGCGCTGTAGAGCGCTGGCT 3'
PilZ _{F28A}	5' TCTACAGCGCCTACATGCCGCGCCGTGAAAGGCGGCGGCATCTT 3' 5' AAGATGCCGCGCCTTTACGCGCCGCGATGTAGGCGCTGTAGA 3'
PilZ _{W69A}	5' CGGTGCGCGGCAAGGTCATCGCGACACGCCTGCGCGGCGCAC 3' 5' TGTGCGCGGCGAGCGGTGGTCCGCGATGACCTTGCCGCGCAGCG 3'
PilZ _{R97A}	5' GCCCGGAAGGCGAGGCGGTGGCCAAAGATCGAGACGCTGCT 3' 5' AGCAGCGTCTCGATCTTGTGGCCACCGCCTCGCCTTCCGGGC 3'
PilZ _{E101A}	5' AGGCGGTGCGCAACAAGATCGCGACGCTGCTGGCCGGATTGAC 3' 5' GTCAATCCGGCCAGCAGCGTCGCGATCTTGTTCGCGACCGCCT 3'
PilZ _{L104A}	5' GCAACAAGATCGAGACGCTGGCGGCCGATTGACCACTCGGA 3' 5' TCCGAGGTGGTCAATCCGGCGCCAGCGTCTCGATCTTGTTCG 3'
PilZ _{Δ107-117}	5' TCGAGACGCTGCTGGCCGGATGAACCACTCGGACAAGCCGAC 3' 5' GTCGGCTTGTCCGAGGTGGTTCATCCGGCCAGCAGCGTCTCGA 3'

References

1. Finn, R. D., Tate, J., Mistry, J., Coghill, P. C., Sammut, S. J., Hotz, H. R., Ceric, G., Forslund, K., Eddy, S. R., Sonnhammer, E. L. & Bateman, A. (2008). The Pfam protein families database. *Nucleic Acids Res* 36, D281-8.
2. Altschul, S. F. & Lipman, D. J. (1990). Protein database searches for multiple alignments. *Proc Natl Acad Sci U S A* 87, 5509-13.
3. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947-8.
4. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25, 3389-402.
5. Holm, L., Kaariainen, S., Rosenstrom, P. & Schenkel, A. (2008). Searching protein structure databases with DALI-Lite v.3. *Bioinformatics* 24, 2780-1.
6. Laskowski, R. A., Watson, J. D. & Thornton, J. M. (2005). ProFunc: a server for predicting protein function from 3D structure. *Nucleic Acids Res* 33, W89-93.
7. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6, 277-93.
8. Delano, W. L. (2002). The PyMOL molecular graphics system <http://www.pymol.org>.
9. James, P., Halladay, J. & Craig, E. A. (1996). Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* 144, 1425-36.
10. Guzzo, C. R. & Farah, C. S. (2009). Expression, crystallization and preliminary crystallographic analysis of PilZ(XAC1133) from *Xanthomonas axonopodis* pv. *citri*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 65, 304-6.
11. De Feyter, R., Yang, Y. & Gabriel, D. W. (1993). Gene-for-genes interactions between cotton R genes and *Xanthomonas campestris* pv. *malvacearum* avr genes. *Mol Plant Microbe Interact* 6, 225-37.
12. Roberts, R. C., Mohr, C. D. & Shapiro, L. (1996). Developmental programs in bacteria. *Curr Top Dev Biol* 34, 207-57.

3.3.3 Ensaios de motilidade de diferentes cepas de *Xac*

Células nocautes de *pilZ*, *fimX* e *pilB* aumentam a motilidade de *Xac* em superfícies semi-sólidas quando comparadas com células tipo selvagem: Ensaios de motilidade das células nocautes dos genes de *pilZ*_{XAC1133} e *fimX*_{XAC2398} causaram o aumento de motilidade de *Xac* em meio semi-sólido (**Figura 12**). Ao adicionar 300mM de NaCl nestes meios semi-sólidos a motilidade reduz e não se observa mais diferença de motilidade entre as células tipo selvagem com as células nocautes (**Figura 12 a**). Por outro lado, ao adicionar triptona ao meio, as células de todas as cepas se dispersam mais rapidamente, mas ainda se nota o aumento de motilidade das células nocautes de *Xac* (**Figura 12 a**).

O movimento sendo observado não requer de flagelo, já que células nocautes de *fliC*, gene que codifica a subunidade majoritária do flagelo, não apresenta nenhuma mudança de motilidade em meio STON quando comparadas com as células tipo selvagem (**Figura 12 c**). Desta forma o movimento não se trata de *swarming*, que é dependente de flagelo^{141,142}. Infelizmente não conseguimos encontrar uma condição em que *Xac* apresentasse motilidade *twitching*, o que é dependente de T4P em *P. aeruginosa*^{143,144}. Este tipo de movimento requer que as células andem sobre uma superfície sólida que pode ser no fundo da placa de petri, sobre o plástico^{3,93}. O ensaio que realizamos aparentemente não se trata deste movimento, pois as células migram sobre o meio semi-sólido.

Em algumas condições, células nocautes de *pilA* de *P. aeruginosa* apresentam um aumento de motilidade em meio semi-sólido¹⁴⁵. Recentemente foi mostrado que *P. aeruginosa* em meio semi-sólido exibe outro tipo de motilidade, chamada *sliding*¹⁴⁶. O movimento *sliding* é um transporte bacteriano passivo que envolve a secreção de substâncias que diminuem a tensão superficial e a fricção entre as células, como lipopeptídeos, liposacarídeos entre outros^{141,147}. Este tipo de movimento é inibido por estruturas na superfície da bactéria, como flagelo e T4P que promovem adesão entre as células e entre as células e a superfície¹⁴⁵. Desta forma, a motilidade sendo observada pelas células nocautes de *pilZ*_{XAC1133}, *fimX*_{XAC2398} e *pilB*_{XAC3239} é condizente com este tipo de motilidade. Provavelmente estas células nocautes não contem o T4P na superfície ou o T4P não está funcional.

O fenótipo de motilidade foi revertido quando células mutantes de *pilZ*_{XAC1133} foram complementadas com o gene de *pilZ*_{XAC1133} tipo selvagem (**Figura 12 a-c**).

Interessantemente, a complementação das células nocautes de *pilZ*_{XAC1133} com *pilZ*_{Δ107-117} não revertem o fenótipo. Esta observação sugere que a interação de PilZ_{XAC1133} com PilB_{XAC3239} é essencial para que o fenótipo seja revertido⁶ porque os últimos aminoácidos de *PilZ*_{XAC1133} são importantes para a interação com PilB_{XAC3239}. A complementação das células nocautes de *pilZ*_{XAC1133} com *pilZ*_{W69A} também não reverteu o fenótipo, mas a complementação com *pilZ*_{E101A} reverteu. Interessantemente, a complementação das células nocautes de *pilZ*_{XAC1133} com *pilZ*_{Y25A} ocasionou uma diminuição na motilidade das células quando comparadas com as células tipo selvagem (**Figura 12 b**).

Finalmente, observamos que células nocautes de *pilZ*_{XAC1133} crescidas em meio semi-sólido STON formam grumos, populações de células separadas, quando observadas pelo microscópio, enquanto que as células tipo selvagem e complementadas com *pilZ*_{XAC1133} formam populações de células mais homogêneas (**Figura 12 d**).

a

SB 0.6% Agar, 2 dias

SB 0.6% Agar + 300mM NaCl, 4 dias

Eiken 0.3% Agar, 2 dias

Eiken 0.3% Agar + Triptona,
2 dias Eiken 0.3% Agar + 300mM NaCl, 4
dias

Ston 0.3% Agar, 2 dias

Ston 0.3% Agar + Triptona, 2
dias Ston 0.3% Agar + 300mM NaCl, 4
dias

b

c **d**

Figura 12: Motilidade de *Xac* em meio semi-sólido.

a) Os meios e a concentração de agar assim como os dias de crescimento das células estão indicados abaixo de cada figura. As cepas de *Xac* utilizadas para os ensaios são: tipo selvagem (wt), *pilZ*_{XAC1133} ($\Delta pilZ$), *pilZ*_{XAC1133}+pUFR047_*pilZ* ($\Delta pilZ$ + *pilZ*) e *fimX*_{XAC2398} ($\Delta fimX$). **b)** As cepas de *Xac* utilizadas para os ensaios de motilidade em meio semi-sólido SB com 0.6% agar são: tipo selvagem (wt), *pilZ*_{XAC1133} ($\Delta pilZ$), *pilZ*_{XAC1133}+pMR20_*pilZ* ($\Delta pilZ$ + *pilZ*), *pilZ*_{XAC1133}+pMR20_*pilZ*_{Y25A} ($\Delta pilZ$ + *pilZ*_{Y25A}), *pilZ*_{XAC1133}+pMR20_*pilZ*_{E101A} ($\Delta pilZ$ + *pilZ*_{E101A}), *pilZ*_{XAC1133}+pMR20_*pilZ*_{W69A} ($\Delta pilZ$ + *pilZ*_{W69A}), *pilZ*_{XAC1133}+pMR20_*pilZ* _{Δ 107-117} ($\Delta pilZ$ + *pilZ* _{Δ 107-117}) e *fimX*_{XAC2398} ($\Delta fimX$). O vetor pMR20 expressa o gene de interesse de forma constitutiva. **(c)** Ensaio de motilidade de diferentes cepas de *Xac* em meio semi-sólido STON, mostrando que células nocautes de *fliC* (produzida pelo aluno do laboratório Mestre Maxuel de Oliveira Andrade) não sofrem mudanças de motilidade quando comparadas com as células tipo selvagem. **(d)** imagens por microscópio de diferentes cepas de *Xac* crescidas em meio semi-sólido. As setas indicam uma amplificação da imagem para destacar o comportamento das

células na colônia, células nocautes de *pilZ* formam grupos enquanto que as células tipo selvagem e células nocautes de *pilZ* complementadas com *pilZ* formam colônias homogêneas.

3.3.4 Microscopia eletrônica das células nocautes de *PilZ*_{XAC1133} e *FimX*_{XAC2398}

A análise de estruturas filamentosas na superfície de *Xac* quando crescidas em meio semi-sólido foram analisadas por microscopia de transmissão. Observamos diferentes estruturas filamentosas na superfície de *Xac* que varia conforme a composição do meio, como está descrito a seguir.

Células de *Xac* tipo selvagem assim como células nocautes de *fimX*_{XAC2398} crescidas em meio semi-sólido SB (0.6% agar) formam vários filamentos finos e longos na superfície da bactéria (**Figura 13 a**), enquanto que as células nocautes de *pilZ*_{XAC1133} não formam estes filamentos na superfície (**Figura 13 a**). Células tipo selvagem e do nocaute de *fimX*_{XAC2398} crescidas em meio semi-sólido XVM2 (0.6% agar) formam filamentos muito longos que conectam a lateral de uma bactéria ao pólo de outra bactéria (**Figura 13 b**). Estes filamentos formam bolhas como se fossem formadas por vários filamentos que quando são torcidos conjuntamente formam superhélices (**Figura 13 b**).

As células nocautes de *pilZ*_{XAC1133} e as células complementadas com *pilZ*_{XAC1133} crescidas em meio semi-sólido XVM2 (0.6% agar) não possuem estes filamentos na superfície, mas as células complementadas com *pilZ*_{XAC1133} possuem maior quantidade de flagelo (**Figura 13 b**) o que explicaria sua reversão parcial em meio semi-sólido XVM2 (0.6% agar) quando comparada com as células tipo selvagem (**Figura suplementar 11 c do Anexo 2**).

Células de *Xac* tipo selvagem assim como do nocaute de *fimX*_{XAC2398} crescidas em meio semi-sólido STON (0.3% agar) formam filamentos na superfície da bactéria parecidos com aqueles observados em meio semi-sólido XVM2 (**Figura 13 c**). Células nocautes de *pilZ*_{XAC1133} possui menor quantidade destes filamentos na superfície (**Figura 13 c**) e sua complementação com *pilZ*_{XAC1133} possui o mesmo fenótipo que as células tipo selvagem.

Células de *Xac* tipo selvagem crescidas em meio semi-sólido Eiken (0.6% agar) formam filamentos finos e longos na superfície da bactéria parecidos com aqueles observados em meio SB (**Figura 13 d**). Células nocautes de *pilZ*_{XAC1133} possuem menor quantidade destes filamentos na superfície (**Figura 13 d**) e sua

complementação com *pilZ*_{XAC1133} possui o mesmo fenótipo que as células tipo selvagem.

As observações relacionadas a **Figura 13** estão resumidas na **Tabela 12**. Desta forma, *Xac* possui diferentes estruturas de filamentos na superfície da bactéria de forma dependente da composição do meio. Em meio rico, como no caso de SB e Eiken, *Xac* forma filamentos finos e longos, enquanto que em meios pobres de nutrientes, como no caso de XVM2 e STON, *Xac* forma filamentos longos que formam bolhas que muitas vezes conectam a lateral de uma bactéria ao pólo de outra bactéria. No entanto, não se observa uma correlação clara de ausência e presença destes filamentos nas células nocautes, o que sugere que possa haver outros motivos para o aumento de motilidade das células nocautes nestes meios.

Tabela 12: Resumo dos resultados mostrados na Figura 13.

+++ : alta quantidade e + : baixa quantidade. Nid: não identificado.

Cepas	Meio	Estruturas observadas			Figura
		Flagelo	Filamentos Finos e longos	Bolhas	
wt	SB	+	+++	nid	12a
	XVM2	+	nid	+++	12b
	STON	+	nid	+++	12c
	Eiken	+	+++	nid	12d
$\Delta pilZ$	SB	+	+	nid	12a
	XVM2	+	nid	+	12b
	STON	+	nid	+	12c
	Eiken	+	+	nid	12d
$\Delta pilZ + pilZ$	SB	+	+++	nid	12a
	XVM2	+++	nid	nid	12b
	STON	+	nid	+++	12c
	Eiken	+	++	nid	12d
$\Delta fimX$	SB	+	+++	nid	12a
	XVM2	+	nid	+++	12b
	STON	+	nid	+++	12c

Legenda

- f: Filamento fino
 .....▶ F: Flagelo
 - - - - -▶ B: Filamento que forma bolhas

a

SB, 0,6% agar

b

XVM2, 0,6% agar

Figura 13: Microscopia eletrônica de *Xac* em diferentes meios.

Microscopia eletrônica de diferentes cepas de *Xac* crescidas em meio semi-sólido SB (a), XVM2 (b), STON (c) e Eiken (d). As cepas de *Xac* utilizadas para os ensaios são: tipo selvagem (wt), *pilZ*_{XAC1133} ($\Delta pilZ$), *pilZ*_{XAC1133}+pUFR047_ *pilZ* ($\Delta pilZ + pilZ$) e *fimX*_{XAC2398} ($\Delta fimX$). As setas em vermelho indicam os filamentos finos e longos (f) enquanto que as setas vermelhas e pontilhadas indicam os filamentos que formam bolhas (B) como se fossem um conjunto de filamentos muito finos juntos, as setas pontilhadas em azul indicam flagelos (F).

Células nocautes de *pilZ* e *fimX* continuam aderindo em folhas de laranja.

Ensaio de aderência de diferentes cepas de *Xac* em folhas de laranja, utilizando células crescidas em meio líquido XVM2, que mimetiza as condições de crescimento de *Xac* na planta¹⁴⁸⁻¹⁵¹, e em meio líquido STON, mostrou que todas as cepas aderem de forma equivalentes na folha de laranja (**Figura 14**). Estes resultados mostram que *fimX*_{XAC2398} e *pilZ*_{XAC1133} não afetam a aderência de *Xac* na superfície da planta nas condições testadas.

Figura 14: Ensaio de aderência em folhas de laranja.

Ensaio de aderência de células de *Xac* crescidas em meio líquido SB (**a**) e XVM2 (**b**) coradas com 1% cristal violeta. As cepas de *Xac* utilizadas para os ensaios são: tipo selvagem (wt), *pilZ*_{XAC1133} ($\Delta pilZ$), *pilZ*_{XAC1133}+pUFR047_ *pilZ* ($\Delta pilZ$ + *pilZ*) e *fimX*_{XAC2398} ($\Delta fimX$).

3.3.5 Imunolocalização de *PilA*_{XAC3240(32)_his}

Como já foi mencionado o nocaute de *pilZ* em *P. aeruginosa* acaba com a motilidade *twitching*, por não secretar PilA na superfície da bactéria³. Infelizmente não conseguimos encontrar condições em que *Xac* apresenta-se este tipo de motilidade, mas não foi descartada a possibilidade de que estes filamentos finos observados na superfície de *Xac* nos meios semi-sólidos pudessem ser T4P (**Figura 13**).

Para testar esta hipótese analisamos os dois genes que codificam para PilA em *Xac* (XAC3240 e XAC3241). O alinhamento destes genes com o PilA de *P. aeruginosa* está mostrado na **Figura 15 a**. Purificamos o fragmento de 32-142 de *PilA*_{XAC3240(32)_his} que não contém o peptídeo sinal, **Figura 15 b**, e produzimos anticorpos policlonais contra *PilA*_{XAC3240(32)_his} (dados não mostrados).

Ensaio de imunolocalização das células crescidas em meio semi-sólido nos diferentes meios testados não foram marcadas contra o soro contendo anticorpo

policlonal contra PilA_{XAC3240(32)_his} (dados não mostrados). Uma possível explicação para isso é que estes filamentos na superfície de *Xac* não sejam T4P. Outra possibilidade é que estes filamentos sejam T4P formados por uma outra pilina que não é PilA_{XAC3240}. As outras pilinas codificadas pelo genoma de *Xac* são: PilA_{XAC3241}, PilA_{XAC3805}, FimT_{XAC2669}, PilV_{XAC2668}, PilW_{XAC2667}, PilX_{XAC2666}, PilY1_{XAC2665} ou PilE_{XAC2665}. Eles apresentam grande variabilidade de sequência primária.

a

Pseudomonas XAC3241 XAC3240	--MKAQKGFTLIELMIVVAIIIGILAAIAIEQYQNYVARSEGASALATINPLKTTVEESLS --MKKQGFTLIELMIVIAIIATLAIALEQYQNYVAKSQVTAGLAELSPGKTQYETLLN MANKVIKGFTELIELMIVIAIIAVLASIAIEQYQIYVAKSQVAAGLAEVSPGKTRYEILVN * :*****:***:****:***** :*:** :* ** * :.
Pseudomonas XAC3241 XAC3240	RGAGSKIKIGTTASTATETVGVPEPDANKLGVIAVAIEDSGAGDITFTFQTGTS-SPKNA EGK-----GTLITADAVGLQATSTINRCSAITVLAPDATTGEQTNALSCTLAGSPRIV DSE-----GSTLTSASAIGLTATATS-RCSSISVLAPDAN-GAQAHAIDCTLRGNPRVN . :*: ** : . . : : *:* * : : . :*:
Pseudomonas XAC3241 XAC3240	TKVITLNRADGVWACKS----TQDPMFTPKGCDN GKKTWARDINGVWTCSTDIATADKAKFAPATCQ- GAKLTWARDANGIWTCTDLPLVDKDRFAPASCR- :* * :*:***: : . . *:* *

b KDa

Figura 15: Alinhamento múltiplo entre PilA de *P. aeruginosa* e os PilA de *Xac* (XAC3241 e XAC3240) e gel de SDS-PAGE contendo amostras puras de PilA_{XAC3240(32)}

Os resíduos vermelhos destacados na sequência de XAC3241 são idênticos ao da proteína PilA de *P. aeruginosa* mas diferentes da sequência de XAC3240 e os resíduos vermelhos destacados na sequência de XAC3240 são idênticos ao da proteína PilA de *P. aeruginosa* mas diferentes da sequência de XAC3240. O quadrado em vermelho destaca os aminoácidos pertencentes ao peptídeo sinal localizado no N-terminal de PilA. **b)** Amostras purificadas de PilA_{XAC3240(32)_his} (peso molecular 13.9 KDa) (2-3) e marcador de peso molecular (1).

3.4 Discussão

Receptores de c-diGMP. A molécula de c-diGMP é um segundo mensageiro ubíquo entre as bactérias que pode desempenhar papéis centrais em diferentes processos bacterianos como na produção de polissacarídeos extracelulares, diferenciação celular, motilidade, expressão gênica e virulência em algumas bactérias.

Receptores de c-diGMP são derivados de diferentes famílias de proteínas como domínios GGDEF ou EAL não-catalíticos, e em algumas proteínas como PelD, FleQ, Clp (CAP-like protein) e domínios PilZ e até em domínios de RNA (“riboswitches”)^{4,115,133,152}. Desta forma, a molécula de c-diGMP atua como uma molécula reguladora na atividade e na função de uma vasta gama de proteínas e domínios de RNA.

Em algumas bactérias como em *Bacillus cereus*, *V. cholerae* e *Clostridium difficile*, a interação de c-diGMP por domínios riboswitches de RNAm regulam a expressão gênica de uma enorme variedade de genes envolvidos em processos celulares fundamentais, como genes envolvidos em virulência, formação de pilus e na biossíntese de flagelo¹³³ (notar que riboswitches para c-diGMP ainda não foram identificados em espécies de *Xanthomonas*). A interação de c-diGMP pela proteínas Clp e FleQ diminuem a suas afinidades pelo DNA^{152,153}, e no caso de FleQ ocorre o aumento da expressão gênica de genes envolvidos na síntese de EPS em *P. aeruginosa*¹⁵³. PelD pertence a via de síntese EPS e sua interação com c-diGMP aumenta a produção de EPS em *P. aeruginosa*¹⁵⁴.

Proteínas quiméricas que contêm ambos os domínios EAL e GGDEF, em geral, apenas um dos domínios é ativo e o outro inativo (“degenerado”), possivelmente atuando como um regulador^{155,156}. Isso é o caso da proteína CC3396 de *C. crescentus* que contem o domínio GGDEF degenerado e sua interação com GTP regula alostericamente a atividade PDE-A do domínio EAL¹⁵⁷. Há ainda alguns casos em que ambos os domínios são ativos. Qual domínio é ativo depende das condições ambientais, como no caso da proteína MSDGC-1 de *Mycobacterium smegmatis*¹⁵⁸ e da proteína ScrC que possui atividade PDE-A na presença de ScrA e ScrB e atividade DGC-A na ausência de ScrA e ScrB¹⁵⁹. Há casos também em que os dois domínios são degenerados, ou seja, inativos. Alguns domínios GGDEF inativos podem interagir com GTP e alguns domínios EAL inativos podem interagir com c-diGMP, como ocorre com FimX de *P. aeruginosa*^{5,111,160}.

O primeiro receptor de c-diGMP descrito foi PilZ da subunidade A da celulose sintetase, em que a interação de PilZ com c-diGMP ativa a atividade da enzima¹⁶¹. Nem todos os domínios PilZ tem a capacidade de interagir com c-diGMP; por exemplo as cinco proteínas de *V. cholerae* que contem domínios PilZ (VC0697, VC1885, VC2344, VCA0042, and VCA0735) apenas duas (VC2344 e VCA0042) interagem com c-diGMP em ensaios de ligação *in vitro*¹¹⁷. *P. aeruginosa* contem oito proteínas que contem domínios PilZ, sete deles interagem com c-diGMP, mas um deles, o PilZ original, PA2960, envolvida no controle da biogênese de T4P, não interage com c-diGMP¹¹⁹. Este resultado é condizente com os nossos resultados de que PilZ_{XAC1133}, ortólogo de PA2960, também não interage com c-diGMP (**Figura suplementar 3 do Anexo 2**).

Estrutura tridimensional de PilZ_{XAC1133} mostra uma diversidade nas estruturas das proteínas da família PilZ e explica a sua incapacidade de interagir com c-diGMP. PilZ_{XAC1133} é um ortólogo da PA2960 de *P. aeruginosa*. A comparação da sua estrutura tridimensional com outras estruturas de PilZ depositadas no PDB mostrou, surpreendentemente, diferenças significativas no sítio de interação ao c-diGMP (**Figura 1 do Anexo 2**). Esta observação, mostra de forma clara que PilZ_{XAC1133} não possui o sítio de interação a c-diGMP. Esta mudança estrutural em PilZ_{XAC1133} explica a sua incapacidade de interagir com este dinucleotídeo.

Aminoácidos conservados em membros da família PA2960/XAC1133: Membros da família PA2960/XAC1133 são caracterizados por possuírem cinco motivos de seqüência conservados, MI-MV, sendo que os primeiros quatro motivos estão localizados em uma face da estrutura de PilZ_{XAC1133} (**Figura 3 e Figura suplementar 1 do Anexo 2**). O último motivo, MV, corresponde a resíduos desordenados localizados no C-terminal de PilZ_{XAC1133}. Alguns resíduos hidrofóbicos conservados estão expostos ao solvente levantando, portanto, a hipótese de que estes poderiam pertencer a uma face de interação de PilZ_{XAC1133} com outras proteínas. Mutações sítio dirigida de alguns destes resíduos mostrou que Y25, R97 e L104 diminuíram a estabilidade da proteína, mas Y22, F28, W69, E101 e a deleção dos dez últimos resíduos de PilZ_{XAC1133} (PilZ_{Δ107-117}) não afetaram sua estabilidade (**Figura suplementar 4 do Anexo 2**).

Mutação na Y22, localizada na $\alpha 1$, para uma alanina causou mudanças no deslocamento químico de resíduos no loop entre $\beta 1$ e $\alpha 1$, em resíduos localizados na

$\alpha 1$ e no C-terminal (**Figura complementar 5B do Anexo 2**). Este resultado sugere que possa haver interações entre resíduos destruturados, localizados no C-terminal de PilZ_{XAC1133}, com resíduos localizados na $\alpha 1$ (**Figuras suplementares 5 e 6B do Anexo 2**). Esta hipótese é reforçada, principalmente, pela observação de que a deleção dos dez últimos resíduos de PilZ_{XAC1133} ocasiona mudanças no deslocamento químico de resíduos localizados na $\alpha 1$ (**Figuras suplementares 5 e 6F do Anexo 2**). Além disso, a mutação na F28A, localizada na $\alpha 1$, também provoca mudanças no deslocamento químico de resíduos localizados no C-terminal (**Figura complementar 5 e 6C do Anexo 2**).

PilZ interage com o domínio EAL de FimX e com PilB, envolvidos no controle da biogênese do T4P. A interação de PilZ_{XAC1133} com FimX_{XAC2398}, proteína envolvida na regulação do T4P, e com PilB_{XAC3239}, componente ATPásico chave no processo de polimerização do T4P (vide em baixo), são os primeiros parceiros observados para um membro da família PA2960/XAC1133 (**Figuras 4-5, 6 e Figura complementar 7 do Anexo 2**). Estas interações inéditas coloca PilZ como um novo componente da maquinaria de biogênese do T4P.

PilZ_{XAC1133} interage com o domínio EAL de FimX_{XAC2398} (**Figura 4 e 5 do Anexo 2**). Esse resultado não é necessariamente surpreendente já que existe algumas proteínas que contem o domínio de PilZ, não pertencente a família PA2960/XAC1133, fusionado ao C-terminal do domínio EAL, como por exemplo, as proteínas ELI_07280 e ED21_28508 de *Erythrobacter* sp^{115,120}. Interessantemente, o domínio EAL de membros homólogos da proteína ELI_07280 e ED21_28508 possuem os resíduos envolvidos na degradação ao c-diGMP conservados (**Figura 16 e 17**) e a maioria dos domínios PilZ destas proteínas também possuem os resíduos envolvidos na interação ao c-diGMP (**Figura 16 e 17**). Esta observação sugere que, diferentemente da interação PilZ_{XAC1133}-FimX_{XAC2398}, em outros sistemas os domínios PilZ, receptores de c-diGMP, possam interagir com domínios EAL ativos.

Estudos recentes mostraram que a síntese de alginato em *P. aeruginosa* é controlada por duas proteínas associadas á membrana bacteriana, Alg44, que liga a c-diGMP via o domínio PilZ, e MucR, que contem os domínio GGDEF e EAL localizados no C-terminal¹⁶¹. Foi proposto que MucR regula a síntese de alginato pela síntese local de c-diGMP próximo ao domínio PilZ de Alg44¹⁶¹. Outra possibilidade para a regulação da síntese de alginato, com base nos nossos resultados, é que haja uma interação

direta entre o domínio PilZ de Alg44 e o domínio EAL de MucR, e esta interação pode ser modulada pela interação com c-diGMP.

Figura 16: Análise da presença de resíduos essenciais para a atividade PDE-A de domínios EAL e de resíduos envolvidos na interação de domínios PilZ com c-diGMP.

a) Análise do sítio ativo do domínio EAL com base na classificação feita por Romling¹⁶². Os aminoácidos coloridos em verde claro estão envolvidos na coordenação ao íon divalente (magnésio ou cálcio) e os coloridos em verde escuro na ligação indireta ao íon divalente, os aminoácidos coloridos em azul estão envolvidos na ligação a molécula de c-diGMP, em vermelho corresponde a base geral envolvida na ativação da molécula de água nucleofílica, os aminoácidos coloridos em lilás correspondem ao loop 6 em que E₂₇₅ é importante para a estabilização deste loop. Detalhes do mecanismo de ação dos domínios EAL está descrito a seguir e na **Figura 19**. **b)** Alinhamento das seqüências dos domínios PilZ das proteína ED21_28508_{PilZ(576-658)} e ELI_07280_{PilZ(806-895)} com o domínio PilZ da proteína VCA0042 de *Vibrio cholerae*¹¹⁶. Em azul estão destacados os resíduos conservados na super-família de PilZ¹²⁰ em que os 5 primeiros resíduos conservados, RxxxR, D/NxSxxG são os envolvidos na interação ao segundo mensageiro bacteriano c-di-GMP^{115,116}. Mutações nestes resíduos acabam com a capacidade destes domínios de PilZ ligarem ao c-diGMP^{4,117,119,163}. Vale a pena ressaltar que apenas os primeiros 100 resíduos de cada domínio PilZ possui similaridade em seqüência primária com a seqüência consenso de PilZ¹²⁰. Como pode ser visto na Figura, apenas a região N-terminal destes domínios é conservada, justamente na porção onde os domínios PilZ estão envolvidos na interação ao c-diGMP. Esta análise sugere que os domínios EAL de ED21_28508 e ELI_07280 tenham atividade PDE-A e seus domínios PilZ liguem ao c-diGMP.

a

ED21_28508

ED21_28508_{EAL-PilZ(320-710)}

b

ELI_07280

138

ELI_07280_{EAL-PilZ(551-895)}

Figura 17: Alinhamento múltiplo de membros ortólogos das proteínas ED21_28508 e ELI_07280.

A parte de cima de cada diagrama do WebLogo ilustra a organização de domínios de cada proteína. ED21_28508 é composta por quatro domínios, PAS-GGDEF-EAL-PilZ e ELI_07280 é composta por três domínios, GGDEF-EAL-PilZ¹²⁰. As caixas em vermelho mostram os resíduos envolvidos na atividade PDE-A dos domínios EALs e as caixas em violeta destacam os resíduos envolvidos na capacidade de domínios PilZ ligarem ao c-diGMP. Foram utilizados 15 seqüências de proteínas ortólogas a ED21_28508_{EAL-PilZ(320-710)} (**a**) e 11 ortólogos da proteína ELI_07280_{EAL-PilZ(551-895)} (**b**). Esta análise mostra que os resíduos importantes para a atividade do domínio EAL e os resíduos importantes para a interação ao c-diGMP pelo domínio PilZ são conservados. Há outros resíduos conservados nestas famílias que devem desempenhar outras funções ainda desconhecidas, principalmente no domínio EAL e na região N-terminal do domínio PilZ. Estas seqüências foram obtidas no site do Pfam e posteriormente utilizadas para o desenho gráfico pelo programa WebLogo^{164, 165, 166}.

O C-terminal conservado e desordenado de PilZ contribui para sua interação com PilB. A Y22 possivelmente tem um papel regulatório na capacidade de PilZ interagir com FimX e PilB. Ensaios de mutação sítio dirigida de resíduos conservados e expostos na estrutura de PilZ_{XAC1133} mostrou que os dez últimos resíduos de PilZ_{XAC1133}, que contem o motivo MV, afetou a interação de PilZ_{XAC1133} com PilB_{XAC3239} mas não com FimX_{XAC2398} (**Figura 7 do Anexo 4**). Estes resultados foram confirmados por ensaios de 2-híbridos em levedura (**Figura 4 e Figura suplementar 7 do Anexo 4**). Estes resultados sugerem que o C-terminal de PilZ_{XAC1133}, conservado mas desordenado, faz parte da interface de interação com PilB_{XAC3239}. Esta interação aparentemente tem um papel fisiológico porque o mutante *pilZ*_{Δ107-117} não foi capaz de reverter o fenótipo do nocaute *ΔpilZ* (**Figura 11b**).

O mutante PilZ_{Y22A} apresentou um reduzida interação com PilB_{XAC3239} e com FimX_{XAC2398}. Como foi descrito acima (Resultados), existe a possibilidade da hélice $\alpha 1$, que contem o resíduo Y22, fazer contatos com resíduos localizados no C-terminal de PilZ_{XAC1133}, que faz parte da interface de interação com PilB_{XAC3239}. Desta forma, é possível que a Y22 tenha um papel regulatório na interface de interação com PilB_{XAC3239}, e na interface de interação com FimX_{XAC2398}, que todavia é desconhecida.

Análise funcional dos domínios EAL, GGDEF e HD-GYP de Xac. A interação PilZ-FimX é muito interessante pelo fato de envolver 3 domínios envolvidos em sinalização por c-diGMP, apesar de todos eles serem inativos. Antes de analisar as razões pelas quais os domínios GGDEF e EAL de FimX_{XAC2398} são inativos, vamos analisar todas as proteínas com domínios GGDEF e/ou EAL ou HD-GYP encontradas no genoma de Xac. Xac possui 34 proteínas com domínios GGDEF e/ou EAL e mais 03 proteínas com domínios HD-GYP⁷. Assim, Xac apresenta um bom exemplo da

diversidade de domínios que metabolizam c-diGMP. O que segue é uma breve análise das seqüências primárias destes domínios com o intuito de prever se apresentarão ou não atividade enzimática.

Análise funcional dos domínios GGDEF de Xac. Os domínios GGDEF e EAL podem ser encontrados sozinhos ou em proteínas com multi-domínios que contêm domínios regulatórios, tais como: RR (Regulador de Resposta), PAS (domínio sinalizador de ligação a cofatores como haem, Flavinas entre outros), GAF (domínio sensor que liga cofatores como o grupo heme)¹⁶⁷ entre outros^{127,130,168}.

Estudos de algumas proteínas contendo domínio GGDEF mostraram que o domínio GGDEF sozinho, tem baixa atividade DGC-A, enquanto que na presença de seus domínios regulatórios, a sua atividade é aumentada significativamente¹⁶⁹. Vários estudos recentemente publicados mostraram que estes outros domínios são responsáveis pela dimerização dos domínios GGDEF, em que cada domínio de GGDEF liga a uma molécula de GTP, para que ocorra a catálise de conversão de duas moléculas de GTP a uma molécula de c-diGMP^{126,131,157,170-172}.

A análise funcional dos domínios GGDEF das 30 proteínas encontradas em *Xac* está esquematizada na **Figura 18** em que foi utilizado como base a estrutura de PleD de *C. crescentus* (PDBID: 2VON)^{126,173} e a análise feita por Navarro *et al.* para FimX de *P. aeruginosa*¹⁶⁰. O sítio ativo do domínio GGDEF de PleD é composto por 7 aminoácidos em que D₃₂₇, E₃₇₀ (do motivo consenso GGDEF) estão envolvidos na coordenação a dois íons de magnésio (colorido em verde na **Figura 18**) e os demais, F₃₃₁, N₃₃₅, D₃₄₄, K₄₄₂ e R₄₄₆ estão envolvidos na ligação ao GTP (colorido em azul na **Figura 18**), em que K₄₄₂ e R₄₄₆ interagem com o fosfato gama do GTP e F₃₃₁ e D₃₄₄ interagem com a base do nucleotídeo (colorido em vermelho na **Figura 18 b**).

Esta análise possibilitou a classificação das proteínas de *Xac* em três classes: *Classe 1*- proteínas ativas, ou seja possuem todos os resíduos importantes para a catálise conservados, totalizam 16 proteínas com domínios GGDEF funcionais; *Classe 2* – podem ser ativas, pois não possuem alguns resíduos importantes para catálise conservados, como a ausência de resíduos importantes no sítio alósterico e/ ou no sítio ativo, totalizando 8 proteínas e *Classe 3* - são inativas, pois não possuem vários resíduos importantes para a catálise, totalizam 6 proteínas (**Figura 18**).

a

Sítio Ativo

Sítio Alostérico - I

b

Ativos, proteínas que possuem todos os resíduos importantes para a catálise

	327	DIDFF	335	N	344	DEV	sítio I	GGDEF	370	EEF	442	KASGR
PleD-C.crescentus-GGDEF-290-454		I	D	I	D	F	V	R	A	I	D	I
XAC1420-330-490		I	D	L	D	D	V	R	T	S	D	K
XAC1345-400-570		L	D	V	D	H	F	K	Q	F	N	D
XAC2897-170-340		I	D	V	D	H	F	K	L	V	N	D
XAC1486-160-320		L	D	L	D	H	F	K	E	V	N	D
XAC1570-860-1011		L	D	I	D	H	F	K	R	I	N	D
XAC0424-380-540		L	D	L	D	D	F	R	R	L	N	E
xac2812-360-520		V	D	L	D	H	F	K	L	V	N	D
XAC0614-140-326		V	D	I	D	H	F	K	R	V	N	D
rrpX-XAC1345-250-420		V	D	I	D	H	F	K	H	I	N	D
XAC4358-430-570		V	D	L	D	K	F	K	L	V	N	D
XAC1940-400-570		I	D	L	D	H	F	K	L	V	N	D
XAC1488-400-570		L	D	V	D	H	F	K	R	F	N	D
pdeA-XAC1887_301-490		I	N	I	D	H	F	K	L	I	N	D
XAC1939-783-939		V	D	I	D	H	F	K	H	I	N	D
XAC2482-254-411		I	D	L	D	H	F	K	Q	V	N	D
XAC1488-400-566		L	D	V	D	H	F	K	R	F	N	D

Possivelmente ativos, proteínas com ausência de alguns resíduos no sítio I e/ou sem resíduos importantes para a interação ao GTP

XAC2382_279-440	A	D	I	D	D	H	F	K	R	V	N	D	T	L	G	H	E	A	G	D	E	A	A	L	G	G	G	G	D	E	F	K	V	A	G	K
XAC2810-200-360	L	D	V	D	H	F	K	T	V	N	D	V	H	G	H	D	Q	G	D	A	V	M	A	G	A	G	G	G	E	E	F	K	R	G	G	R
XAC1938_260-420	I	N	L	D	G	F	K	V	I	N	D	G	Y	G	H	A	L	G	D	R	L	R	A	H	G	-	G	A	D	E	F	K	H	A	G	R
XAC3460-840-1000	M	D	I	D	H	F	K	R	I	N	D	D	Y	G	H	E	V	G	D	A	V	D	Q	A	H	C	G	G	D	E	F	K	R	Q	G	R
XAC0258-180-340	I	D	C	D	R	F	K	R	I	N	D	E	F	G	H	L	A	G	D	R	V	I	A	E	G	D	G	G	D	E	F	K	R	Q	G	G
XAC0593-480-620	I	D	C	D	R	F	K	R	I	N	D	E	F	G	H	L	A	G	D	R	V	I	A	E	G	D	G	G	D	E	F	K	R	Q	G	G
XAC0644-190-343	L	D	M	D	N	L	K	Q	I	N	D	R	F	G	H	L	A	G	D	R	V	L	G	A	Q	D	G	G	D	E	F	K	F	S	G	G
XAC0610-710-870	L	D	L	D	G	F	K	R	I	N	D	V	Y	G	H	R	A	G	D	A	V	A	G	E	R	Y	A	G	D	E	F	K	R	A	R	Q

Inativos, proteínas com ausência de vários resíduos importantes para a catálise

xac2398-270-430	I	E	P	D	H	Y	A	R	I	L	P	E	F	G	L	D	S	A	D	A	L	L	D	S	V	G	E	H	S	F	A	E	L	G	G	
xac1992-20-170	V	D	I	D	H	F	A	S	I	N	E	N	M	S	G	E	V	G	D	Q	A	L	G	A	R	G	G	S	D	E	M	A	R	E	G	
xac2711-350-500	A	Q	A	D	A	S	V	Q	Q	D	A	Q	R	L	L	Q	L	R	R		A	S	I	G	I	H	G	A	D	P	H	D	V	E	Q	
xac0495-200-340	I	E	I	A	S	A	L	G	L	R	E	R	Y	G	Y	A	A	E	R	L	A	Q	P	Y	P	N	D	N	S	F	R	L	R	S	D	
xac1993-400-560	I	E	V	D	T	R	R	M	Q	W	T	E	E	C	Y	I	H	G	L	R		Y	T	Q	D	S	A	F	E	A	A	L	T	G	L	P
XAC1795-180-350	C	G	S	D	Y	F	R	D	M	V	K	A	L	G	W	E	Y	A	D	G	Y	L	P	G	G	T	D	P	T	T	F	S	M	Y	E	R

Ligação ao cdiGMP ou GTP

Ligação ao Mg

Liga na Base do GTP

Figura 18: Análise funcional de todos os domínio GGDEF encontrados no genoma de *Xac*.

a) Sítio ativo e alostérico (I) do domínio GGDEF da estrutura de PleD de *C. crescentus* (PDBID: 2VON)^{126,173}. Os aminoácidos coloridos em verde estão envolvidos na coordenação aos íons de magnésio, os aminoácidos coloridos em azul escuro estão envolvidos na ligação a molécula de GTP e os aminoácidos coloridos em azul claro, *R₃₁₃ (painel do sítio alostérico de PleD) corresponde a outra cadeia do dímero. Em laranja claro está a estrutura de GTP (painel do sítio ativo) e de c-diGMP (no painel do sítio alostérico). **b)** classificação de todas as proteínas encontradas no genoma de *Xac*, que possui domínio GGDEF. Esta análise foi baseada na estrutura de PleD¹²⁶ e na análise feita para FimX de *P. aeruginosa*¹⁶⁰. Os resíduos coloridos em verde estão envolvidos na coordenação aos íons de magnésio, em azul ao nucleotídeo e em vermelho a base do GTP no sítio ativo. Interessantemente, todas as proteínas inativas possuem um domínio EAL fusionado.

Alguns domínios GGDEF de proteínas preditas a serem ativas ou possivelmente ativas como: XAC0424, XAC1420, XAC1940, e XAC0258, foram estudados (Materiais e Métodos). O domínio GGDEF de XAC1420 não apresentou atividade DGC-A (dados não mostrados). Este resultado é condizente com vários trabalhos recentemente publicados, que mostraram que os demais domínios fusionados aos domínios GGDEF desempenham um papel importante no processo de dimerização do domínio GGDEF e conseqüente ativação da atividade DGC-A do domínio GGDEF^{126,131,157,170-172}. Desta forma, apenas XAC1420 inteira teria atividade DGC-A.

Interessantemente, células de *E. coli* quando expressas com o plasmídio contendo o gene que codifica a proteína XAC0258 inteira formam agregados em meio líquido sugerindo que a atividade enzimática da proteína XAC0258 estava preservada (dados não mostrados). Ensaio de cristalização da proteína XAC0258 e do domínio GGDEF da XAC1420 foram feitos, mas não observamos formação de cristais nas condições testadas.

Um resultado interessante, obtido a partir dos ensaios de 2-híbridos usando domínios GGDEF como isas das proteínas XAC0258, XAC0424, XAC1420, e XAC1940, é a interação do domínio GGDEF de XAC0258_{GGDEF(182-339)} com domínios MCPs “Methyl-accepting chemotaxis domains” (**Tabela 13**). A fim de comprovar a interação de XAC0258 com o domínio MCP *in vitro* clonamos e expressamos em *E. coli* o domínio MCP de XAC3768₄₇₆₋₇₁₀. Ensaio preliminares de overlay utilizando a amostra do domínio MCP induzida imobilizada na membrana de nitrocelulose, incubada com XAC0258_{his} pura (XAC0258 fusionada a uma cauda de histidina) e revelada com anticorpo contra cauda de histidina, não mostrou nenhuma interação entre XAC0258_{his} e o domínio MCP de XAC3768₄₇₆₋₇₁₀. Ensaio de pull-down foram feitos, mas também não observamos nenhuma interação. Estes resultados são

preliminares e devem ser refeitos em outras condições para determinar se de fato estas proteínas interagem.

Tabela 13: Resultado do ensaio de 2-híbrido dos domínios GGDEF das proteínas XAC0258, XAC1420 d XAC1940.

Em parênteses se refere ao número do primeiro aminoácido da proteína que interagiu com a isca. Em coloração verde são as proteínas que acreditamos que sejam interações reais. TolC é um falso positivo⁸¹. XAC0145 é uma proteína hipotética conservada, com 1322 aminoácidos sem domínios caracterizados, encontrada em apenas quatro organismos: *Xanthomonas axonopodis* pv *citri*, *Xanthomonas campestris* pv. *Vesicatoria*, *Xanthomonas campestris* pv. *campestris* e *Caulobacter crescentus* CB.

Isca	Interagiu	Número de clones independentes
Domínio GGDEF de XAC0258	TolC - XAC3463 XAC1892 (266) XAC3768 (466) Frame negativo	3 1 1 1
Domínio GGDEF de XAC1420	TolC - XAC3463 XAC0145 (327) Frame negativo	7 3 7
Domínio GGDEF de XAC1940	TolC - XAC3463 Frame negativo	3 4

Análise funcional dos domínios EAL de Xac. A análise dos 14 domínios EAL codificados pelo genoma de *Xac* foi feita com base em estudos da estrutura das proteínas tdEAL (PDB ID: 2R6O) e BlrP1 em complexo com c-diGMP e cálcio¹⁷⁴, assim como os extensivos estudos da cinética enzimática de mutantes sítio-dirigido de RocR^{162,175}. BlrP1 possui 3 aminoácidos envolvidos na interação ao c-diGMP, Q₁₇₃, R₁₉₂ e D₃₂₅, coloridos em azul na **Figura 19**, e 4 aminoácidos envolvidos na coordenação ao metal de magnésio (na estrutura é um íon de cálcio) E₁₈₈, N₃₂₉, E₂₇₂, D₃₀₂, e três envolvidos indiretamente na ligação ao metal, K₃₂₃, E₂₆₂, e Q₃₇₉, coloridos em verde e verde escuro na **Figura 19**, respectivamente. O aminoácido E₃₅₉ é a base geral essencial para a ativação da molécula de água que irá atacar nucleofilicamente o c-diGMP, e E₂₇₅ é essencial para a estabilização do loop 6 (DDFGA/TGYSS, colorido em violeta na **Figura 19**), muito importante para o processo de dimerização e ativação do sítio ativo^{162,174,175}.

Romiling classificou os domínios EAL em três classes, *classe 1*: domínios ativos, domínios com todos os resíduos importantes conservados, *Classe 2*: domínios que podem ser ativados, ou seja são potencialmente ativos, mas não contêm os

aminoácidos conservados no loop 6, e *Classe 3*: domínios inativos, pois não contêm vários resíduos importantes para a catalise conservados. Com base nesta classificação analisamos as 14 proteínas de *Xac* com domínios EAL e 8 são domínios preditos a serem ativos (*classe 1*), 3 são potencialmente ativos (*classe 2*) e o restante são inativos (*classe 3*) **Figura 19**.

a

Sítio Ativo

b

ATIVOS

	173	188	239	272	275	302	323	359	362	379
	QA	EALIR	N	EVTE	EVTE	DDFGAGYSG	KVDAEL	EGVE	QG	QG
BlrP1_BLUF_3GFX_128-400	ALQA	EALIR	N	EVTE	EVTE	DDFGAGYSG	KVDAEL	EGVE	QG	QG
RocR-P.aerug._121-392	YYQP	EALIR	N	EITE	EITE	DDFGAGYSS	KLDRTF	EGVE	QG	QG
XAC1795-360-618	HFQP	EALLR	N	EFTE	EFTE	DDFGTGYSN	KIDQSF	EGIE	QG	QG
XAC2152-350-603	VFQP	EALLR	N	EITE	EITE	DDFGTGYSN	KLDRSL	EGIE	QG	QG
VieA-XAC2868_162-427	AYQP	EALAR	N	ELTE	ELTE	DDYGTGFSS	KLDRSF	EGVE	QG	QG
XAC1938A_438-695	HYQP	EALIR	N	EITE	EITE	DDFGTGYSN	KIDRSF	EGVE	QG	QG
xac3958-270-520	HYQP	EALVR	N	EATE	EATE	DDFGTGYSS	KIDKSF	EGVE	QG	QG
XAC2382_458-705	VYQP	EVLRL	N	ELTE	ELTE	DDFGTGFSG	KIDKSF	EGIE	QG	QG
xac2711-481-766	RFQP	EVLVR	N	ELTE	ELTE	DDFGTGYSS	KIDKSF	EGVE	QG	QG
pdeA-XAC1887_478-754	HYQP	EALLR	N	ELTE	ELTE	DDFGTGYSS	KLDMSF	EGVE	QG	QG

Podem ser Ativos - sem o loop 6

	173	188	239	272	275	302	323	359	362	379
	QA	EALIR	N	EVTE	EVTE	DDFGAGYSG	KVDAEL	EGVE	QG	QG
xac1184-1-240	DGQD	DVSTR	N	EVSE	EVSE	DDFGAGYAG	KLDIAL	EGIE	QG	QG
xac1992-180-430	RYQP	EALLR	N	EINE	EINE	DNFGTGDSS	KIDRSF	NGVE	QG	QG
xac2812-540-780	YAQK	EVLVR	N	EITE	EITE	DDFGSGLAS	KIDCGF	EGVE	QG	QG

Inativos

	173	188	239	272	275	302	323	359	362	379
	QA	EALIR	N	EVTE	EVTE	DDFGAGYSG	KVDAEL	EGVE	QG	QG
xac0495-350-600	AYQP	QVLLR	S	DLRL	DLRL	SQFEPGDEA	RMANRF	QRIE	QG	QG
xac1993-580-820	INDS	SARLH	I	NTYH	NTYH	EFTLVYQP	EALMRW	IGEA	AV	AV
xac2398-450-690	HYQP	QAFLR	R	QTPE	QTPE	EQFGSGLDS	KLDIRGI	EFVA	QG	QG

Ligação ao cdiGMP

Coordenação ao MG

Ligação indireta ao MG

Estabiliza Loop6

Loop 6

Base Geral

Figura 19: Análise funcional de todos os domínio EAL encontrados no genoma de Xac.

a) Painel a esquerda corresponde ao sítio ativo do domínio EAL da proteínas BlrP1 (PDBID: 3GFX)¹⁷⁴, e o painel a esquerda corresponde ao loop 6, colorido em lilás, envolvido na dimerização da proteína, cadeia A está colorida em laranja e a cadeia B está colorida em azul claro. Tanto na parte de cima da figura quanto na parte de baixo da figura os aminoácidos coloridos em verde claro estão envolvidos na coordenação e os em verde opaco na ligação

indireta aos íons de magnésio (neste caso é cálcio), os aminoácidos coloridos em azul estão envolvidos na ligação a molécula de c-diGMP, em vermelho corresponde a base geral envolvida na ativação da molécula de água nucleofílica, os aminoácidos coloridos em lilás correspondem a loop 6 em que E₂₇₅ é importante para a estabilização deste loop, mostrado no painel a direita. **b)** O alinhamento de todos os domínios EAL encontrados no genoma de *Xac*. Estes foram classificados com base na classificação feita por Romling¹⁶². A coloração dos resíduos é a mesma que a ilustrada no item **a** com exceção dos resíduos envolvidos indiretamente na ligação aos íons de magnésio (neste caso é cálcio) que estão coloridos em verde escuro.

Análise funcional dos domínios HD-GYP de Xac. Pouco se sabe sobre o papel biológico das proteínas que contêm domínios HD-GYP, apesar de ser um domínio encontrado em várias proteínas codificadas por muitos genomas bacterianos¹⁷⁶. A maioria das informações sobre a atividade PDE-A deste domínio foi obtida por estudos funcionais da RpfG de *Xcc*¹⁷⁷. Apenas o domínio HD-GYP da RpfG é capaz de degradar o c-diGMP à duas moléculas de GMP¹⁷⁷, em que os resíduos HD do motivo consenso HD-GYP estão envolvidos na atividade PDE-A da RpfG⁴¹ e os resíduos GYP estão envolvidos na capacidade deste domínio interagir com domínios GGDEF de algumas proteínas¹⁷⁸. *P. aeruginosa* possui três proteínas com domínios HD-GYP, sendo que duas delas, PA4108 e PA4781, possuem o motivo consenso, HD-GYP, intacto enquanto que, PA2572 possui o motivo consenso corrompido, YN-GYP¹⁷⁹. PA4108 e PA4781 possuem atividade PDE-A mas PA2572 não possui atividade enzimática¹⁷⁹ mostrando que os resíduos HD do motivo consenso HD-GYP tem papel essencial na atividade PDE-A dos domínios HD-GYPs.

Com base nestas observações analisamos a presença do motivo consenso HD-GYP nas três proteínas codificadas pelo genoma de *Xac*, RpfG_{XAC1877}, XAC2493 e XAC0350. Todas possuem os resíduos HD importantes para a atividade PDE-A (**Figura 20**). A análise de bioinformática de 211 proteínas contendo domínios HD-GYP mostrou que a família HD-GYP possui o motivo HD-HHExxDGxGYPxxL conservado¹⁷⁸ e todas as proteínas de *Xac* possuem estes resíduos (**Figura 20**). Todas estas informações indicam que RpfG_{XAC1877}, XAC2493 e XAC0350 sejam enzimaticamente ativas. É importante salientar que deva haver outros resíduos, além do motivo consenso HD-GYP, importantes para a atividade enzimática e/ou para a capacidade deste domínio interagir com outras proteínas, que serão desvendados futuramente com mais estudos funcionais desta classe de proteínas.

Figura 20: Alinhamento múltiplo das seqüências de aa dos domínios HD-GYP.

O quadrado mostra o motivo consenso HD e GYP das proteínas RpfG_{XAC877}, XAC2493 e XAC0350. HD-HHExxDGxGYPxxL é o motivo conservado na família HD-GYP¹⁷⁸, mostrado em violeta acima do alinhamento múltiplo.

Os domínios HD-GYP das proteínas XAC0350 (resíduos de 171 à 311, XAC0350_{His171-311}), XAC1877 (resíduos de 199 à 333, XAC1877_{His199-333}) e XAC2493 (resíduos de 146 à 288, XAC2493_{His146-288}) foram clonados no vetor de expressão pET28a e todos os domínios vão para corpos de inclusão. O protocolo de purificação destas proteínas estão descritos nos Materiais e Métodos assim como os ensaios de atividade fosfodiesterase destes domínios. Nenhuma atividade PDE-A foi observada para estas construções (dados não mostrados).

Existem várias hipóteses para não ter sido observado atividade PDE-A para estes domínios: a) Os domínios realmente não possuem atividade fosfodiesterase; b) Os domínios HD-GYP não foram reenovelados corretamente ou c) Os domínios HD-GYP necessitam interagir com outras proteínas para que sua atividade seja ativada. Ensaio de cristalização para os três domínios foram feitos. No entanto, nenhum cristalizou nas condições testadas.

Resumo das previsões sobre as atividades DGC-A e PDE-A para as proteínas com domínios GGDEF, EAL e HD-GYP. A Tabela 14 resume a previsão de atividade PDE-A, DGC-A de todas as 34 proteínas de *Xac* que contem domínios GGDEF e/ou EAL e para as 3 proteínas que contem domínios HD-GYP (Figura 18-20). A Tabela 14 demonstra que *Xac* é um excelente exemplo da diversidade de arquiteturas e de atividades que proteínas com domínios GGDEF e/ou EAL podem apresentar. Dentre as proteínas de *Xac* que possuem domínios GGDEF e EAL, XAC1887 possui ambos os domínios ativos, XAC1795 possui apenas o domínio EAL ativo, XAC1938 possui o domínio EAL ativo e o GGDEF possivelmente ativo, XAC2812 possui o domínio

GGDEF ativo e o EAL possivelmente ativo, XAC1992 possui o domínio EAL possivelmente ativo e o GGDEF inativo e XAC0495 e XAC2398 possuem ambos os domínios inativos. Esta variedade de atividade pode sugerir a diversidade de mecanismos que estas proteínas possam assumir para regular a sua própria atividade ou a atividade de receptores a c-diGMP.

Tabela 14: Previsão da atividade DGC-A e/ou PDE-A de todas as proteínas de *Xac* que contêm domínio GGDEF, EAL e HD-GYP.

+: possui atividade, -: não possui atividade, ?: pode possuir atividade e em branco não possui o domínio.

Proteína	GGDEF (DGC-A)	EAL (PDE-A)	HD-GYP (PDE-A)
XAC1420	+		
XAC1345	+		
XAC2897	+		
XAC1486	+		
XAC1570	+		
XAC0424	+		
XAC2812	+	?	
XAC0614	+		
XAC1345	+		
XAC4358	+		
XAC1940	+		
XAC1488	+		
XAC1887	+	+	
XAC2382	?	+	
XAC2810	?		
XAC1938	?	+	
XAC3460	?		
XAC0258	?		
XAC0593	?		
XAC0644	?		
XAC0610	?		
XAC2398	-	-	
XAC1992	-	?	
XAC2711	-	+	
XAC0495	-	-	
XAC1993	-	-	
XAC1795	-	+	
XAC2152		+	
XAC2868		+	
XAC3958		+	
XAC1939	+		
XAC2482	+		
XAC1184		?	
XAC1488	+		
XAC1877			+
XAC2493			+
XAC0350			+

Análise do sítio ativo dos domínios GGDEF e EAL de FimX. A comparação do sítio ativo do domínio GGDEF de PleD com o sítio ativo dos domínios GGDEF de FimX de *P. aeruginosa*¹⁶⁰ e de FimX_{XAC2398} de *Xac* mostrou que os aminoácidos envolvidos na interação com a molécula de c-diGMP no sítio alostérico do domínio GGDEF, RxxD, não estão presentes na sequência de FimX de *P. aeruginosa* nem em FimX_{XAC2398} de *Xac*. Esta análise inviabiliza a possibilidade dos FimX interagirem com c-diGMP neste sítio (**Figura 21**). Ambas as proteínas FimX não possuem todos os resíduos importantes para a interação ao GTP e ao magnésio no sítio ativo do domínio GGDEF. Isto sugere que as proteínas não interajam com GTP nem com íons de magnésio no sítio ativo do domínio GGDEF (**Figura 21**). Esta análise é condizente com os dados experimentais observados por Navarro e colaboradores e pelo nosso grupo, de que FimX de *P. aeruginosa* e de *Xac* não possuem atividade diguanilato ciclase e não interagem com GTP^{6,160} (**Figura 6 do Anexo 2 e Figura 11**).

O mecanismo de catálise dos domínios EAL é composto por uma base geral, três resíduos envolvidos na ligação ao c-diGMP, quatro resíduos envolvidos na coordenação ao íon de magnésio, três resíduos envolvidos indiretamente na ligação ao íon de magnésio e cinco resíduos importantes para a estabilidade e função do loop 6 na dimerização da proteína e na ativação do sítio ativo (**Figura 21**)^{175, 162,168,174}. A predição da atividade dos domínios EAL de FimX de *P. aeruginosa* e de FimX_{XAC2398} de *Xac* usando como base a estrutura da proteína BlrP1 (PDBID: 3GFX)¹⁷⁴ em complexo com c-diGMP e cálcio, assim como a classificação feita por Romling dos domínios EAL¹⁶² mostrou que tanto FimX de *P. aeruginosa* quanto FimX_{XAC2398} de *Xac* possuem os resíduos envolvidos na interação ao c-diGMP (resíduos coloridos em azul na **Figura 21 b e c** e vide abaixo). No entanto, FimX de *P. aeruginosa* não possui o glutamato importante para a ativação da molécula de água que irá atacar nucleofilicamente o c-diGMP, mas FimX de *Xac* o possui (colorido em vermelho na **Figura 21 b e c**). Estas observações sugerem que FimX de *P. aeruginosa* não seria capaz de clivar a molécula de c-diGMP, condizente com os resultados observados por Navarro e colaboradores¹³³.

Apesar de FimX_{XAC2398} possuir a base geral, E₆₅₃, a sobreposição do sítio de ligação ao magnésio do domínio EAL da proteína BlrP1 (PDBID: 3GFX)¹⁷⁴ com os resíduos correspondentes na estrutura de FimX de *P. aeruginosa* (PDBID: 3HV8) e com o modelo da estrutura de FimX_{XAC2398} (criada pelo programa PHYRE “Protein Homology/analogY Recognition Engine” utilizando o modelo 2r6o) mostra que ambas

as proteínas possuem substituições nas posições importantes para coordenação o íon de magnésio (**Figura 21 c**).

Tanto FimX de *Xac* quanto de *P. aeruginosa* possuem os resíduos importantes para a ligação ao c-diGMP, sugerindo que ambos seriam capazes de ligarem ao c-diGMP (**Figura 21**). Ambas as proteínas também possuem o glutamato (E₅₇₉ em FimX_{XAC2398}) essencial para a estabilização do loop 6 (**Figura 21 b**), mas possuem um grande número de substituições nas posições conservadas no loop 6 descritas como importantes para a catálise enzimática^{174,175}. Todas estas observações são condizentes com os resultados obtidos por Navarro e colaboradores e por nosso grupo de que ambas as proteínas não possuem atividade PDE-A, mas conseguem interagir com c-diGMP^{6,160} (**Figura 6 e Figura suplementar 8 do Anexo 2**).

Apesar de nós não termos observado nenhuma atividade PDE-A para FimX_{XAC2398}, mesmo na presença de PilZ_{XAC1133}, não pode ser excluída a possibilidade de FimX de *Xac* possuir atividade PDE-A em condições específicas. A ativação poderia vir por algum tipo de sinalização intracelular que seria transmitida pelos demais domínios de FimX, RR, PAS e/ou GGDEF, ou/e pela interação de FimX com outras proteínas.

lilás correspondem ao loop 6 em que E₂₇₅ é importante para a estabilização deste loop. **c**: Sobreposição do sítio de interação a c-diGMP da estrutura de FimX de *P. aeruginosa* em complexo com c-diGMP (PDBID: 3HV8) com os resíduos correspondentes na estrutura do modelo de FimX de *Xac* criada pelo programa PHYRE “Protein Homology/analogY Recognition Engine”¹⁸⁰ usando como modelo a estrutura de PDBID: 2r6o. Os resíduos coloridos em laranja claro correspondem a estrutura de FimX de *P. aeruginosa* e os coloridos em azul e vermelho (o resíduo colorido em vermelho corresponde a base geral necessária para a ativação da água nucleofílica) pertencem a estrutura do modelo de FimX de *Xac*. **d**: Sobreposição do sítio de interação ao íon de magnésio (neste caso é cálcio) da estrutura de BlrP1 (PDBID: 3GFX) colorido em rosa claro, com os correspondentes na estrutura de FimX de *P. aeruginosa* em complexo com c-diGMP (PDBID: 3HV8) coloridos em laranja claro e com os correspondentes no modelo da estrutura de FimX_{XAC2398} coloridos em azul claro. O cátion de cálcio está mostrado na forma de esfera colorida em amarelo.

Análise estrutural do domínio EAL de FimX de *P. aeruginosa* em complexo com c-diGMP. Observamos que apenas o fragmento contendo os domínios GGDEF-EAL de FimX_{XAC2398} interagiu com c-diGMP, mas os domínios isolados não interagem com c-diGMP (dados não mostrados). No entanto, o domínio EAL isolado de FimX de *P. aeruginosa* foi demonstrado de ser capaz de ligar ao c-diGMP^{6,160}.

Esta diferença de resultados sugere que o domínio EAL de FimX_{XAC2398} isolado utilizado em nossos estudos (resíduos 438-689) possa estar parcialmente desenovelado que poderia de alguma maneira diminuir sua afinidade ao c-diGMP. Uma outra possibilidade é que o sítio ativo do domínio EAL somente se forme corretamente quando os dois domínios (GGDEF-EAL) estão juntos. Esta segunda hipótese é consistente com a análise da estrutura do domínio EAL de FimX de *P. aeruginosa* em complexo com c-diGMP (PDBID: 3HV8). A **Figura 22 a** mostra que a superfície de ligação ao c-diGMP é próxima a região N-terminal deste domínio que por sua vez está conectada ao domínio GGDEF por uma alça de somente seis aminoácidos^{6,160}. Isso sugere, que o domínio GGDEF possa fazer contatos diretos com c-diGMP ou possa afetar a ligação do domínio EAL ao c-diGMP de forma indireta. A **Figura 22 b** apresenta uma sobreposição das estruturas dos domínios EAL de FimX de *P. aeruginosa* determinadas na presença e na ausência do domínio GGDEF. Interessantemente o domínio GGDEF é móvel o suficiente no cristal que não possui densidade eletrônica¹⁶⁰, **Figura 22 b**.

Figura 22: Estrutura do domínio EAL de FimX de *P. aeruginosa* em complexo com c-diGMP.

Estruturas do domínio EAL de FimX de *P. aeruginosa* na conformação apo e em complexo com c-diGMP¹⁶⁰ **(a)** Estrutura do domínio EAL isolada (PDBID: 3HV8). **(b)** Sobreposição da estrutura em (a) com a estrutura do domínio EAL dentro do fragmento GGDEF-EAL em que o domínio GGDEF é desordenada (PDBID: 3HVB). A região N-terminal dos domínios EAL estão coloridas em verde. A estrutura do domínio EAL cristalizado isoladamente está colorido em azul escuro e a estrutura do fragmento contendo os domínio GGDEF-EAL está colorido em azul claro¹⁶⁰.

Células nocautes de *pilZ*, *fimX* e *pilB* dispersam mais em superfície semi-sólida do que células tipo selvagem. Células nocautes de *pilZ*_{XAC1133}, *pilB*_{XAC3239} e *fimX*_{XAC2398} dispersam mais em meio semi-sólido SB quando comparadas com células tipo selvagem (**Figura 8 do Anexo 2 e Figura 12**). No entanto, como *pilZ*_{XAC1133}, *pilB*_{XAC3239} e *fimX*_{XAC2398} modulam estas mudanças na motilidade em *Xac* não está clara.

A análise das estruturas filamentosas na superfície de *Xac* quando as células são crescidas em meio semi-sólido contendo 0,6% agar mostrou que *Xac* possui diferentes estruturas filamentosas na superfície da bactéria (**Figura 13**). A forma destes filamento é dependente da composição do meio. Em meios ricos (SB e Eiken) *Xac* apresenta filamentos finos e longos, possivelmente T4P, na sua superfície, enquanto que em meio mínimo, como no caso de XVM2 e STON, *Xac* forma filamentos longos que formam bolhas que geralmente conectam duas bactérias. A frequência destes filamentos é severamente reduzida em células nocautes de *pilZ*_{XAC1133} (**Tabela 12, Figura 13**).

Mecanismos de polimerização de PilA via PilB. Em outras bactérias, membros ortólogos de PilZ, FimX e PilB foram implicados na biogênese do T4P^{3,5,107,111,181-183}. PilZ é necessária para a secreção do T4P em *P. aeruginosa*³ enquanto que em *Neisseria meningitidis* o mutante de *pilZ* apresenta T4P mas perdeu a capacidade de agregar¹¹³. PilB é uma ATPase necessária para a polimerização de subunidades de pilinas para a formação do T4P¹⁸⁴ e sua atividade é o oposto da atividade de PilT, cuja atividade ATPásica é necessária para o processo de despolimerização do T4P¹⁸⁴. FimX de *P. aeruginosa* localiza-se no pólo da célula de forma dependente do domínio RR e da atividade PDE-A do domínio EAL⁵. FimX contendo apenas o domínio RR se localiza no pólo da célula provavelmente via interação com outra proteína desconhecida de membrana⁵.

O complexo ternário observado em nosso estudo, FimX_{XAC2398}–PilZ_{XAC1133}–PilB_{XAC3239}, indica que FimX_{XAC2398} possa estar presa na membrana via o domínio N-terminal RR e o domínio C-terminal EAL. A **Figura 9 do Anexo 2** apresenta estas interações no contexto da estrutura do T4P e os processos responsáveis pela sua polimerização e despolimerização. O que é importante agora é de tentar entender de qual maneira estas interações modulam o funcionamento do T4P. Para isso, precisamos primeiro entender um pouco sobre a maneira em que as pilinas são polimerizadas por PilB para formar um T4P de forma dependente de ATP.

Um mecanismo geral de polimerização de pilinas proposto por Craig e colaboradores⁹² está mostrado na **Figura 23 a**. PilB, proteína citossólica, forma uma estrutura hexamérica associada à face citossólica da membrana interna em que sua localização no pólo da célula é dependente de PilC, proteína de membrana interna¹⁰⁷. A hidrólise de ATP por PilB ocasiona mudanças conformacionais na estrutura hexámerica de PilB que causa um efeito pistão na proteína MBP (*membrane binding partner*; provavelmente PilC). Este movimento ocasiona a translação do filamento de TP4 sendo formado em 10,5Å para cima, criando um buraco que será preenchido futuramente por uma nova subunidade de PilA **Figura 23 a**^{92,181}. A troca de ADP para ATP neste processo de polimerização reverteria a mudança conformacional de PilB que está na forma helicoidal para a forma relaxada, estrutura de toróide¹⁸⁵.

Este modelo leva em consideração as informações estruturais da proteína EpsE (*extracellular protein secretion*) de *Vibrio cholerae*, que pertence ao sistema de secreção tipo II, T2SS^{181,186} e é um homólogo de PilB. Assim sendo, o mecanismo de ação de EpsE deve ser parecido com o que ocorre com PilB. Interessantemente, o sítio de

interação de BfpD de *E. coli*, homólogo ao PilB, com BfpE, homólogo ao PilC, muda conforme o estado de ligação do BfpD ao ATP ou ADP¹⁸⁷. Isso sugere que PilB poderia interagir com PilC, MBP no modelo mostrado na **Figura 23 a**. Estas informações são consistentes com a hipótese de que PilB transmitiria a sua mudança conformacional do citosol para o periplasma via PilC (**Figura 23 a**).

O filamento do T4P é composto principalmente por sub-unidades de PilA, que formam três fitas que se enrolam entre si formando uma estrutura de *3-helix bundle* **Figura 23 b**¹⁰⁵. As interações entre as sub-unidades de PilA de *N. gonorrhoeae* ocorrem entre a α -hélice de PilA, α_1 (**Figura 23 c**)¹⁰⁵. A face que contém as fitas betas da região globular de PilA, que compõe a região C- terminal da α_1 , α_1 -C e as fitas betas, estão expostas ao longo do filamento (**Figura 23 d**)¹⁰⁵.

A porção N-terminal da α_1 , α_1 -N, possui apenas um resíduo hidrofílico E₅ e os demais resíduos são hidrofóbicos (**Figura 23 d**). A cadeia lateral do E₅, altamente conservado, forma um ponte de hidrogênio com o grupo α -amino da F₁ da próxima sub-unidade pertencente a outra fita do *3-helix bundle*. Desta forma, as interações entre sub-unidades de diferentes fitas da estrutura *3-helix bundle* é composta principalmente por interações hidrofóbicas e uma interação de ponte de hidrogênio entre o E₅ e F₁, **Figura 23 c**¹⁰⁵.

No modelo mostrado na **Figura 23 a**, há três sítios ativos (sítios de entrada para uma subunidade de pilina) ao redor do filamento (apenas um está mostrado na Figura). Conforme sub-unidades de PilA são adicionadas, o filamento sobe em 10,5 Å, que cria um buraco na próxima fita do *3-helix bundle*, onde pode ser adicionada uma nova subunidade no próximo passo de hidrólise de ATP por PilB. Portanto, a adição de três sub-unidades, uma para cada fita do *3-helix bundle*, resulta na formação de um buraco de 31,5 Å na primeira fita, que irá acomodar uma nova sub-unidade de PilA. Este processo pode ser revertido por PilT pela dissociação das sub-unidades de PilA pela reversão do pistão¹⁸⁸.

PilA de *N. gonorrhoeae* possui 37% de identidade e 51% de similaridade com XAC3241 (**Figura 23 d**) e 32% de identidade e 48% de similaridade com XAC3240. A região N-terminal de PilA de *N. gonorrhoeae* e de PilA_{XAC3241} são altamente similares, que corresponde a interface de interação entre as sub-unidades de PilA no filamento⁹². A região C-terminal, que é menos conservada (**Figura 23 d**) é o que dá variabilidade funcional entre as pilinas de diferentes organismos⁹². Desta forma, XAC3241 e/ou

XAC3240 poderiam formar o mesmo tipo de filamento, *3-helix bundle*, que PilA de *N. gonorrhoeae*. É interessante ressaltar que apesar de haver baixa similaridade em sequência primária na região C-terminal entre XAC3241 e PilA de *N. gonorrhoeae*, XAC3241 possui as duas cisteínas que formam um grampo β que contem o loop hipervariado, próximo a região D, **Figura 23c** e α^{92} .

Possíveis mecanismos de regulação da biogênese do T4P via o complexo ternário FimX-PilZ-PilB. Nossos resultados de interação de PilZ_{XAC1133} com FimX_{XAC2398} e PilB_{XAC3239}, posiciona PilZ_{XAC1133} e FimX_{XAC2398} diretamente na base do T4P. Estes resultados apontam para as possíveis hipóteses sobre a maneira na qual PilZ e FimX poderiam afetar a atividade de PilB (**Figura 9 do Anexo 2**):

- i) PilZ_{XAC1133} poderia regular a atividade ATPase de PilB_{XAC3239};
- ii) FimX_{XAC2398}, através de sinais alostéricos sentidos pelos domínios regulatórios, REC, PAS, GGDEF ou EAL, poderia modular a interação entre PilZ e PilB,
- iii) FimX_{XAC2398} poderia interagir diretamente com PilB_{XAC3239} regulando sua atividade, através de interações diretas que somente aconteçam na presença de PilZ_{XAC1133} ou
- iv) o complexo FimX_{XAC2398}-PilZ_{XAC1133}-PilB_{XAC3239} poderia controlar a localização de PilB_{XAC3239} ou PilT no processo de polimerização e despolimerização, respectivamente, do pilus tipo IV.

de mão direita (4-start)). Em destaque a interação via ponte de hidrogênio entre a F_1 de uma fita com o E_5 de outra fita, na região da α_1 -N. As demais interações são hidrofóbicas. **c)** Estrutura de PilA de *N. gonorrhoeae* (PDBID: 2HI2). A α_1 é dividida na região N-terminal α_1 -N, composta pelos resíduos 1-29, e a região C-terminal, α_1 -C, composta pelos resíduos 30-53. A parte globular de PilA é composta pela região α_1 -C e as 5 fitas β . A S_{63} e a S_{68} sofrem mudanças pós-traducionais, A S_{63} está ligada ao Gal-DADDGlc (α -D-galactopiranosil-(1->3)-2,4-diacetomido-2,4-dideoxi- β -D-glicopiranosilide), e a S_{68} ao fosfoetanolamina (PE). Em verde está destacando o loop $\alpha\beta$, que faz contatos com o loop super-variável (D-region, colorido em violeta) de outra sub-unidade presente em outra fita. Em cyan está mostrado a ponte tiosulfeto de duas cisteínas, C_{121} e C_{151} . **d)** Alinhamento entre PilA de *N. gonorrhoeae* com PilA_{XAC3240} e PilA_{XAC3241}, resíduos importantes para a função de PilA estão coloridos em vermelho. Os resíduos de PilA que compõem a α_1 estão grifados em cyan, os resíduos que compõem as fitas- β estão grifado em amarelo e os resíduos que formam o loop $\alpha\beta$ estão grifados em verde. Partes a, b e c desta figura são derivadas do manuscrito publica por Craig e colaboradores¹⁰⁵.

3.5 Conclusões

Nesta parte do trabalho determinamos a estrutura tridimensional de PilZ_{XAC1133} utilizando a técnica MAD, sendo a primeira estrutura tridimensional depositada para um membro da família PA2690/XAC1133^{6,136}. A estrutura de PilZ_{XAC1133} assim como a estrutura de XC1028 (recentemente determinada)¹⁴⁰, mostraram diferenças significativas no sítio de interação a c-diGMP quando comparadas com as outras três estruturas de PilZ depositadas. Estas diferenças estruturais explicam claramente porque membros da família PA2690/XAC1133 não interagem com c-diGMP, por não possuírem o sítio de interação a c-diGMP. O assinalamento da cadeia principal de PilZ_{XAC1133}, utilizando a técnica de RMN, mostrou que a estrutura secundária de PilZ_{XAC1133} em solução é condizente com aquela observada na estrutura cristalina.

A primeira correlação entre PilZ com T4P foi observada em 1996³, mas como PilZ está envolvida no controle da função do T4P não tinha sido elucidada. Nossos resultados demonstram, pela primeira vez, a interação física entre PilZ com duas proteínas envolvidas na regulação da biogênese do T4P, PilB e FimX e sugere possíveis mecanismos de como PilZ poderia modular a atividade do pilus tipo IV. Mostramos que PilZ_{XAC1133} interage especificamente com o domínio EAL de FimX_{XAC2398} e que apesar de PilZ_{XAC1133} não ter a habilidade de interagir com c-diGMP, seu parceiro, FimX_{XAC2398} tem. O complexo PilZ-FimX é um dos primeiros sistemas descritos que contem um conjunto de domínios associados a sinalização por c-diGMP degenerados; ou seja, o domínio RR de FimX não contem o aspartato conservado para ser fosforilado, os domínios GGDEF e EAL de FimX são inativos e PilZ_{XAC1133} não é um

receptor de c-diGMP. Estas observações sugerem que a regulação da função destas proteínas evoluiu, ou está evoluindo, para ser independente do segundo mensageiro bacteriano, e ser uma regulação via interações proteína-proteína. Como a ligação de c-diGMP ao domínio EAL de FimX regula este sistema, ainda não está clara.

Ensaio de mutação sítio dirigida de resíduos conservados nos membros da família PA2690/XAC1133, que são hidrofóbicos ou aromáticos e estão expostos na estrutura de PilZ_{XAC1133}, mostrou que alguns deles, Y25, R97 e L104 diminuíram a estabilidade da proteína, mas Y22, F28, W69, E101 e a deleção dos dez últimos resíduos de PilZ_{XAC1133} não afetaram sua estabilidade. A deleção dos dez últimos resíduos de PilZ_{XAC1133} afetou sua interação apenas com PilB_{XAC3239}, enquanto que a mutação na Y22A afetou a sua interação com PilB_{XAC3239} e com FimX_{XAC2398}. Estes resultados mostram que o C-terminal de PilZ_{XAC1133} faz parte da interface de interação com PilB_{XAC3239} e sugerem que a Y22 provavelmente tem um papel regulatório na interação de PilZ_{XAC1133} com ambas as proteínas.

Apesar de não termos conseguido observar motilidade *twitching* para células de *Xac*, células nocautes de *pilZ*_{XAC1133}, *pilB*_{XAC3239} e de *fimX*_{XAC2398} crescidas em meio semi-sólido SB-0,6% agar espalham mais rapidamente do que as células tipo selvagem. Esta motilidade, aparentemente se trata da motilidade *sliding* que, em *P. aeruginosa* não necessita da presença de flagelo e nem T4P na superfície da bactéria. Estes filamentos na superfície da bactéria diminuem a sua dispersão em meio semi-sólido.

Estes resultados abrem novas hipóteses de regulação do T4P envolvendo FimX-PilZ-PilB e c-diGMP. Estudos mais aprofundados destas interações devem ser feitos para entender como FimX_{XAC2398} e PilZ_{XAC1133} afetam a atividade de PilB_{XAC3239} e como estas proteínas estão envolvidas na motilidade em *Xac*. Além disso, a compreensão do papel das pilinas majoritárias e minoritárias na composição e no funcionamento do T4P nos diferentes meios semi-sólidos testados devem ser elucidadas para compreender melhor o mecanismo de funcionamento do T4P em *Xac*.

CAPÍTULO 4

YaeQ Pertence à superfamília de PD-(D/E)XK endonucleases

4.1 Introdução

YaeQ é altamente conservada em proteobactérias. A família de YaeQ é altamente conservada sendo encontrada em 222 espécies que inclui principalmente proteobactérias gram-negativas e actinobactérias gram-positivas ricas em CG. Genes que codificam proteínas homólogas a YaeQ podem ser encontrada também em alguns organismos eucarióticos como *Ricinus communis* (mamona), *Populus trichocarpa* (árvore) e *Nematostella vectensis* (anêmona marinha)¹⁸⁹.

O fato do gene de *yaeQ* ser encontrado em tantos organismos sugere que a sua função biológica seja importante. A sequência primária de YaeQ é altamente conservada entre membros da família de YaeQ, em especial na região N-terminal (**Figura 2 do Anexo 4**).

A função de YaeQ é desconhecida. Existem dois trabalhos publicados que descrevem a possível função de YaeQ. O primeiro trabalho apresenta evidências de que YaeQ estaria envolvida no controle de transcrição de fatores de virulência em bactérias, pois o gene de *yaeQ* suprime os efeitos causados pela mutação no gene de *rfaH* em *Salmonella typhimurium* e em *E.coli*.¹⁹⁰ A proteína RfaH, possivelmente atua como um anti-terminador de transcrição, regula a transcrição dos genes no operon *hlyCABD* responsáveis pela produção de hemolisinas. No entanto, um segundo trabalho verificou que o gene de *yaeQ* não restaurou a atividade hemolítica em células nocautes de *rfaH* em *E.coli*¹⁹¹, contradizendo os resultados do primeiro trabalho. Portanto, a função de YaeQ se mantém desconhecida.

Contexto gênico de *yaeQ*. A ausência de informações sobre a função de YaeQ na literatura nos levou a analisar a organização gênica de *yaeQ* nos diferentes organismos que a possui. Em espécies de *Xanthomonas* (**Figura 24**), como *Xac*, *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), *Xanthomonas campestris* pv. *vesicatoria* (*Xcv*), *Xanthomonas campestris* pv. *campestris* (*Xcc*), e *Stenotrophomonas maltophilia* (*Sml*) esta organização gênica é composta por três genes: *yaeQ*_{XAC2396}, *XAC2397* e *XAC2398*. O gene de *yaeQ*_{XAC2396} e *XAC2397* formam um possível “cistron” (o stop códon de *yaeQ*_{XAC2396} sobrepõe ao códon de iniciação de *XAC2397*). No entanto, este possível operon não é encontrado em outros organismos (de fato, nenhum padrão de organização gênica envolvendo o gene de *yaeQ* está observado na maioria dos genomas sequenciados).

Em relação as proteínas codificadas por este *operon*, XAC2397 é uma proteína hipotética conservada com domínio HI0933-like caracterizado por possuir motivo de ligação a dinucleotídeo¹²⁰. XAC2398 pertencente a família FimX, (descrita no **Capítulo 3**) que possui quatro domínios envolvidas em processos de sinalização intracelular: a) RR (regulador de resposta); b) PAS (domínio sensor); c) EAL e d) GGDEF.

Figura 24: Operon que *yaeQ* pertence em espécies de *Xanthomonas*¹⁸⁹.

A estrutura tridimensional protéica pode fornecer dados funcionais chaves para desvendar a função biológica de proteínas sem função conhecida. Dois exemplos que utilizam informações da estrutura tridimensional da proteína para inferir sua função biológica que era desconhecida é no caso da proteína TM841 de *Thermotoga marítima* e MTH149 de *Methanobacterium thermoautotrophicum*^{192,193}. TM841 pertence a uma família grande de proteínas, DegV ou COG1307, de função desconhecida, e a sua estrutura tridimensional possui baixa similaridade com outras estruturas depositadas no PDB¹⁹². No entanto, TM841 cristalizou com uma molécula de ácido graxo e em conjunto com uma análise de bioinformática dos genes ao redor de TM841 foi possível chegar na conclusão de que TM841 possa ter um papel no processo de transporte de lipídios ou possa fazer parte do metabolismo de ácido graxo em complexo com outra proteína, já que TM841 não possui os resíduos para catalise enzimática¹⁹². MTH1491 é uma proteína hipotética com domínio DrsE em que sua estrutura tridimensional possui similaridade estrutural com dehidrogenases, amidohidrolases e oxidoreductases¹⁹³. A análise estrutural de MTH1491 em conjunto com a análise dos genes próximos ao MTH1491 sugere que ela possa estar envolvida no processo de oxidação de enxofre¹⁹³. Desta forma, informações estruturais fornecem dados importantes para direcionar estudos funcionais de proteínas sem função conhecida.

4.2 Objetivos

Tínhamos como objetivo determinar a estrutura tridimensional da proteína hipotética conservada YaeQ_{XAC2396} utilizando a técnica de difração de raios-X. O intuito era que a determinação da sua estrutura tridimensional poderia fornecer indícios sobre a sua função biológica com base na similaridade com outras estruturas de proteínas com funções conhecidas depositadas no PDB.

4.3 Resultados

4.3.1 Cristalização de YaeQ

YaeQ_{XAC2396} cristalizou em diferentes condições provenientes dos Kits de cristalização, descritos nos Materiais e Métodos. Todas as condições de cristalização foram refinadas e a melhor condição de cristalização, que resultou na determinação da estrutura de YaeQ_{XAC2396}, assim como a coleta e processamento dados de difração estão descritos no **Anexo 3**⁸³.

Anexo 3 – Expressão, purificação, cristalização e análise preliminar de Raios-X de YaeQ de *Xanthomonas axonopodis* pv. *citri*

- Expression, purification, crystallization and preliminary X-ray analysis of YaeQ (XAC2396) from *Xanthomonas axonopodis* pv. *citri*

crystallization communications

Acta Crystallographica Section F
Structural Biology
and Crystallization
Communications
ISSN 1744-3091

Cristiane R. Guzzo,^a
Ronaldo A. P. Nagem,^b
Leonor M. P. Galvão-Botton,^a
Beatriz G. Guimarães,^c
Francisco J. Medrano,^c
João A. R. G. Barbosa^{c*} and
Chuck S. Farah^{a*}

^aDepartamento de Bioquímica, Instituto de Química, Universidade de São Paulo, CEP 05508-000, São Paulo, SP, Brazil, ^bInstituto de Física de São Carlos, Universidade de São Paulo, CEP 13560-970, São Carlos, SP, Brazil, and ^cCentro de Biologia Molecular Estrutural, Laboratório Nacional de Luz Síncrotron, CEP 13084-971, Campinas, SP, Brazil

Correspondence e-mail: joao@lnls.br, chsfarah@iq.usp.br

Received 25 February 2005
Accepted 7 April 2005
Online 22 April 2005

Expression, purification, crystallization and preliminary X-ray analysis of YaeQ (XAC2396) from *Xanthomonas axonopodis* pv. *citri*

Xanthomonas axonopodis pv. *citri* YaeQ (XAC2396) is a member of a family of bacterial proteins conserved in several Gram-negative pathogens. Here, the cloning, expression, purification and crystallization of the 182-residue (20.6 kDa) YaeQ protein are described. Recombinant YaeQ containing selenomethionine was crystallized in space group $P2_1$ and crystals diffracted to 1.9 Å resolution at a synchrotron source. The unit-cell parameters are $a = 39.75$, $b = 91.88$, $c = 48.03$ Å, $\beta = 108.37^\circ$. The calculated Matthews coefficient suggests the presence of two YaeQ molecules in the asymmetric unit. Initial experimental phases were calculated by the multiple-wavelength anomalous dispersion technique and an interpretable electron-density map was obtained.

1. Introduction

Xanthomonas is a genus of phytopathogenic bacteria that attack a variety of hosts with economic importance, including citrus fruits, rice, beans, grapes and cotton. Citrus canker is caused by the phytopathogen *Xanthomonas axonopodis* pv. *citri* (*Xac*), a rod-like Gram-negative motile bacterium. The *Xac* genome (Da Silva *et al.*, 2002) has a 5 175 554 bp chromosome and two megaplasmids: pXAC33 (33 699 bp) and pXAC64 (64 920 bp). Genome annotation identified over 4400 open reading frames, 2770 of which were assigned putative functions based on similarities with sequences in the public databases (Da Silva *et al.*, 2002).

Xac YaeQ (XAC2396) is a member of a family of bacterial proteins that are conserved in several Gram-negative pathogens, including *Salmonella* ssp., *Escherichia coli*, *Pseudomonas* ssp., *Shigella* ssp., *Yersinia* ssp., *Ralstonia* ssp., *Erwinia* ssp., *Bordetella* ssp., *Burkholderia* ssp., *Vibrio parahaemolyticus* and *Chromobacterium violaceum*, as well as *Azotobacter vinelandii*, *Dechloromonas aromatica*, *Shewanella oneidensis*, *Geobacter metallireducens* and *G. sulfurreducens*. This 182-amino-acid residue protein was selected for structural analysis owing to its well folded and non-aggregated state in solution, both in bacterial lysates and in the purified form (Galvão-Botton *et al.*, 2003).

It has been suggested that YaeQ is involved in the regulation of transcription of genes encoding virulence factors such as haemolysin and enzymes involved in the production of lipopolysaccharides. This hypothesis was presented as a YaeQ-encoding fragment from *Salmonella typhimurium* was shown to complement a defect in the RfaH-dependent expression of the *hlyCABD* operon (Wong *et al.*, 1998). However, Vicari & Artsimovitch (2004) have recently presented evidence that *E. coli* YaeQ does not complement the absence of RfaH in *hlyCABD* expression assays *in vitro* or *in vivo*. Thus, the precise function of YaeQ and its orthologues remains unknown. The YaeQ family does not show any sequence similarity with proteins of known structure or function. Determination of the three-dimensional structure of YaeQ may provide insights regarding its function. Several recent examples have demonstrated the feasibility of obtaining functional information from structure (Zarembinski *et al.*, 1998; Bhattacharyya *et al.*, 2002; Christendat *et al.*, 2002; Jackson & Russell, 2001). In this report, we describe the cloning and expression of recombinant *Xac* YaeQ containing selenomethionine and its crystallization and also the determination of an initial

© 2005 International Union of Crystallography
All rights reserved

crystallization communications

electron-density map of the *Xac* YaeQ crystal structure by multi-wavelength anomalous dispersion (MAD).

2. Cloning and expression of YaeQ

The YaeQ gene (XAC2396) was amplified by PCR from *Xac* genomic DNA using the following primers that were designed based on the published *Xac* genome sequence (Da Silva *et al.*, 2002): 5'-CATGC-CATGGCTCATATGGCCCTCACCGCCA-3' and 5'-GGAATT-CAAGCTTTCATTCCGGCCGGGGCT-3'. The PCR product was digested with *Hind*III and *Nde*I and subcloned into the pET-3a vector (Studier *et al.*, 1990) previously digested with the same endonucleases. Selenomethionine-containing YaeQ was expressed in *E. coli* strain BL21(DE3)pLysS (Studier *et al.*, 1990) by growing a 500 ml culture in M9 medium to an optical density (600 nm) of 0.8, at which point 100 mg l⁻¹ lysine, 100 mg l⁻¹ phenylalanine, 100 mg l⁻¹ threonine, 50 mg l⁻¹ isoleucine, 50 mg l⁻¹ valine and 60 mg l⁻¹ selenomethionine were added (adapted from Berne *et al.*, 1999). After 15 min, heterologous protein expression was induced by adding 1 mM isopropyl- β -D-thiogalactopyranoside (IPTG) and the cells were grown for 4 h before harvesting and storage at 203 K.

3. Protein purification

Cells from 1 l culture were resuspended in 25 ml 50 mM Tris-HCl pH 8.0, 25% sucrose, 1 mM EDTA, 1 mM PMSF and lysed using a French press. 2 U ml⁻¹ DNase and 4 mM magnesium chloride were added to the soluble fraction, followed by incubation on ice for 30 min. This mixture was applied onto a Q-Sepharose Fast Flow (FF) Hiload 16/10 column (Amersham Pharmacia) previously equilibrated with 50 mM Tris-HCl pH 7.0, 1 mM EDTA and 14 mM β -mercaptoethanol. Bound proteins were eluted using a 0–300 mM NaCl gradient over 12 column volumes. Fractions containing YaeQ were concentrated using an Amicon system with a 10 kDa membrane and then further purified by gel filtration on a Superdex 75 prep-grade column (Amersham Pharmacia) equilibrated with 5 mM Tris-HCl pH 7.0. The molecular mass of the purified protein, as determined by MALDI-TOF mass spectrometry, was 20 982.749 $\pm$ 5 for the selenomethionine protein and 20 668.582 $\pm$ 5 for the methionine-containing protein. The difference in mass between the two preparations demonstrated that all seven methionines were exchanged for selenomethionines. The total protein yield was 46.5 mg l⁻¹, with approximately 99% purity.

Figure 1
Typical crystals of YaeQ. The approximate dimensions of individual plates are 0.40 $\times$ 0.40 $\times$ 0.03 mm.

Table 1

Recombinant SeMet-labelled *Xac* YaeQ crystal parameters and data-reduction statistics.

Values in parentheses refer to the highest resolution shell.

Space group	$P2_1$		
Unit-cell parameters			
<i>a</i> (Å)	39.75		
<i>b</i> (Å)	91.88		
<i>c</i> (Å)	48.03		
β (°)	108.37		
Data-set statistics	Peak	Inflection	Remote
Resolution range (Å)	37.80–1.90	46.13–2.40	45.64–2.40
	(1.97–1.90)	(2.53–2.40)	(2.53–2.40)
No. of observed reflections	99056 (8105)	49405 (7232)	51465 (7478)
No. of unique reflections	24507 (2207)	12499 (1816)	12469 (1799)
$\langle I/\sigma(I) \rangle$	25.1 (2.1)	11.1 (2.0)	13.0 (2.2)
Multiplicity	4.0 (3.7)	4.0 (4.0)	4.1 (4.2)
Multiplicity, anomalous†	2.1 (1.9)	2.1 (2.0)	2.1 (2.0)
Completeness (%)	95.0 (86.7)	97.5 (96.9)	97.5 (96.7)
Completeness, anomalous† (%)	93.5 (82.8)	94.0 (94.8)	94.4 (95.1)
<i>R</i> (%)	3.2 (34.9)	6.0 (35.2)	5.1 (31.7)
No. of images	200	195	199
λ (Å)	0.97952	0.97962	1.00000

† Each member of a Friedel pair is counted as a separate reflection.

4. Crystallization

YaeQ crystals (Fig. 1) were grown using the hanging-drop vapour-diffusion technique at 291 K. Initial crystallization conditions of the methionine-containing protein were screened by the sparse-matrix sampling approach. Crystals were obtained in several conditions using the Crystal Screen (Hampton Research) and Clear Strategy (Molecular Dimensions Limited) screening kits. Optimization was then pursued by varying the precipitant concentration and the buffer pH. Suitable crystals for diffraction experiments were obtained by mixing equal volumes (1 μ l) of 10 mg ml⁻¹ protein solution (in 5 mM Tris-HCl pH 7.0) with reservoir solution, which consisted of either (i) 30–32% (w/v) PEG 4000, 0.2 M ammonium acetate and 0.1 M Tris-HCl pH 8.5–9.0 or (ii) 30% (w/v) PEG 8000, 0.1 M sodium cacodylate pH 6.7 and 0.2 M sodium acetate. Hanging drops were then equilibrated against 0.4 ml reservoir solutions. Crystals appeared within a few days and grew to mature size within a few weeks. Both methods gave the same crystal form.

5. Data collection and preliminary structure analysis

Initial X-ray diffraction data from the methionine-containing protein were collected at the D03B beamline of the Laboratório Nacional de Luz Síncrotron, Campinas, Brazil using a MAR CCD detector (data not shown). MAD X-ray diffraction data for selenomethionine-containing crystals were collected at the protein crystallography beamline 8.3.1 of the Advanced Light Source (ALS), Berkeley, California using a ADSC CCD detector. SeMet-YaeQ crystals equilibrated against 32% PEG 4000, 0.1 M Tris-HCl pH 9.0, 0.2 M ammonium acetate were flash-frozen in liquid nitrogen and maintained at 100 K in a nitrogen-gas stream during data acquisition (this crystallization solution was sufficient to cryoprotect the crystals). MAD data sets were collected using a single crystal at three wavelengths, 0.97952, 0.97962 and 1.00000 Å, corresponding to the peak, inflection (approximate) and remote points of the fluorescence spectrum, respectively. The peak data set was collected initially and the inflection and remote data sets were then both collected from another region of the same crystal. The exposure time for each image of the peak data set (5 s) was ten times that for the inflection and remote data sets (0.5 s). The oscillation range for each image was 1°. The diffraction patterns from these data sets extended to approxi-

mately 1.9 and 2.4 Å resolution, respectively (Table 1). The crystal belongs to space group $P2_1$, with unit-cell parameters $a = 39.75$, $b = 91.88$, $c = 48.03$ Å, $\beta = 108.37^\circ$. The Matthews coefficient ($V_M = 2.0$ Å³ Da⁻¹ and 38.4% solvent content) suggests that there are two protein molecules per asymmetric unit.

Diffraction images data were indexed, integrated, scaled and merged using the *HKL2000* package (Otwinowski & Minor, 1997) for the peak data set and the programs *MOSFLM* (Leslie, 1992) and *SCALA* (Collaborative Computational Project, Number 4, 1994) for the inflection and remote data sets. The Friedel mates were scaled separately during data processing. Data sets were merged using *CAD* (Collaborative Computational Project, Number 4, 1994). Details of data-acquisition and data-processing statistics are shown in Table 1.

The program *SHELXD* (Sheldrick, 1998) was used to locate 14 selenium sites in the asymmetric unit of the crystal (seven selenomethionines per protein molecule are expected) using the anomalous differences from the peak data set. However, only the coordinates of 12 major selenium sites were used to calculate MAD phases using the program *SHARP* (de La Fortelle & Bricogne, 1997). Experimental phases obtained by *SHARP* were improved by density-modification protocols using the programs *SOLOMON* (Abrahams, 1997) and *DM* (Collaborative Computational Project, Number 4, 1994; Cowtan, 1994). The resulting electron-density map was used by the program *ARP/wARP* (Perrakis *et al.*, 1999) to produce a preliminary poly-alanine model that included 285 of the 364 amino-acid residues found within the asymmetric unit of the crystal. On the basis of this preliminary model, the non-crystallographic symmetry relating the two monomers was determined and used for phase improvement and model building. Interpretation of electron-density maps and construction of missing residues is being performed using the program *O* (Jones *et al.*, 1991). Structural refinement of the XAC2396 atomic model is currently under way using *REFMAC* (Murshudov *et al.*, 1997) and *CNS* (Brünger *et al.*, 1998).

We would like to thank the ALS 8.3.1 beamline staff and Dr Mark Glover (University of Alberta) and members of his laboratory for the opportunity to collect MAD data at the ALS. We thank the laboratory of Dr Paolo di Mascio (Instituto de Química, Universidade de São Paulo) for MALDI-TOF MS analysis. This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) through grant 01/07534-3, Conselho Nacional de Desen-

volvimento Científico e Tecnológico (CNPq) (Brazil) and Associação Brasileira de Tecnologia de Luz Síncrotron (ABTLuS). CRG is a FAPESP fellow and LMPG was supported with a PhD studentship from FCT/MCT (Portugal).

References

- Abrahams, J. P. (1997). *Acta Cryst.* **D53**, 371–376.
- Berne, P. F., Doublé, S. & Carter, C. W. Jr (1999). *Crystallization of Nucleic Acids and Proteins: A Practical Approach*, 2nd ed., edited by A. Ducruix & R. Giegé, pp. 45–73. Oxford University Press.
- Bhattacharyya, S., Habibi-Nazhad, B., Amegbey, G., Slupsky, C. M., Yee, A., Arrowsmith, C. & Wishart, D. S. (2002). *Biochemistry*, **41**, 4760–4770.
- Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* **D54**, 905–921.
- Christendat, D., Saridakis, V., Kim, Y., Kumar, P. A., Xu, X., Semesi, A., Joachimiak, A., Arrowsmith, C. H. & Edwards, A. M. (2002). *Protein Sci.* **11**, 1409–1414.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Cowtan, K. (1994). *Int. CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **31**, 34–38.
- Da Silva, A. C. R. *et al.* (2002). *Nature (London)*, **417**, 459–463.
- Galvão-Botton, L. M. P., Katsuyama, A. M., Guzzo, C. R., Almeida, F. C. L., Farah, C. S. & Valente, A. P. (2003). *FEBS Lett.* **552**, 207–213.
- Jackson, R. M. & Russell, R. B. (2001). *Comput. Chem.* **26**, 31–39.
- Jones, T. A., Zhou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* **A47**, 110–119.
- La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **244**, 472–494.
- Leslie, A. G. W. (1992). *Int. CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **26**.
- Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* **D53**, 240–255.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Perrakis, A., Morris, R. M. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
- Sheldrick, G. M. (1998). *Direct Methods for Solving Macromolecular Structures*, edited by S. Fortier, pp. 401–411. Dordrecht: Kluwer Academic Publishers.
- Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). *Methods Enzymol.* **185**, 60–89.
- Vicari, D. & Artimovitch, I. (2004). *Mol. Genet. Genomics*, **272**, 489–496.
- Wong, K. R., Hughes, C. & Koronakis, V. (1998). *Mol. Gen. Genet.* **257**, 693–696.
- Zarembinski, T. I., Hung, L. W., Mueller-Dieckmann, H. J., Kim, K. K., Yokota, H., Kim, R. & Kim, S.-H. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 15189–15193.

4.3.2 Incorporação de selênio metionina na proteína YaeQ

A produção de YaeQ_{XAC2396} recombinante contendo selênio metionina (YaeQ_{XAC2396}-SelMet) era essencial para a determinação da estrutura de YaeQ_{XAC2396} pela técnica MAD. Tentativas anteriores de preparar derivados (através de técnica de *soaking*) com mais de 20 diferentes sais de átomos pesados e resolver as fases pela técnica de MIR (Multiple Isomorphous Replacement) utilizando dados coletados na linha MX1 do LNLS não tiveram sucesso. Na época a linha MX2 no LNLS não estava disponível e a linha MX1, que possui um comprimento de onda fixo de aproximadamente 1,45Å, não é configurada para coletas de dados em mais de um comprimento de onda.

A preparação de YaeQ_{XAC2396}-SelMet está descrita nos Materiais e Métodos. Verificamos a incorporação de selênio-metionina na proteína recombinante através de espectrometria de massa (**Figura 25**). A massa molecular de YaeQ_{XAC2396} medida por espectrometria de massa foi de 20661g/mol, muito próximo da massa molecular teórica que é de 20635.2g/mol ou 20663g/mol com a primeira metionina formilada. A massa molecular de YaeQ_{XAC2396}-SelMet medida por espectrometria de massa foi de 20987.8g/mol, muito próxima da massa molecular teórica para YaeQ_{XAC2396}-SelMet, que é de 20992 g/mol. [Observa que: YaeQ possui 7 metioninas. Se todas foram trocadas para selênio metioninas a diferença de massa esperada é de: 7×79.0 (os enxofres) – 7×32.0 (selênios) = $553 - 224 = 329$ g/mol]. A coleta de dados MAD dos cristais de YaeQ_{XAC2396}-SelMet foi feita na linha “Crystallography Beam Line 8.3.1” no Advanced Light Source (ALS), Berkeley, Califórnia.

Figura 25: Espectro de massa da proteína YaeQ.

m/z: Relação massa (m) carga (z) da proteína. **a)** sem selênio-metionina e **b)** com selênio-metionina. M^{+1} e $M^{+1/2}$, carga 1 e carga 2, respectivamente. M^{+1} corresponde a massa de YaeQ dividida por 1 (carga 1), e $M^{+1/2}$ corresponde a massa de YaeQ dividida por dois (duas cargas).

4.3.3 Coleta de dados de difração de raios-X e determinação das fases

Escolha dos comprimentos de onda para a coleta de dados de difração de raios-X. Os dados de difração de raios-X foram coletados em três comprimentos de onda e devidamente processados conforme descrito nos Materiais e Métodos. Estes comprimentos de onda foram selecionados com o objetivo de otimizar o sinal anômalo - a diferença de intensidade entre os pares de Bijvoet. A determinação do melhor comprimento de onda para a coleta de dados de difração foram determinados com base no espectro de excitação de raios-X e medindo a emissão de fluorescência pelos átomos de selênio de YaeQ_{XAC2396} no cristal. Este espectro de fluorescência, **Figura 26**, é característico para átomos de selênio – uma banda intensa de absorção em volta de 0.97 Å (12657 eV).

A partir do espectro de fluorescência foi possível determinar os componentes do espalhamento anômalo - f' e f'' – em função do comprimento de onda de excitação. O objetivo é coletar os dados de difração de raios-X nos comprimentos de onda onde o valor de f'' (parte imaginária) tem o maior valor positivo, (“peak” = 12657.88 eV), e o valor de f' (parte real) tem o maior valor negativo, (“inflection” = 12655.09 eV) (**Figura 26**). Desta forma, foram coletados três conjuntos de dados em três comprimentos de

onda : $\lambda_1 = 0,97962 \text{ \AA}$ (12656.4 eV) que representa a média do comprimento de onda do “peak” e do “inflection”, $\lambda_2 = 0,979524 \text{ \AA}$ (12657.88eV) que representa o “peak” e $\lambda_3 = 1,00000 \text{ \AA}$ (12398.4 eV) que representa o “remote”. O comprimento de onda do “remote” contém pouca contribuição de sinal anômalo dos átomos de selênio.

Uma descrição do processamento dos dados está mostrado na **Tabela 1** do **Anexo 3**, em que os conjuntos de dados do “inflection” (λ_1) e do “remote” (λ_3) foram processados utilizando o programa Mosflm⁶⁶ e foram escalonados pelo programa Scala do Kit ccp4i⁶⁷. O conjunto de dados do “peak” (λ_2) foi processado pelo programa HKL2000⁶⁵.

Figura 26: Espectro de emissão de fluorescência do cristal da proteína YaeQ_{XAC2396} com selênio-metionina.

O valor mínimo na curva de $f'(\lambda_1)$, curva de baixo na figura, corresponde ao ponto de inflexão na curva f'' , curva de cima na figura.

Localização das posições dos átomos de Selênio no cristal de YaeQ_{XAC2396}.

Após o processamento das imagens, as posições dos selênios na proteína foram inicialmente calculadas pelo programa Shelxd⁶⁹. A **Figura 27** mostra como varia o coeficiente de correlação das soluções em relação ao número de rodadas (ciclos). Apenas os ciclos com alto coeficiente de correlação contêm as posições dos átomos pesados verdadeiras.

A curva que correlaciona o número de respostas em função do coeficiente de correlação (**Figura 27**) deveria conter um salto no valor do coeficiente de correlação entre o conjunto de respostas erradas com o grupo de respostas certas (coeficientes de correlação maiores). A partir deste salto, conseguiríamos distinguir quais respostas contêm as posições verdadeiras dos metais pesados na célula unitária do cristal de $\text{YaeQ}_{\text{XAC2396}}$. O conjunto de dados que tem um maior salto é referente ao conjunto de dados do “peak” no comprimento de onda λ_1 , enquanto que os dados do remote, que não contêm sinal anômalo significativo, não apresenta nenhum salto.

Figura 27: Gráfico que relaciona o coeficiente de correlação com o número de ciclos para localizar os sítios dos átomos pesados, programa SHELXD⁶⁹.

Cálculo das fases e a determinação da estrutura de $\text{YaeQ}_{\text{XAC2396}}$. As posições dos selênios na proteína para cada conjunto de dados foram distintos e as posições que forneceram as melhores fases, utilizando o programa SHARP⁷⁰, foram às posições dos selênios calculadas a partir dos dados processados do peak, o que já era esperado, por este conter a máxima dispersão anômala.

Para cada proteína se esperaria 7 posições de selênio, já que no espectro de massa verificou-se a incorporação de 7 selênios-metioninas (**Figura 28**). Como há duas moléculas de $\text{YaeQ}_{\text{XAC2396}}$ na unidade assimétrica esperaríamos 14 átomos de selênio por unidade assimétrica. Encontramos apenas 12 sítios de selênio. Este resultado sugere que há uma metionina em cada molécula de $\text{YaeQ}_{\text{XAC2396}}$ -SelMet móvel o suficiente que nós não conseguimos detectá-las.

As posições dos átomos de selênio refinadas e seus valores de ocupância e fator B, calculados pelo programa SHARP, estão mostrados na **Tabela 15**. Utilizando estas

posições dos selênios calculamos as fases e conseqüentemente o mapa de densidade eletrônica de YaeQ_{XAC2396}-SelMet, utilizando o programa SHARP. Os dados estatísticos do refinamento das fases para gerar o primeiro mapa de densidade eletrônica estão mostrados na **Tabela 16**.

O primeiro mapa da densidade eletrônica mostrou claramente regiões de densidade eletrônica correspondentes a cadeia polipeptídica principal e ao solvente. Desta forma, foi possível construir um modelo inicial da proteína YaeQ_{XAC2396}-SelMet usando o programa ARP/WARP⁷¹, que construiu 285 resíduos (de 364 resíduos totais), distribuídos em 14 cadeias.

Tabela 15: As posições dos selênios refinadas pelo programa SHARP.

As posições dos selênios foram refinadas a partir das posições iniciais calculadas pelo sinal anômalo dos dados processados do Peak pelo programa SHELXD, e verificadas posteriormente no mapa de densidade eletrônica. Está descrita na tabela os valores da ocupância, fator B e as posições em coordenadas fracionadas de cada átomo de selênio na unidade assimétrica do cristal e YaeQ_{XAC2396}-SelMet.

Ocupância	Fator B (Å)	Posições		
		X	Y	Z
2,13	35,84	0,65	0,36	0,52
1,01	32,31	0,61	0,39	0,53
2,01	44,11	0,81	0,36	0,67
1,20	37,66	1,08	0,18	0,73
1,85	46,57	1,11	0,18	0,46
1,47	40,01	1,08	0,39	0,92
1,45	52,79	1,65	0,09	0,86
1,61	53,33	1,65	0,10	0,79
0,93	39,31	1,10	0,04	0,69
1,35	65,37	1,49	0,08	1,33
1,37	38,67	1,43	-0,05	1,11
0,70	24,37	1,44	0,02	0,76

Tabela 16: Dados estatísticos do refinamento das fases do cristal de YaeQ_{XAC2396}-SelMet. Parâmetros relevantes no cálculo do primeiro mapa de densidade eletrônica de YaeQ_{XAC2396}-SelMet. A quantidade de solvente na célula unitária corresponde a 34,6%, em que está mostrado na tabela os valores de FOM (*Figure of Merit*, ou $|F_{p_{best}}/F_{p_{obs}}|$) dos dados no eixo de simetria (*centric*) e fora do eixo de simetria (*acentric*). Está mostrado também os valores do poder de faseamento (*Phasing Power*, ou F_H/σ) calculados segundo o método isomorfo e anômalo.

Conjunto de Dados	FOM (centric/ acentric)	Phasing Power			
		Isomorfo		Anômalo	
		Acentric	Centric	Acentric	Centric
Peak		0	0	2,072	0
Inflection	0,31092 / 0,43243	0,179	0,199	1,009	0
Remote		1,191	1,280	0,156	0

4.3.4 Estrutura Tridimensional de YaeQ

Estrutura de YaeQ. A topologia de YaeQ_{XAC2396} está mostrada na **Figura 29** e na **Figura suplementar 1A** do **Anexo 4**. A estrutura é composta por uma hélice central ($\alpha 1$) envolvida por duas folhas betas ($\beta 3$, $\beta 4$, $\beta 5$, $\beta 6$, $\beta 7$, $\beta 11$ e $\beta 1$, $\beta 2$, $\beta 8$, $\beta 9$, $\beta 10$) que se conectam pelas fitas $\beta 10$ e $\beta 11$, assim como pela fita $\beta 7$, formando um barril quase perfeito (**Figura 21** e **Figuras suplementares 1B** do **Anexo 4**). A abertura neste barril ocorre entre as fitas $\beta 2$ e $\beta 3$ em que os aminoácidos R39, R43 e F47, pertencentes a uma face da hélice central, interagem com moléculas do solvente (**Figura suplementar 2A** do **Anexo 4**).

As três fitas paralelas $\beta 5$, $\beta 6$ e $\beta 7$ ligadas a duas hélices externas ao barril de folhas betas, $\alpha 2$ e $\alpha 3$, tem topologia β - α - β - α - β . Esta topologia é similar ao do motivo de Rossmann-fold, envolvida na ligação a dinucleotídeos ou nucleotídeos. No entanto, YaeQ_{XAC2396} não possui a sequência consenso G-x-G-x-x-G do motivo de Rossmann-fold envolvida na ligação a NAD(P). As duas primeiras glicinas do motivo consenso, em geral, possuem ângulos diedrais permitidos apenas para resíduos de glicina e estão envolvidos na interação com os pirofosfatos dos dinucleotídeos¹⁹⁴.

Há duas moléculas (A e B) de YaeQ_{XAC2396} na unidade assimétrica. A sobreposição das cadeias A e B revela um RMSD igual a 0.82 Å para os átomos da cadeia principal. Os resíduos presentes na parte de cima da estrutura de YaeQ_{XAC2396} contêm os maiores valores de fator B (coloridos de azul para vermelho na **Figura 29**). Estes resíduos estão localizados no amino-terminal da molécula, no loop entre as fitas $\beta 3$ e $\beta 4$ e no amino-terminal da $\alpha 3$. Os resíduos 63 e 64 da cadeia A e os resíduos 61,

62 e 64 da cadeia B, localizados no loop β 3- β 4 possuem densidade eletrônica pobre, os três primeiros aminoácidos das cadeias A e B e os dois últimos aminoácidos da cadeia B estão desestruturados e não apresentam densidade eletrônica no mapa.

A superfície enterrada entre as duas moléculas é de $648 \text{ \AA}^2$, sugerindo que este dímero não é biologicamente relevante¹³⁹. YaeQ_{XAC2396} em tampão PBS elui com volume de 67,5 mL (**Figura 30**). A curva de calibração desta coluna é dada pela equação $\log MW = -1,3014V_e/45,7 + 3,2624$, que correlaciona a massa da proteína (MW) com o volume de eluição da proteína (V_e). YaeQ_{XAC2396} elui com massa aparente de 21,5 KDa muito próxima da massa molecular teórica de 20,6 KDa. Desta forma, YaeQ deve ser monomérica.

Figura 28: Estrutura de YaeQ colorida segundo o valor do fator-B

As regiões coloridas em vermelho são os de maior valor de fator-B e as coloridas em azul são os de menor valor. A maioria dos resíduos mais móveis estão localizados na parte de cima da estrutura de YaeQ_{XAC2396}.

Figura 29: Cromatograma da eluição de YaeQ_{XAC2396} na coluna de gel filtração 75 .

As setas indicam o volume de eluição de diferentes proteínas com massa molecular diferentes (KDa). Estas proteínas foram utilizadas para o cálculo da curva de calibração desta coluna de gel filtração. Condições: 8mM dihidrogênio fosfato de sódio, 2,3mM monohidrogênio fosfato de sódio e 130mM cloreto de sódio.

4.3.5 Resíduos Altamente Conservados na Família de YaeQ

A família de YaeQ é altamente conservada. O grau de conservação em sequência primária entre ortólogos pertencentes à família de YaeQ_{XAC2396} está mostrado na **Figura 2** do **Anexo 4**. O alinhamento da sequência primária de membros da família de YaeQ revela que alguns ortólogos possuem inserções de aminoácidos entre estruturas secundárias. A inserção mais acentuada ocorre no loop entre a alfa hélice $\alpha 2$ e a fita beta $\beta 5$ (resíduos 54-55 de YaeQ_{XAC2396}).

Os resíduos conservados na família de YaeQ localizam em uma única face da estrutura de YaeQ. Os aminoácidos absolutamente conservados nos ortólogos de YaeQ são: S₃₄, E₃₅, R₄₃, W₈₁ e G₈₅. Estes aminoácidos estão localizados em uma face da estrutura de YaeQ_{XAC2396} (**Figura suplementar 2B** do **Anexo 4**). Três dos cinco

aminoácidos absolutamente conservados (S_{34} , E_{35} , R_{43}) formam uma rede de interações como ilustrado na **Figura complementar 2C do Anexo 4**. Os nitrogênios da cadeia lateral e da cadeia principal da H_{32} interagem com o hidrogênio do grupo hidroxil da S_{34} e do grupo carboxi do E_{35} , respectivamente. O grupo carboxi do E_{35} faz mais três pontes de hidrogênio, uma com a cadeia lateral da R_{39} e as outras duas com a cadeia lateral da R_{43} . A R_{43} faz mais três pontes de hidrogênio sendo uma delas com o grupo carboxi da cadeia principal da A_{30} e as outras duas pontes de hidrogênio com duas moléculas de água. Além disso, a R_{43} faz interação eletrostática com o anel da F_{57} . A G_{85} está localizada no loop entre $\beta 5$ - $\alpha 2$ e possui ângulos diedrais apenas permitidos para resíduos de glicina, sugerindo que este resíduo seja importante para a estabilidade da estrutura de $YaeQ_{XAC2396}$. O W_{81} se localiza na fita $\beta 5$ em um ambiente hidrofóbico em que o W_{81} interage com o anel da P_{64} e com o aminoácido A_{95} .

A família de YaeQ possui 4 motivos conservados. A família de YaeQ contém 4 motivos de sequência conservados (**Figura complementar 2D-F do Anexo 4**): 1- *Motivo AXHPSE* (resíduos de 30-35) localizado no loop $\beta 2$ - $\alpha 1$; 2- *Motivo ácido DEP(D/E)* (resíduos 65-68) localizado no loop $\beta 3$ - $\beta 4$ onde há alta mobilidade como pode ser visto na **Figura 29**; 3- G_{85} localizado no loop $\beta 5$ - $\alpha 2$ e 4- R_{148} , altamente conservada, pertencente ao loop $\alpha 4$ - $\beta 8$.

4.3.6 YaeQ Pertence à Família PD-(D/E)XK de Nucleases Dependentes de Magnésio

A comparação da estrutura de $YaeQ_{XAC2396}$ com outras depositadas em bancos de dados públicos (**Tabela 1 complementar do Anexo 4**) mostrou que $YaeQ_{XAC2396}$ pertence a uma variação do motivo PD-(D/E)XK encontrado em endonucleases como as enzimas de restrição *SdaI*, *FokI*, *BglI*, *Ecl18kI*, *BamHI*, *NgoMIV*, *BstYI*, em membros da família de endonuclease XPF (Xeroderma Pigmentosa) envolvida em reparo de DNA e recombinação que possui o sítio catalítico composto pelos resíduos GD-ERKX₃D¹⁹⁵, em outras proteínas envolvidas no processo de reparo de DNA como MthH¹⁹⁶, nas Holiday junction resolvases¹⁹⁷ e RecB¹⁹⁸. YaeQ também possui baixa similaridade estrutural com metiltransferases putativas e oxidoreductases (**Tabela 1 complementar do Anexo 4**).

Membros da família PD-(D/E)XK possuem baixa ou nenhuma similaridade em sequência primária, mas possuem o sítio ativo (os aminoácidos PD-(D/E)X₁₄K) espacialmente conservados¹⁹⁹. Os dois aminoácidos negativos deste motivo estão

envolvidos na ligação aos cátions (na maioria dos casos são cátions de magnésio) que também interagem com o DNA. A lisina (não muito bem conservada) possivelmente está envolvida na ativação da água nucleofílica que ataca o fosfato a ser clivado¹⁹⁹.

O sítio catalítico putativo de *YaeQ*_{XAC2396} é *P*₆₇(D/E)₆₈-*W*₈₁*X*(D/E)*83XG85XP*₈₇*X*₃*R*₉₁. A **Figura 30 a e c** mostra a superposição do sítio ativo das enzimas de restrição do tipo II, *Sdal*, *NgoMIV*, da proteína hipotética *Tt1808* com a estrutura de *YaeQ*_{XAC2396}. Assim como *Sdal*, *YaeQ*_{XAC2396} não possui a lisina do motivo PD-(D/E)*X*₁₄K. No caso de *Sdal*, Tamulaitiene e colaboradores (2006) propôs que o papel da lisina foi transferida possivelmente para Q₁₇₈, pois mutações na Q178 acaba com a sua atividade endonuclease²⁰⁰.

Na posição da lisina no PD-(D/E)XK, *YaeQ*_{XAC2396} possui a G₈₅, absolutamente conservada na família de *YaeQ*_{XAC2396} (**Figura 2 do Anexo 4**). A G₈₅ faz parte do loop β₅ - α₂ que contem a P₈₇ também altamente conservada. A comparação deste loop com os correspondentes loops nas estruturas de *Sdal*, *NgoMIV*, *Tt1808*, *BamHI* e *Hef* revelou que a família de *YaeQ* é aquela que possui o loop de menor tamanho (**Figura 30 b**). Desta forma, provavelmente a G₈₅ e a P₈₇ estão envolvidas na estruturação deste loop, β₅-α₂, que conecta vários aminoácidos do sítio ativo putativo de *YaeQ*_{XAC2396}, *P*₆₇(D/E)₆₈-*W*₈₁*X*(D/E)*83XG85XP*₈₇*X*₃*R*₉₁.

O sítio catalítico putativo de *YaeQ*_{XAC2396} está mostrado na **Figura 3A do Anexo 4**. O D₆₈ interage com o D₈₃, com o E₃₅, absolutamente conservado na família de *YaeQ*, e com duas moléculas de água (água 75 e água 137). O D₈₃ interage com o aminoácido absolutamente conservado W₈₁, assim como com o aminoácido altamente conservado R₉₁ e a água 75. O E₃₅ interage também com duas moléculas de água (água 104 e água 139).

Figura 30: Superposição estrutural de YaeQ com outras proteínas estruturalmente similares¹⁹⁵.

a) Superposição do sítio ativo de *NgoMIV* (PDBID: 1FIU), *SdaI* (PDBID: 1IXS), *Tt1801* (PDBID: 1WDJ) com *YaeQ*_{XAC2396} (PDBID: 2G3W); **b)** sobreposição dos loops que conectam os aminoácidos do sítio ativo PD-(D/E)XK e **c)** Alinhamento do sítio ativo de *SdaI*, *NgoMIV*, *Tt1801*, *Hef* (PDBID: 1J22), *BamHI* (PDBID: 1BAM) com *YaeQ*_{XAC2396}. *YaeQ*_{XAC2396} está em vermelho, *SdaI* em rosa claro, *NgoMIV* está em ciano, *Tt1801* está em azul, *Hef* está em preto e *BamHI* em laranja. O ponto vermelho se refere ao E₈₉ de *YaeQ*_{XAC2396} e o ponto em preto se refere ao D₅₉ de *Hef* pertencente a família XPF²⁰¹.

Anexo 4 – A Estrutura de YaeQ de *Xanthomonas axonopodis* pv. citri revela um novo padrão de enovelamento construído ao redor da variação do motivo PD-(D/E)XK de endonucleases.

- Structure of *Xanthomonas axonopodis* pv. Citri YaeQ reveals a new compact protein fold built around a variation of the PS-(D/E)XK nuclease motif -

Structure of *Xanthomonas axonopodis* pv. citri YaeQ reveals a new compact protein fold built around a variation of the PD-(D/E)XK nuclease motif

Cristiane R. Guzzo,¹ Ronaldo A. P. Nagem,² João A. R. G. Barbosa,³ and Chuck S. Farah^{1*}

¹ Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, Avenida Prof. Lineu Prestes 748, São Paulo, SP, CEP 05508-000, Brazil

² Departamento de Bioquímica e Imunologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Avenida Antônio Carlos 6627, Belo Horizonte, MG, CEP 31270-901, Brazil

³ Centro de Biologia Molecular Estrutural, Laboratório Nacional de Luz Síncrotron, C.P. 6192, Campinas, SP, CEP 13084-971, Brazil

ABSTRACT

The YaeQ family of proteins are found in many Gram-negative and a few Gram-positive bacteria. We have determined the first structure of a member of the YaeQ family by X-ray crystallography. Comparisons with other structures indicate that YaeQ represents a new compact protein fold built around a variation of the PD-(D/E)XK nuclease motif found in type II endonucleases and enzymes involved in DNA replication, repair, and recombination. We show that catalytically important residues in the PD-(D/E)XK nuclease superfamily are spatially conserved in YaeQ and other highly conserved YaeQ residues may be poised to interact with nucleic acid structures.

Proteins 2007; 69:644–651.
© 2007 Wiley-Liss, Inc.

Key words: YaeQ PD-(D/E)XK nuclease motif; PD-(D/E)XK superfamily; *Xanthomonas axonopodis* pv. citri; crystal structure; protein topology; protein structure; X-ray diffraction.

INTRODUCTION

Homologous proteins are grouped into families based on the principle that related proteins adopt similar structures and are, in turn, functionally related. Indeed, protein function is much more strongly dependent on three-dimensional structures than on similarities in primary structure¹ and structural comparisons may reveal evolutionary relationships between two or more protein families with low or no recognizable similarities.² One good example of this is provided by the type II restriction endonucleases and other nucleic acid cleaving enzymes that form a large and extremely diverse protein superfamily that display little sequence and low three-dimensional structure similarities. However, they all retain a common core fold with a PD-(D/E)XK motif responsible for cleavage,³ even though some members of this family lack nuclease function. Diversity within the PD-(D/E)XK superfamily is so great that the identification of new members is often only possible as three-dimensional structure information becomes available.⁴

Xanthomonas axonopodis pv. citri (Xac) YaeQ (XAC2396) is a member of a family of bacterial proteins conserved in proteobacteria. No YaeQ homologs have been identified in eukaryotes or archaeal species. Within bacteria, YaeQ homologs are restricted to 130 of the 343 proteobacteria whose genomes have been sequenced and deposited in the public databases as well as seven others, six of which are gram positive actinobacter species with high G+C contents. Xac YaeQ was selected for structural analysis because of its well-folded and non-aggregated state in solution, both in bacterial lysates and in the purified form.⁵

It has been suggested that YaeQ is involved in the regulation of transcription of genes encoding virulence factors such as hemolysin and enzymes involved in the production of lipopolysaccharides. This hypothesis was based on the observation that a YaeQ-encoding fragment from *Salmonella typhimurium* complements a defect in the RfaH-dependent expression of the *hlyCABD* operon.⁶ However, Vicari and Artsimovitch⁷ have recently presented evidence that *E. coli* YaeQ does not complement the absence of RfaH in *in vitro* or *in vivo* *hlyCABD* expression assays. Thus, the precise function of YaeQ and its orthologs remains unknown.

The Supplementary Material referred to in this article can be found online at <http://www.interscience.wiley.com/jpages/0887-3585/suppmat>.

Grant sponsors: Fundação de Amparo à Pesquisa do Estado de São Paulo, Conselho Nacional de Pesquisa e Tecnologia (Brazil).
*Correspondence to: Chuck S. Farah, Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, Av. Prof. Lineu Prestes 748, Cidade Universitária, São Paulo, SP, Brazil 05508-000. E-mail: chsfarah@iq.usp.br

Received 21 February 2007; Revised 19 March 2007; Accepted 27 March 2007

Published online 10 July 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21556

Here, we report the first high resolution structure of a member of the YaeQ protein family. The structure of recombinant *Xac* YaeQ containing selenomethionine was determined to a resolution of 1.9 Å by multiple-wavelength anomalous dispersion (MAD). Comparison of its structure with that of other proteins deposited in the public databases indicates that YaeQ represents a new compact protein fold that contains within it a conserved three-dimensional motif that corresponds to the active site of the PD-(D/E)XK superfamily of proteins. The mapping of other well conserved residues within the YaeQ family of proteins on the *Xac* YaeQ structure allows us to identify other sites of possible functional importance.

MATERIALS AND METHODS

Structure determination and refinement

Recombinant YaeQ was expressed in minimal media containing selenomethionine, purified and crystallized as described.^{8,9} Anomalous scattering from Se atoms from a recombinant SeMet *Xac* YaeQ crystal⁸ was used for phase determination by MAD. The peak dataset was processed to 1.8 Å. Data-reduction statistics of this dataset to 1.9 Å (highest resolution shell $I/\sigma(I) = 2.1$, multiplicity = 3.7, completeness = 86.7%, $R = 34.9\%$) have been published elsewhere.⁸ The program *SHELXD*¹⁰ was used to obtain initial estimates of 12 of the 14 selenium sites present in the two copies of the protein in the asymmetric unit. Se positions were subsequently refined and phases were calculated up to 1.8 Å resolution using the program *SHARP*.¹¹ Phases were refined by density modification using the programs *SOLOMON*¹² and *DM*.^{13,14} The overall figure of merit (FOM) for centric and acentric reflections were 0.31 and 0.43, respectively.

Polypeptide segments could be traced in the resulting electron-density map and a preliminary hybrid model (amino acids and dummy atoms) was generated by ARP/wARP¹⁵ that included 285 of the 364 residues distributed over 14 peptide segments in the two molecules found within the asymmetric unit. The data collected at the peak wavelength (high resolution dataset⁸) were used for refinement of the model. Using non-crystallographic symmetry, we were able to trace almost all of the remaining residues in the asymmetric unit except for N- and C-terminal segments and an internal loop (residues 60–65) of both chains. The position of the Se sites was confirmed by analysis of the anomalous difference map using the peak data. Structural refinement of the YaeQ atomic model done using ARP/wARP,¹⁵ REFMAC,¹⁶ CNS,¹⁷ and COOT.¹⁸ TLS was used in the final cycles of refinement with each chain in the asymmetric unit assigned as an independent group. Water molecules were added with ARP/wARP, REFMAC, and directly by COOT. The final R_{work} and R_{free} values are 0.192 and 0.247, respectively.

Table I
Refinement Statistics

Refinement (30.0–1.8 Å)	
No. of molecules/A.U.	2
Residues/chain	182
Water	218
R_{factor}^a	0.192
R_{free}^a	0.247
R.m.s.d.	
Bond lengths (Å)	0.015
Bond angles (°)	1.481
% of residues in Ramachandran regions:	
Most favored	95.1
Additional allowed	4.9
Generously allowed	0.0
Disallowed	0.0

^a $R_{\text{factor}} = \sum |F_o - F_c| / \sum F_o$. 5% of the data were excluded for the R_{free} calculation.

The refinement data statistics are shown in Table I. [At the time of submission of this manuscript, we became aware of a report of the crystallization of a YaeQ ortholog from *Xanthomonas campestris*.¹⁹]

Accession numbers

The YaeQ coordinates and structure factors have been deposited in the Protein Data Bank under accession code 2G3W.

RESULTS

YaeQ structure and fold

Xac YaeQ is the first member of the YaeQ family for which the structure has been determined. The YaeQ structure consists of a central α -helix ($\alpha 1$) surrounded by two curved β -sheets [Fig. 1(A) and Fig. 1 of Supplementary Online Material]. One β -sheet is composed of $\beta 2$, $\beta 1$, $\beta 8$, $\beta 9$, and $\beta 10$ and the other is composed of $\beta 3$, $\beta 4$, $\beta 5$, $\beta 6$, $\beta 7$, and $\beta 11$. These β -sheets form an almost complete barrel in which only part of one face of the central α -helix is exposed to the solvent [Fig. 2(A) of Supplementary Online Material]. Three adjacent parallel strands ($\beta 5$, $\beta 6$, $\beta 7$) are linked to two external α -helices ($\alpha 2$, $\alpha 3$) to form a beta-alpha-beta-alpha-beta motif in which α -helix $\alpha 3$ is kinked significantly at A120 [Fig. 1(A)].

The first three residues of chain A and B and the last two residues of chain B are not visible in the electron density map and the C α atoms of the two chains superpose with a RMSD of 0.82 Å. Residues with the greatest temperature factors are found on the top of YaeQ barrel at the N-terminal of $\alpha 3$ and in two loops located between $\beta 2$ and $\alpha 1$ and between $\beta 3$ and $\beta 4$. Furthermore, electron density for specific residues within the

Figure 1

YaeQ structure and similarity to PD-(D/E)XK superfamily proteins. (A) Front view of ribbon representation of the YaeQ structure. Cartoons show α -helices (red), β -strands (yellow), and coil regions and turns (green) using chain B of the asymmetric unit. (B) Structural superposition of YaeQ residues 81–95 (red) with residues 246–265 of SdaI (pink, 1IXS), 183–205 of NgoMIV (cyan, 1FIU), 114–134 of Td808 (yellow, 1WDJ), 46–70 of the archaeal Hef XPF/Rad1/Mus81 family nuclease (black, 1J22), and 109–130 of BamHI (orange, 1BAM). The side chains W81, D83, G85, P87, and R91 of YaeQ are indicated as red sticks. The small red and black circles indicate the Ca position of YaeQ E89 and Hef nuclease D54, respectively. Note that YaeQ E89 does not superpose on the conserved Asp residue in the last position of the GDX₂ERKX₂D motif in the structures of the XPF/Rad1/Mus81 family of proteins (for example residue D54 of the Hef nuclease²⁰). (C) Alignment of the PD-(D/E)XK motifs in the six PD-(D/E)XK family proteins shown in part (B). Because of the high level of conservation of W81, (D/E)83, G85, P87, and R91 in the YaeQ family of proteins this motif may be extended to PD-WX(D/E)XGXPX₂R (see text). Figures were produced with PyMol.²¹

loop between $\beta 3$ and $\beta 4$ (residues 63 and 64 of chain A and 61, 62, and 64 of chain B) is poorly defined. In the presence of 130 mM NaCl, YaeQ elutes as a monomer in gel filtration chromatography (data not shown). Therefore, the contacts observed between the two YaeQ molecules in the asymmetric unit are probably not physiologically relevant. Furthermore, analysis of the protein–protein contacts within the crystal lattice using the PISA server²² failed to identify significant interfaces: the largest contact surface buries 648 Å² and all possible interfaces have a complexation significance score of zero.

Well conserved residues

At the moment, 130 YaeQ homologs have been identified in bacteria, all but seven of which are gram negative proteobacteria from the α , β , δ , and γ subdivisions. Of the remaining, six YaeQ proteins have been identified in gram positive bacteria with high G+C content. No YaeQ homologs have been identified in archaea or eukaryotes. A profile of the degree of conservation of residues within the YaeQ family is presented in Figure 2 based on an

alignment of 121 full-length YaeQ orthologs (including all 82 members of Pfam family PF07152). The numbering scheme in this figure corresponds to that of the Xac YaeQ protein and indicates that residues within the N-terminal half of the YaeQ family are more conserved than the C-terminal half. The alignment also shows that some homologs have insertions in loops between secondary structure elements, the most significant of which occurs before the short strand $\beta 3$ (between residues 54 and 55; Fig. 2). Small insertions are also found in the loops between $\beta 3$ and $\beta 4$, $\beta 4$ and $\beta 5$, $\alpha 2$ and $\beta 6$, and $\beta 8$ and $\beta 9$ (Fig. 2).

Only five residues are absolutely conserved in 119 of the 121 full-length YaeQ orthologs used to produce the alignment profile in Figure 2. The locations of these residues—S34, E35, R43, W81, and G85—are shown and described in Figure 2(B) of Supplementary Online Material. The top face of the YaeQ structure [as displayed in Fig. 2(E) of Supplementary Online Material] is made up of several loops or linkers between secondary structure elements. Four of these structures deserve special attention since they contain motifs well-conserved within the

PD-(D/E)XXK Nuclease Motif in YaeQ Structure

Figure 2

Multiple sequence representation of the YaeQ family. Multiple sequence representation of 121 full-length YaeQ orthologs including 82 derived from the Pfam family PF07152. The height of the residue symbol indicates the degree of conservation. The YaeQ sequence and numbering scheme is shown below the alignment. Secondary structure elements are colored red (α -helix) and blue (β -strands). The figure was generated using the WebLogo server (<http://weblogo.berkeley.edu/>).

YaeQ family: (1) Residues 30–35 are made up of an AXHPSE motif that is well defined in the structure. (2) An acidic DEP(D/E) motif (residues 65–68) is also well defined and is in fact part of a larger loop that extends from residues 58 to 68. As mentioned above, residues 61–64 of this loop are poorly defined in both chains of the asymmetric unit and were modeled with different conformations which should only be interpreted as reflecting the structural mobility of this region and not as two well defined structures [Fig. 2(F) of Supplementary Online Material]. (3) The loop between β 5 and α 2 contains the absolutely conserved G85 (see above). (4)

Finally, an almost absolutely conserved arginine residue (R148) is found in the linker between α 4 and β 8. Figure 2(D,E) of Supplementary Online Material present front and top views that highlight the positions of the AXHPSE motif, the acidic DEP(D/E) motif, and the conserved G95 and R148 residues in the YaeQ structure.

Comparison of YaeQ with proteins in the 3D structure databases

Comparison of the YaeQ structure with other protein structures deposited in public databases using the Dali²³ and ProFunc²⁴ servers identified proteins with low structural similarity whose best hits could be classified as endonucleases (restriction endonuclease of PD-(D/E)XXK superfamily including the related GD-ERKX₃D family), putative methyltransferases and oxidoreductases (Table 1 of Supplementary Online Material). There is no primary structure similarity between YaeQ and any of the above-mentioned structures and tertiary structure similarity is restricted in all cases to the amino-terminal section of YaeQ and, in most cases specifically to the α 1– β 3– β 4– β 5– α 2– β 6– α 3– β 7 motif beginning with the central α -helix (α 1) (Fig. 3 of Supplementary Online Material and data not shown). Thus, the YaeQ fold differs from all of these proteins in two very general aspects: (i) the structure adopted by the last 50 residues which fold into α 4 and beta strands β 8– β 11 and (ii) with the exception of the SdaI C-terminal domain, by the presence of two initial β -strands (β 1– β 2) that precede α 1 [Fig. 3(A) of Supplementary Online Material]. As mentioned above, these two regions come together to form one of the two β -sheets of the β -barrel that encases the central α -helix. Thus, in contrast to the YaeQ central α -helix α 1, the corresponding α -helices in all the structures listed in Table 1 of Supplementary Online Information are invariably highly exposed to the solvent. The most extensive tertiary structure superposition that could be achieved was between YaeQ and the C-terminal domain (residues 160–323) of the recently described restriction endonuclease SdaI.²⁵ A description of this superposition as well as that of the YaeQ structure with the yeast RNA polymerase RPB5 subunit is given in the legend to Figure 3 of Supplementary Online Material.

DISCUSSION

The YaeQ structure represents a new compact protein fold built around a variation of the PD-(D/E)XXK nuclease motif

The majority of structures listed in Table I of Supplementary Online Material belong to the PD-(DE)XXK superfamily of Mg²⁺-dependent nucleases: SdaI, FokI, BglII, Ecl18kI, BamHI, NgoMIV, BstYI restriction endonucleases, and Tt1808, a putative nuclease remotely related to Holliday junction resolvases. Members of this

superfamily share little or no sequence similarity except for the presence of two acidic residues, separated by a variable number of residues (most commonly 14), whose spatial configuration is conserved within a common structural core made up of a mixed β -sheet flanked on both sides by α -helices.^{4,26,27} However, a wide variety of structural variation has been observed within the structural core because of insertions of secondary structure elements and migrations of key catalytic residues.²⁸ Within the PD-(D/E)XK motif, the first acidic residue is almost always an Asp while the second acidic residue can be Asp or Glu and even sometimes another amino acid.²⁹ The conserved acidic residues have been observed to coordinate Mg^{2+} and water molecules in the active site.³⁰ The Lys is even less conserved and has been shown to “migrate” to other positions in the primary structure or can be replaced by Asn, Gln, or acidic residues.²⁸

The superposition of specific YaeQ residues with those corresponding to the PD-(D/E)XK motif in some other

members of this superfamily (SdaI, NgoMIV, BamHI, Hef XPF/Rad1/Mus81 family nuclease, and the conserved hypothetical protein Tt1808) which were identified in the database searches for similar structures are presented in Figure 1(B) and in Figure 4 of Supplementary Online Material. YaeQ residues P67, D68, and D83 correspond structurally to the P, D, and (D/E) residues of the PD-(D/E)XK motif [Fig. 1(C)]. Instead of a Lys at position 85, YaeQ has an absolutely conserved Gly at this position (Figs. 1 and 2). This Gly forms part of a well structured loop that connects $\beta 5$ with $\alpha 2$ and contains a well-conserved Pro at position 87 [Fig. 1(B)]. The importance of the Gly and Pro residues at these positions may be related to the fact that in YaeQ this loop is significantly shorter and more compact than that observed in other PD-(D/E)XK family members [Fig. 1(B)] and that it contributes several functional groups that may be important for catalysis (see below and Fig. 3). Furthermore, the indole nitrogen of the absolutely conserved Trp residue

Figure 3

Active site analysis. (A) The hydrogen bond network involving residues (red) of the putative catalytic site of YaeQ. Also shown are the interactions between R39 and E35 which takes part in the hydrogen bond network shown in Figure 3(C) and between W81 and D83 [see also Fig. 1(B)]. Hydrogen bonds (up to 3.3 Å) are indicated with thin dotted lines. Water molecules are shown as green spheres. (B, C) Structural superposition of YaeQ (red) with the active site of BamHI. Three BamHI structures are shown: BamHI in the absence of DNA (blue, 1BAM),²⁹ the pre-cleavage complex with DNA in the presence of Ca^{2+} (cyan, 2BAM)³⁰ and the post-cleavage complex with DNA in the presence of Mg^{2+} (magenta, 3BAM).³⁰ Ca^{2+} and Mg^{2+} ions are shown in cyan and magenta spheres, respectively. The putative nucleophilic water molecule from the 2BAM structure is shown in green. Three nucleotides of the pre- and post-cleavage complexes are shown with the scissile bond at the center and the post-cleavage 5' phosphate colored olive. In (B) the four acidic residues in the BamHI active site and the corresponding residues in YaeQ are shown. Also shown are the backbone atoms of L84 and F112 of YaeQ and BamHI respectively. Interactions in the BamHI-DNA pre-cleavage complex between specific BamHI residues, Ca^{2+} ions and nucleophilic water (thin dashed lines) and between the nucleophilic water and phosphate (thick dashed line) are shown. In (C) a structural superposition of two BamHI lysine residues with two conserved YaeQ arginines is shown. In the BamHI-DNA complexes,³⁰ these basic residues interact with the phosphate groups on the DNA backbone one nucleotide 3' and one nucleotide 5' of the scissile bond (dashed lines). Hydrogen bonds (up to 3.6 Å) observed in the 2BAM structure shown are indicated with thin dashed lines. Figures were produced with Pymol.²¹

at position 81 H-bonds with the D83 side chain which is implicated in metal coordination and catalysis in other members of the PD-(D/E)XK family [Fig. 1(B) and Fig. 3(A)]. Finally, the highly conserved R91 in YaeQ superposes well with functionally important basic residues found at corresponding positions in the structures of PD-(D/E)XK members (Fig. 4 of Supplementary Online Material; also see Fig. 3 and Discussion below). In all of the few cases in the YaeQ family where R91 is substituted with a hydrophobic residue (usually a tryptophan, Fig. 2) the adjacent amino acid at position 90 is a Lys or Arg (Fig. 2 and data not shown). Thus, the YaeQ family is characterized by a variation of the PD-(D/E)XK motif that is best described as P(D/E)-WX(D/E)XGXPX₃R [Fig. 1(C)].

Since in YaeQ the position of Lys in the PD-(D/E)XK motif is occupied by the absolutely conserved G85 [Fig. 1(B,C)], we searched for basic residues that could be candidates for the migrating lysine. Figure 4 of Supplementary Online Material shows that YaeQ residue R39 superposes with residues Q178 in SdaI. R39 is well conserved in the YaeQ family of proteins (Fig. 2). YaeQ residues R39, D68, D83, R91 interact together through a hydrogen bonding network that includes three bound water molecules (O75, O104, and O137) as well as the indole nitrogen of W81 [Fig. 3(A)]. This highly solvent exposed network is also connected to the hydrogen bond network of the AXHPSE motif via a pair of direct hydrogen bonds between the R39 and E35 side chains [Fig. 2(C) of Supplementary Online Material].

Two of the best characterized nucleases from the PD-(D/E)XK family, in terms of structural features important for the catalytic mechanism are the NgoMIV endonuclease which has a P₁₃₉D₁₄₀S₁₈₅K₁₈₇ motif and BamHI which has a I₉₃D₉₄E₁₁₁E₁₁₃ motif [Fig. 1(C)]. Comparison of the YaeQ structure with the NgoMIV-DNA product complex³¹ is presented in Figure 4 of the Supplementary Online Material. BamHI has been crystallized without DNA²⁹ and in two DNA-bound forms: a pre-cleavage state with two bound Ca²⁺ ions and a post-cleavage state with two bound Mn²⁺ ions.³⁰ Figure 3(B,C) presents the superposition of the BamHI active site in these three structures with the corresponding residues in YaeQ. While BamHI has three acidic residues which contribute directly to metal ion coordination (E77, D94, E111) and a fourth (E113) which binds a putative nucleophilic water molecule, YaeQ has only two acidic residues, D68 and D83, that adopt similar conformations in the positions corresponding to D94 and E111, respectively, in BamHI [Fig. 3(B)]. We note that a small but significant fraction of YaeQ family members (<20%) have an Ala residue at position 68 (Fig. 2). Interestingly, over half of the variants with an alanine residue at this position come from the *Pseudomonadae* family of proteobacteria. Furthermore, the main chain carbonyl group of residues L84 in YaeQ is oriented precisely in the same

manner as that of residue F112 in BamHI which coordinates one of the active site metal ions [Fig. 3(B)]. The position of E113 in BamHI is occupied by the absolutely conserved G85 in YaeQ [Fig. 3(B)]. This residue may be important to maintain the structure of the loop containing D83 and L84 since it adopts main chain angles that are prohibited for non-glycine residues.

BamHI residues K61, E111, K126 undergo significant structural changes upon DNA binding. For example, the K61 side chain swings through a distance of ~2.5 Å upon BamHI binding to DNA [Fig. 3(C)].³⁰ In all cases, the corresponding YaeQ residue (R39, D83, and R91) overlaps best with the free BamHI structure [Fig. 3(B,C)]. Most interestingly YaeQ R39 and R91 superpose well with the K61 and K126 side chains in the BamHI apo structure which interact, respectively, with phosphate groups on the DNA backbone one nucleotide 3' and one nucleotide 5' of the scissile bond in both pre- and post-cleavage BamHI-DNA complexes [Fig. 3(C)].³⁰

We used the superposition of the YaeQ structure with the BamHI-DNA-Ca²⁺ pre-cleavage complex to analyze the disposition of other conserved YaeQ residues in relation to a putative DNA ligand. In this superposition, the loop formed by the highly conserved and YaeQ-specific AXHPSE motif is pointed towards the minor groove of the DNA double helix (Fig. 5 of Supplementary Online Material). Furthermore, the loop (residues G58-D68) that contains the disordered residues 61–64 as well as the highly conserved DEP(D/E) motif is also pointed towards the minor groove (Fig. 5 of Supplementary Online Material). However, the corresponding loop in the BamHI-DNA-Ca²⁺ complex (residues R76-D94) is splayed out and away from the minor groove in order to make contacts with the sugar-phosphate backbone³⁰ [Fig. 5(A) of Supplementary Online Material]. Interestingly, residues 79–91 of this loop are not visible in the BamHI structure determined in the absence of DNA.²⁹ This suggests that the corresponding motif in YaeQ may tend to adopt a more defined structure only in the presence of a specific nucleic acid ligand. The above observations therefore point towards a possible role of both of these motifs in contacts with nucleic acid structures.

Possible YaeQ function

YaeQ homologs have been identified in over 130 bacterial species. The distribution of YaeQ homolog genes within sequenced genomes does not, however, reveal any conserved pattern or preference for specific neighboring genes and specifically, no apparent genes coding for cognate methyltransferases. As described in the Introduction, the only studies on YaeQ function arrived at contradictory results: the first study⁶ provided evidence for a role in the control of RfaH-dependent transcription termination in *E. coli* and *Salmonella* while the second study⁷ found no evidence for a role for YaeQ in *in vitro* or *in*

vivo transcription expression assays. While the YaeQ structure was most closely related to restriction endonucleases of the PD-(D/E)XK family, we believe that its function will eventually be shown to be distinct from that of restriction endonucleases. This is because while the highly divergent restriction endonucleases are species or genera-specific, the YaeQ family of proteins is highly conserved among a wide variety of proteobacteria.

The PD-(D/E)XK superfamily of proteins includes enzymes such as MthH, an essential protein for initiation of methyl-directed DNA mismatch repair,³² archeal Holliday junction resolvases,^{33–35} XPF/Rad1/Mus81-dependent nuclease family members that cleave branched structures generated during DNA repair, replication, and recombination,^{20,36,37} and RecB, part of the RecBCD multifunctional enzyme complex that processes DNA ends resulting from a double-stranded breaks.³⁸ Therefore, it is likely that YaeQ plays a role in a central process of importance to a wide variety of bacteria that most likely involves the remodeling or modification of nucleic acid (DNA or RNA) structures.

ACKNOWLEDGMENTS

We thank the Mark Glover laboratory (University of Alberta) for the opportunity to collect MAD data at the Advanced Light Source.

REFERENCES

- Oliver SG. From DNA sequence topological function. *Nature* 1996;379:597–600.
- Moult J, Melamud E. From fold to function. *Curr Opin Struct Biol* 2000;10:384–389.
- Kinch LN, Ginalski K, Rychlewski L, Grishin NV. Identification of novel restriction endonuclease-like fold families among hypothetical proteins. *Nucleic Acids Res* 2005;33:3598–3605.
- Kosinski J, Feder M, Bujnicki JM. The PD-(D/E)XK superfamily revisited: identification of new members among proteins involved in DNA metabolism and functional predictions for domains of (hitherto) unknown function. *BMC Bioinformatics* 2005;12:172–184.
- Galvao-Botton LM, Katsuyama AM, Guzzo CR, Almeida FC, Farah CS, Valente AP. High-throughput screening of structural proteomics targets using NMR. *FEBS Lett* 2003;552:207–213.
- Wong KR, Hughes C, Koronakis V. A gene, *yaeQ*, that suppresses reduced operon expression caused by mutations in the transcription elongation gene *rfaH* in *Escherichia coli* and *Salmonella typhimurium*. *Mol Gen Genet* 1998;257:693–696.
- Vicari D, Arsimovitch I. Virulence regulators RfaH and YaeQ do not operate in the same pathway. *Mol Genet Genomics* 2004;272:489–496.
- Guzzo CR, Nagem RA, Galvao-Botton LM, Guimaraes BG, Medrano FJ, Barbosa JARG, Farah CS. Expression, purification, crystallization and preliminary X-ray analysis of YaeQ (XAC2396) from *Xanthomonas axonopodis* pv. citri. *Acta Crystallogr F* 2005;61:493–495.
- Barbosa JARG, Soares-Netto LE, Farah CS, Schenkman S, Meneghini R. The structural molecular biology network of the State of São Paulo, Brazil. *Anais da Academia Brasileira de Ciências* 2006;78:241–253.
- Schneider TR, Sheldrick GM. Substructure solution with SHELXD. *Acta Crystallogr D* 2002;58:1772–1779.
- La Fortelle E de, Bricogne G. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol* 1997;276:472–493.
- Abrahams JP, Leslie AGW. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr D* 1996;52:30–42.
- Cowan K. Joint CCP4 and ESR-EACBM newsletter on protein. *Crystallography* 1994;31:34–38.
- Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D* 1994;50:760–763.
- Perrakis A, Morris R, Lamzin VS. Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* 1999;6:458–463.
- Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* 1997;53:240–255.
- Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D* 1998;54:905–921.
- Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D* 2004;60:2126–2132.
- Chio KC, Chin KH, Gao FP, Lyu PC, Shr HL, Wang AH, Chou SH. The cloning, crystallization and preliminary X-ray analysis of XC2113, a YaeQ protein from *Xanthomonas campestris*. *Acta Crystallogr F* 2006;62:1046–1048.
- Nishino T, Komori K, Ishino Y, Morikawa K. X-ray and biochemical anatomy of an archaeal XPF/Rad1/Mus81 family nuclease. Similarity between its endonuclease domain and restriction enzymes. *Structure* 2003;11:445–457.
- DeLano WL. The PyMOL molecular graphics system. 2002. <http://www.pymol.org>.
- Krissinel E, Henrick K. Detection of protein assemblies in crystals. In: Berthold MR, Glen R, Diederichs K, Kohlbaier O, Fischer I, editors. *CompLife 2005*. LNBI 3695. Berlin: Springer Verlag; 2005. pp 163–174.
- Holm L, Sander C. Touring protein fold space with Dali/FSSP. *Nucleic Acids Res* 1998;26:316–319.
- Laskowski RA, Watson JD, Thornton JD. ProFunc: a server for predicting protein function from 3D structure. *Nucleic Acids Res* 2005;33:W89–W93.
- Tamulaitiene G, Jakubauskas A, Urbanke C, Huber R, Grazulis S, Siksnys V. The crystal structure of the rare-cutting restriction enzyme SdaI reveals unexpected domain architecture. *Structure* 2006;14:1389–1400.
- Aggarwal AK. Structure and function of restriction endonucleases. *Curr Opin Struct Biol* 1995;5:11–19.
- Venclovas C, Timinskas A, Siksnys V. Five-stranded β -sheet sandwiched with two α -helices: a structural link between restriction endonucleases EcoRI and EcoRV. *Proteins* 1994;20:279–282.
- Feder M, Bujnicki JM. Identification of a new family of putative PD-(D/E)XK nucleases with unusual phylogenomic distribution and a new type of the active site. *BMC Genomics* 2005;6:21–33.
- Newman M, Strzelecka T, Dorner LF, Schildkraut I, Aggarwal AK. Structure of restriction endonuclease BamHI phased at 1.95 Å resolution by MAD analysis. *Structure* 1994;2:439–452.
- Viadiu H, Aggarwal AK. The role of metals in catalysis by the restriction endonuclease BamHI. *Nat Struct Biol* 1998;5:910–916.
- Deibert M, Grazulis S, Sasnauskas G, Siksnys V, Huber R. Structure of the tetrameric restriction endonuclease NgoMIV in complex with cleaved DNA. *Nat Struct Biol* 7; 2000:792–799.
- Ban C, Yang W. Structural basis for MthH activation in *E. coli* mismatch repair and relationship of MthH to restriction endonucleases. *EMBO J* 1998;17:1526–1534.

PD-ID/E)XK Nuclease Motif in YaeQ Structure

33. Nishino T, Komori K, Tsuchiya D, Ishino Y, Morikawa K. Crystal structure of the archaeal holliday junction resolvase Hjc and implications for DNA recognition. *Structure* 2001;9:197–204.
34. Bond CS, Kvaratskhelia M, Richard D, White MF, Hunter WN. Structure of Hjc, a Holliday junction resolvase, from *Sulfolobus solfataricus*. *Proc Natl Acad Sci USA* 2001;98:5509–5514.
35. Middleton CL, Parker JL, Richard DJ, White MF, Bond CS. Substrate recognition and catalysis by the Holliday junction resolving enzyme Hje. *Nucleic Acids Res* 2004;32:5442–5451.
36. Tsutakawa SE, Muto T, Kawate T, Jingami H, Kunishima N, Ariyoshi M, Kohda D, Nakagawa M, Morikawa K. Crystallographic and functional studies of very short patch repair endonuclease. *Mol Cell* 1999;3:621–628.
37. Kovall RA, Matthews BW. Structural, functional, and evolutionary relationships between lambda-exonuclease and the type II restriction endonucleases. *Proc Natl Acad Sci USA* 1998;95:7893–7897.
38. Singleton MR, Dillingham MS, Gaudier M, Kowalczykowski SC, Wigley DB. Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature* 2004;432:187–193.

Supplementary Online Material

The structure of *Xanthomonas axonopodis*
pv. *citri*
YaeQ reveals new compact protein fold built
around
a variation of the PD-(D/E)XK nuclease motif

Cristiane R. Guzzo¹, Ronaldo A. P. Nagem²,
João A. R. G. Barbosa³,
and Chuck S. Farah^{1*}

1 Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, Avenida Prof. Lineu Prestes 748, São Paulo, SP, CEP 05508-000, Brazil.

2 Departamento de Bioquímica e Imunologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Avenida Antônio Carlos 6627, Belo Horizonte, MG, CEP 31270-901, Brazil.

3 Centro de Biologia Molecular Estrutural, Laboratório Nacional de Luz Síncrotron, C.P. 6192, Campinas, SP, CEP 13084-971, Brazil

* To whom correspondence should be sent.
Chuck S. Farah

Dept. de Bioquímica, Instituto de Química
Universidade de São Paulo
Av. Prof. Lineu Prestes 748
Cidade Universitária, São Paulo, SP
Brasil 05508-000

Tel: (+55 11) 3091-3312 ext 104

FAX: (+55 11) 3091-3312 ext 204; 3815-5579

email: chsfarah@iq.usp.br

Figure 1. YaeQ topology and structure

A) YaeQ topology. Numbers indicate the first and last residues in each secondary structure element. Nomenclature ($\alpha 1$, $\beta 1$, etc.) used to refer to secondary structure elements is shown below the diagram. B) Back view of ribbon representations of the YaeQ structure. Strands $\beta 10$ and $\beta 11$ are almost colinear and are separated by a light bend at T172 and P173 at the back of the molecule. This linker between $\beta 10$ and $\beta 11$ as well as that between $\beta 7$ and $\alpha 4$ connect the two β -sheets together and in this way bury the central α -helix at the back of the molecule. Solvent access to the central α -helix is thereby limited to only three residues (R39, R43 and F47) at the front face whose side chains can be accessed between the edge strands ($\beta 2$ and $\beta 3$) of the two β -sheets (See Fig. 2A of Supplementary Online Material). Cartoons show α -helices (red), β -strands (yellow) and coil regions and turns (green). Figures were produced with Pymol¹ using chain B of the asymmetric unit.

Figure 2. Localization of conserved residues on the YaeQ structure

A) Front view of the YaeQ surface showing that R39 (red), R43 (blue) and F47 (green) of the central α -helix are the only residues that can be accessed between strands β 2 and β 3. R39 is even more exposed at the top face of the molecule.

B) Ribbon model of YaeQ showing the locations of the absolutely conserved residues S34, E35, R43, W81 and G85 in blue.

C) Set of interactions involving absolutely conserved residues S34, E35, R43 (blue) and highly conserved residues A30, H32, R39, F57 (red) that extend from the top face of the molecule to within the cleft that partially exposes the central α -helix. Hydrogen bonds (up to 3.7 Å) are indicated with dotted lines. The R43 δ -guanido group stacks perpendicular over the aromatic ring of F57, and makes a set of five specific hydrogen bonds with the two carboxylate oxygens of E35, the carbonyl backbone oxygen of A30 and two water molecules (wat25 and wat54 for chain A and wat8 and wat213 for chain B). The E35 carboxylate group also accepts hydrogen bonds from the main chain amide of H32 and the side chain of R39. A30, H32, R39 and F57 are all highly conserved in the YaeQ family. While the S34 side chain is not well defined in chain B, in chain A the side-chain hydroxyl oxygen is within 3.8 Å of the NE2 atom of the H32 side chain. The two other absolutely conserved residues are W81 and G85. G85 is found in a well structured loop between β 5 and α 2 that includes a well-conserved proline

residue at position 87. This loop is located at the top surface of the molecule. G85 has main chain angles $\phi=93^\circ$ and $\psi=-141^\circ$ which are not permitted for residues other than glycines. W81 is located within the hydrophobic core of the molecule and seems to be involved in the maintenance of the YaeQ structure. W81 takes part in the hydrophobic packing of the β 5- α 2- β 6- α 3- β 7 motif in which its indole ring is sandwiched between P64 and A95. The outer edge of the ring is lined with a set of charged side chains (D83, R91, K94, R98) and the indole NH group makes a hydrogen bond with the carboxyl group of D83. While W81 of chain A is completely buried, in chain B it is partially exposed due to an alternative conformation of the K94 side chain. W81, D83 and G85 are located in a region of probable functional importance

Fig

D, E) Front and top views (respectively) of the YaeQ surface in which the AXHPSE motif is shown in blue, the DEP(D/E) motif is shown in cyan, G85 is shown in yellow and R148 is shown in orange. The DEP(D/E) motif is in fact part of a larger loop that includes residues 61-64 (green). **F)** Superposition of the $C\alpha$ chains of the YaeQ monomers in the asymmetric unit. The AXPSE motif is shown in blue and the DEP(D/E) motif is shown in cyan. Residues 61-64 of chain A (yellow) and chain B (green) have poor electron density and were modeled with different conformations. Figures produced with Pymol¹.

Figure 3. Structural superposition of YaeQ with SdaI and RPB5

A) Structural superposition of YaeQ (red) with the C-terminal domain of restriction endonuclease SdaI² (pink, Top view). The YaeQ and SdaI C-domain (SdaI_C) structures differ by three general features: i) An insertion of a β - α -meander (SdaI residues 204-233)

between $\beta 3$ and $\beta 4$ in the YaeQ fold. This insertion packs against the Sda_C β -sheet but maintains the common topology of the respective β -sheets in the two proteins; ii) Residues 134-174 that form $\alpha 4$ and $\beta 8$ - $\beta 10$ in

YaeQ are replaced in Sda_C by a loop that joins two β strands that overlap with $\beta 7$ and $\beta 11$ in YaeQ. iii) Finally, Sda_C has a C-terminal α -helix (residues 313-323) not present in the YaeQ fold.

B) Structural superposition of YaeQ (red) with yeast RNA polymerase subunit RPB5 (Front view). The N-terminal domain of RPB5 is colored light blue while the C-terminal domain is grey. The superposition of YaeQ on the yeast RNA polymerase RPB5 subunit is also based on the central α/β sandwich motif of the N-terminal domain. This N-terminal domain has no nucleolytic activity and is thought to interact with downstream DNA during eukaryotic transcription.³ The structural relatedness of RPB5 with PD-(D/E)XK superfamily proteins has been previously observed.⁴ The structural superposition of YaeQ with a putative RNA methyltransferase⁵ and a Class 1 acetohydroxy acid isomeroreductase⁶ are restricted to a single β -sheet sandwiched between two α -helices (data not shown). Figures produced with Pymol.¹

Figure 4. Structural superposition of PD-(D/E)XK motif residues.

Structural superposition of YaeQ (red) with the residues of active sites of the type II restriction endonucleases Sda (light pink, 1IXS)², NgoMIV (cyan, 1FIU)⁸ and Tt1808 (blue, 1WDJ)⁷. Residues R39 and R43 of YaeQ and the corresponding residues in the other three structures are also shown since R39 may be involved in YaeQ function (see main text) and R43 (absolutely conserved the YaeQ family of proteins, Fig. 2 of main text) superposes with E70 of NgoMIV which may be involved in catalysis⁸. The NgoMIV structure was obtained in complex with DNA (not shown) and Mg²⁺

(green spheres). In this structure, D140 coordinates two Mg²⁺ ions while K187 is 3.5 Å from the 5' phosphate group of the cleaved phosphate-ester bond. In addition, two other acidic residues E70 and E201 contribute to metal coordination via water bridges.⁸ The side chains of YaeQ residues D68, D83 and R91 occupy structurally similar positions to the active site residues D140, S185 and E201, respectively, involved in metal coordination and catalysis in NgoMIV. NgoMIV residue E70 has also been implicated in metal binding in the catalytic site. While its C α position corresponds to that of YaeQ residue R43, the R43 side chain adopts a well-defined conformation in which it is pointed away from the putative active site. However, the side chain of YaeQ residue R39, one turn up the YaeQ central α -helix points towards the metal-binding site (and see comparison with BamHI catalytic site below).

Sda exhibits a new and unusual variation in which the acidic groups are conserved but the Lys residue migrates in the sequence to Q178 or R260 (Ref 2). In the case of NgoMIV and BamHI (see below) endonucleases, the respective Lys and Glu residues at this position are thought to coordinate and activate the attacking water molecule^{8,9} (the K187 side chain in NgoMIV is shown in Fig. 5A). The actual nature of the Lys migration in Sda is not clear but mutational analysis demonstrated that a substitution of Q178 abolishes catalytic activity.² Figures produced with Pymol.¹

Figure 5. Disposition of YaeQ in relation to a putative DNA ligand based on the structural superposition of YaeQ with BamHI-DNA-Ca²⁺ complex.

A) Structural superposition of ribbon representations of YaeQ with the BamHI-DNA-Ca²⁺ complex.⁹ For clarity, only BamHI residues 58-179, whose secondary structure elements adopt the same topology as that observed in YaeQ, are shown. Secondary structure elements in YaeQ are colored red except for

the AXHPSE motif (dark blue) and residues 58-68 (green) which contains the disordered residues 61-64 as well as the highly conserved DEP(D/E) motif. BamHI is colored orange except for the loop made up of residues R76-D94 (magenta) which corresponds to residues YaeQ residues 58-68. Note that in the BamHI structure determined in the absence of DNA residues 79-91 were not visible.¹⁰ The double-stranded DNA ligand of the BamHI-DNA-Ca²⁺ complex is represented by an orange ribbon and the scissile phosphate is shown as a black stick model. **B)** Ribbon and **C)** surface representations of the structural superposition of YaeQ with the DNA ligand of the BamHI-DNA-Ca²⁺ complex. Coloring is the same as in **(A)** except that the four charged residues of the putative YaeQ active site (R39, D68, D83 and R91 shown as sticks in part **B)** are colored blue-white. Figures produced with Pymol.¹

Table 1. Proteins structurally similar to YaeQ identified using the Dali and Profunc servers.

Z-score	RMSD (Å)	PDB	Enzyme	Organism	Reference
Dali[#]					
5.1	4.0	2IXS	SdaI restriction endonuclease	<i>Streptomyces diastaticus</i>	2
4.9	4.4	2FO	FokI restriction endonuclease	<i>Flavobacterium</i>	11
4.9	3.0	1D2I	BglII restriction endonuclease in	<i>Bacillus subtilis</i>	12
4.5	4.3	2GB	<i>Ec18kl</i> restriction endonuclease	<i>Enterobacter cloacae</i>	13
4.5	3.7	1BA	BamHI restriction endonuclease	<i>Bacillus</i>	10
4.9	2.6	2A1L	DNA excision repair protein ERCC-1	<i>Homo sapiens</i>	14
4.7	3.2	2bG	XPF endonuclease in complex with	<i>Aeropyrum pernix</i>	15
4.8	4.2	1DZF	RNA polymerase - RPB5 subunit	<i>Saccharomyces</i>	4
4.2	3.1	1YV	60-kDa SS-A/Ro ribonucleoprotein	<i>Xenopus laevis</i>	16
5.3	3.3	1Y88	Hypothetical restriction endonuclease	<i>Archaeoglobus fulgidus</i>	17
4.8	3.1	1XM	Hypothetical Protein	<i>Vibrio cholerae</i>	18
4.4	2.8	1WD	Tt1808 - Conserved hypothetical protein	<i>Thermus thermophilus</i>	7
4.1*	2.8	1J22	XPF/Rad1/Mus81 family nuclease	<i>Pyrococcus furiosus</i> DSM	19
4.6*	4.6	1QM	Acetohydroxy acid isomeroreductase	<i>Spinacia oleracea</i>	20
4.1*	4.1	1NP3	Acetohydroxy acid isomeroreductase	<i>Pseudomonas</i>	6
ProFunc^{##}					
4.2	3.0	1FIU	NgoMIV restriction endonuclease	<i>Neisseria gonorrhoeae</i>	8
3.4	2.5	1SD	<i>Bst</i> YI restriction endonuclease	<i>Geobacillus</i>	21
3.1	2.4	1BH	BamHI restriction endonuclease in	<i>Bacillus</i>	22
4.1	2.3	2bH	XPF endonuclease	<i>Aeropyrum pernix</i>	15
4.1	2.9	1WY	RNA methyltransferase	<i>Pyrococcus horikoshii</i>	5
3.0	2.9	1KJN	Hypothetical protein conserved in	<i>Methanoth.</i>	23

(#) Searches with the Dali server²⁴ was performed using YaeQ residues 4-134 and the top 12 hits are shown. (*) Searches using the complete YaeQ structure yielded qualitatively similar results with the addition of two acetohydroxy acid isomeroreductases and the Xpf/Rad1/Mus81 family nuclease domain. (##) All Profunc²⁵ hits with Z-score >2 are listed. Only the best hit of several that were obtained with BamHI structures is shown.

The YaeQ structure showed structural similarity with the catalytic domain of members of the XPF/Rad1/Mus81 family that contains a highly conserved GDX_nERKX₃D motif. This motif is structurally related to the catalytic site of the PD-(D/E)XK endonucleases in which the first two acidic residues are crucial for catalysis.²⁶ Although the two acidic residues D68 and D83 YaeQ superpose well with the corresponding residues in this family, YaeQ residues P₆₇D₆₈X₁₅D₈₃L₈₄G₈₅X₃E₈₉ do not correspond to the consensus conserved catalytic site GDX_nERKX₃D. Furthermore, YaeQ E89 does not superpose on the conserved Asp in the last position of this motif in the structures of the XPF/Rad1/Mus81 family of proteins (for example residue D54 of the Hef nuclease, Fig. 1B of the main text).

References

- DeLano WL The PyMOL Molecular Graphics System on World Wide Web <http://www.pymol.org> 2002
- Tamulaitiene G, Jakubauskas A, Urbanke C, Huber R, Grazulis S, Siksnys V. The Crystal Structure of the Rare-Cutting Restriction Enzyme SdaI Reveals Unexpected Domain Architecture. *Structure* 2006;**14**:1389-1400.
- Cramer P, Bushnell DA, Fu J, Gnatt AL, Maier-Davis B, Thompson NE, Burgess RR, Edwards AM, David PR, Kornberg RD. Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* 2000;**288**:640-649.
- Todone F, Weinzierl RO, Brick P, Onesti S. Crystal structure of RPB5, a universal eukaryotic RNA polymerase subunit and transcription factor interaction target. *Proc Natl Acad Sci USA* 2000;**97**:6306-6310.
- Gao YG, Yao M, Yong Z, Tanaka I. Crystal structure of the putative RNA methyltransferase PH1948 from *Pyrococcus horikoshii*, in complex with the copurified S-adenosyl-L-homocysteine. *Proteins* 2005;**61**:1141-1145.
- Ahn HJ, Eom SJ, Yoon HJ, Lee BI, Cho H, Suh SW. Crystal Structure of Class I Acetohydroxy Acid Isomeroreductase from *Pseudomonas aeruginosa*. *J Mol Biol* 2003;**328**:505-515.
- Idaka M, Wada T, Murayama K, Terada T, Kuramitsu S, Shirouzu M, Yokoyama S. Crystal Structure Of TT1808 From *Thermus Thermophilus* HB8. To be published.
- Deibert M, Grazulis S, Sasnauskas G, Siksnys V, Huber R. Structure of the tetrameric restriction endonuclease NgoMIV in complex with cleaved DNA. *Nat Struct Biol* 2000;**7**:792-799.
- Viadiu H, Aggarwal AK. The role of metals in catalysis by the restriction endonuclease BamHI. *Nat Struct Biol* 1998;**5**:910-916.
- Newman M, Strzelecka T, Dorner LF, Schildkraut I, Aggarwal AK. Structure of restriction endonuclease BamHI phased at 1.95 Å resolution by MAD analysis. *Structure* 1994;**2**:439-452.
- Wah DA, Bitinaite J, Schildkraut I, Aggarwal AK. Structure of FokI has implications for DNA cleavage. *Proc Natl Acad Sci USA* 1998;**95**:10564-10569.
- Lukacs CM, Kucera R, Schildkraut I, Aggarwal AK. Understanding the immutability of restriction enzymes: crystal structure of BglII and its DNA substrate at 1.5 Å resolution. *Nat Struct Biol* 2000;**7**:134-140.
- Bochtler M, Szczepanowski RH, Tamulaitis G, Grazulis S, Czapinska H, Manakova E, Siksnys V. Nucleotide flips determine the specificity of the Ecl18kI restriction endonuclease. *EMBO J* 2006;**25**:2219-2229.
- Tsodikov OV, Enzlin JH, Scharer OD, Ellenberger T. Crystal Structure and DNA Binding Functions of ERCC1, a Subunit of the DNA Structure-Specific Endonuclease XPF-ERCC1. *Proc Natl Acad Sci USA* 2005;**102**:11236-11241.
- Newman M, Murray-Rust J, Lally J, Rudolf J, Fadden A, Knowles PP, White MF, McDonald NQ. Structure of an Xpf Endonuclease with and without DNA Suggests a Model for Substrate Recognition. *EMBO J* 2005;**24**:895-905.
- Stein AJ, Fuchs G, Fu C, Wolin SL, Reinisch KM. Structural Insights into RNA Quality Control: The Ro Autoantigen Binds Misfolded RNAs via Its Central Cavity. *Cell* 2005;**121**:529-639.
- Lunin VV, Evdokimova E, Kudritskaya M, Cuff ME, Joachimiak A, Edwards A, Savchenko A. The crystal structure of hypothetical protein AF1548 from *Archaeoglobus fulgidus*. To be Published.
- Chang C, Joachimiak A, Wu R, Moy S, and Midwest Center for Structural Genomics. Crystal structure of Hypothetical protein VC1899. To be Published.
- Nishino T, Komori K, Ishino Y, Morikawa K. X-Ray and Biochemical Anatomy of an Archaeal XPF/Rad1/Mus81 Family

Nuclease. Similarity between Its Endonuclease Domain and Restriction Enzymes. *Structure* 2003;**11**:445-457.

20. Thomazeau K, Dumas R, Halgand F, Forest E, Douce R, Biou V. Structure of spinach acetohydroxyacid isomeroreductase complexed with its reaction product dihydroxymethylvalerate, manganese and (phospho)-ADP-ribose. *Acta Cryst* 2000;**D56**:389-397.

21. Townson SA, Samuelson JC, Vanamee ES, Edwards TA, Escalante CR, Xu SY, Aggarwal AK. Crystal Structure of BstYI at 1.85 Å Resolution: A Thermophilic Restriction Endonuclease with Overlapping Specificities to BamHI and BglII. *J Mol Biol* 2004;**338**:725-733.

22. Newman M, Strzelecka T, Dorner LF, Schildkraut I, Aggarwal AK. Structure of Bam HI endonuclease bound to DNA: partial folding and unfolding on DNA binding. *Science* 1995;**269**:656-663.

23. Christendat D, Edwards A, Joachimiak A, Korolev S. Crystal structure of MT0777. To be published.

24. Holm L, Sander C. Touring protein fold space with Dali/FSSP. *Nucl. Acids Res* 1998;**26**:316-319.

25. Laskowski RA, Watson JD, Thornton JD. ProFunc: a server for predicting protein function from 3D structure. *Nucleic Acids Res* 2005;**33**:W89-W93.

26. Nishino T, Komori K, Ishino Y, Morikawa K. X-Ray and Biochemical Anatomy of an Archaeal XPF/Rad1/Mus81 Family Nuclease. Similarity between Its Endonuclease Domain and Restriction Enzymes. *Structure* 2003;**11**:445-457.

4.3.7 Ensaios Funcionais de YaeQ

4.3.7.1 Mutantes de YaeQ_{XAC2396}: YaeQ_{D68A} e YaeQ_{D83A}

Expressão de YaeQ_{D68A} e YaeQ_{D83A}. Com o objetivo de testar algumas hipóteses sobre possíveis funções de YaeQ_{XAC2396}, produzimos dois mutantes de YaeQ_{XAC2396} que teoricamente devem abolir sua atividade endonuclease: YaeQ_{D68A} e YaeQ_{D83A} (**Figura 31**). Estes resíduos pertencem ao sítio putativo catalítico de YaeQ_{XAC2396}, P₆₇(D/E)₆₈-(D/E)₈₃XG₈₅XR₉₁, e foram alterados para alaninas.

Mutações nestes aminoácidos em outras enzimas da família PD-(D/E)XK acabam com a atividade nuclease²⁰²⁻²⁰⁵. Desta forma, é interessante analisar a atividade de YaeQ_{XAC2396} e compará-la com a atividade dos mutantes. A expressão dos mutantes e as suas purificações estão mostradas na **Figura 32**. As amostras purificadas dos mutantes foram confirmadas por Western blotting usando anticorpo contra YaeQ_{XAC2396} selvagem (**Figura 33**).

Figura 31: Mutantes de YaeQ, YaeQ_{D68A} e YaeQ_{D83A}.

a) Sequenciamento do clone YaeQ_{D68A} e **(b)** YaeQ_{D83A}. O cromatograma mostra o sequenciamento dos clones destacando ao códon que foi mutado para alanina, A₆₈ e A₈₃. O alinhamento mostrado refere-se a sequência de aminoácidos derivada da tradução do sequenciamento dos clones de YaeQ_{D68A} e YaeQ_{D83A} (Query) e seu alinhamento com a sequência de YaeQ_{XAC2396} tipo selvagem (Sbjct). As caixas vermelhas destacam a mutação ocorrida de D para A.

Figura 32: Purificação dos mutantes de YaeQ: YaeQ_{D68A} e YaeQ_{D83A}.

a) Larga escala de YaeQ_{D68A} e YaeQ_{D83A} em BL21(DE3)RP. 1: Não induzido, 2: amostra induzida de YaeQ_{D68A}, 3: Amostra induzida de YaeQ_{D83A}, 4: Amostra de YaeQ pura, 5 e 6: Amostras induzidas de YaeQ_{D68A}, 7: poço vazio, 8 e 9: Amostras induzidas de YaeQ_{D83A}. **b)** Frações da eluição de YaeQ_{D68A} e YaeQ_{D83A} quando aplicadas na coluna de Gel Filtração, 1-4: Frações de YaeQ_{D83A}, 5: Amostra pura de YaeQ 6-11: Frações de YaeQ_{D68A}. As setas mostram a proteína de interesse no gel de SDS-PAGE.

Figura 33: Western blotting dos mutantes de YaeQ: YaeQ_{D68A} e YaeQ_{D83A}.

a) Membrana de nitrocelulose corado com “ponceau red” e **b)** filme revelado contra anticorpo policlonal de YaeQ_{XAC2396} tratado com anti-IgG fusionado a proteína A peroxidase. 1: Amostra não induzida de YaeQ_{XAC2396}, 2: amostra induzida de YaeQ_{XAC2396}, 3: amostra pura de YaeQ_{XAC2396}, 4: Amostra não induzida de YaeQ_{D68A}, 5: Amostra induzida de YaeQ_{D68A}, 6: Fração eluída da Troca Iônica contendo YaeQ_{D68A}, 7: Fração eluída da Gel Filtração contendo YaeQ_{D68A}, 8: Amostra pura de YaeQ_{D68A}, 9: Amostra não induzida de YaeQ_{D83A}, 10: Amostra induzida de YaeQ_{D83A}, 11: Fração da Troca Iônica contendo YaeQ_{D83A}, 12: Fração da Gel Filtração contendo YaeQ_{D83A}, 13: amostra pura de YaeQ_{D68A}.

4.3.7.2 Cristalização dos Mutantes de YaeQ: YaeQ_{D68A} e YaeQ_{D83A}, e a determinação da estrutura de YaeQ_{D83A}

Os mutantes de YaeQ_{XAC2396} foram submetidos a ensaios de cristalização utilizando, como ponto de partida, as condições em que a proteína selvagem cristalizava. Para ambos os mutantes, cristais foram observados. Os cristais de YaeQ_{D68A} tem formato de agulhas (**Figura 34 a**) e os cristais de YaeQ_{D83A} são muito parecidos com os cristais de YaeQ_{XAC2396} tipo selvagem (**Figura 34 b e c**).

Os cristais de YaeQ_{D83A} difrataram até 1.45 Å na linha MX1 do LNLS e a estatística do processamento dos dados se encontra na **Tabela 17**. As fases iniciais foram calculadas por substituição molecular usando a estrutura de YaeQ_{XAC2396} como modelo (PDBID: 2G3W). A **Figura 35 a** mostra a densidade eletrônica do aminoácido A₈₃, demonstrando a substituição deste resíduo de aspartato para alanina. A **Figura 35 b** mostra a estrutura de YaeQ_{D83A} presente na unidade assimétrica. O refinamento da estrutura de YaeQ_{D83A} ainda não foi terminado, mas uma inspeção inicial da densidade eletrônica deste mutante indica que o loop que contem os resíduos 61-64 possui densidade eletrônica mais definida do que o mapa da estrutura de YaeQ_{XAC2396} tipo selvagem (**Figura 35 c e d**).

Figura 34: Cristalização dos mutantes de YaeQ.

a) Cristais do mutante de YaeQ_{D68A}, que cristalizou em 20% PEG 4000, 0,1M Tris pH 8,5 e 0,2M acetato de sódio; **b)** Cristais de YaeQ_{D83A}, que cristalizou em 28% PEG 4000, 0,1M Tris pH 8,5 e 0,2M acetato de sódio; **c)** Cristais de YaeQ_{D83A}, que cristalizou em 24% PEG 4000, 0,1M Tris pH 8,5 e 0,2M acetato de sódio.

Tabela 17: Dados estatísticos do processamento dos dados de difração de raios-X do cristal de YaeQ_{D83A}. (Coletados na linha MX1 do LNLS).

Grupo Espacial	P2 ₁
<i>a</i> (Å)	40.109
<i>b</i> (Å)	90.686
<i>c</i> (Å)	48.148
γ	109.894
Dados Estatísticos	YaeQ _{D83A}
Resolução (Å)¹	50.00 – 1.45 (1.50 – 1.45)
N. reflexões observadas	290883
N. reflexões únicas	49744
<I/erro>¹	28.9 (1.50)
Multiplicidade¹	5.8 (2.7)
Completeza (%)¹	86.5 (40.9)
R (%)^{1†}	0.050 (0.480)
N. imagens	345
λ (Å)	1.24

¹Os valores em parênteses se referem aos dados da última camada.

$$^{\dagger}R = (\sum(I - \langle I \rangle)^{**2}) / \sum(I)^{**2}$$

d

Figura 35: Estrutura de YaeQ_{D83A}.

a) Densidade eletrônica ($2F_0 - F_C$, sigma de 1,0) do resíduo A₈₃ da estrutura de YaeQ_{D83A}; **b)** Estrutura de duas subunidades de YaeQ_{D83A} na unidade assimétrica; mapa de densidade eletrônica ($2F_0 - F_C$, sigma de 1,0) do loop que contém os resíduos 61-64 de YaeQ tipo selvagem **(c)** e YaeQ_{D83A} **(d)**.

4.3.7.3 Atividade de Nuclease de YaeQ e dos Mutantes YaeQ_{D68A} e YaeQ_{D83A}

YaeQ_{XAC2396} pertence a família de nucleases PD-(D/E)XK dependentes de magnésio. Esta família inclui enzimas de restrição tipo II como *Bam*HI, *Ngo*MIV entre outras. Para testar a possibilidade de YaeQ_{XAC2396} possuir atividade endonuclease na ausência de outros fatores fizemos ensaios de degradação de DNA utilizando como substrato o DNA do fago lambda.

Observamos atividade nucleásica de YaeQ_{XAC2396}, YaeQ_{D68A} e YaeQ_{D83A} apenas em altas concentrações de proteína, na faixa de 14 μ M (**Figura 36**), mas não observamos atividade na faixa de 14 nM (dados não mostrados). Estes resultados sugerem que a atividade DNAásica sendo observada deva ser proveniente de outra proteína copurificada com YaeQ_{XAC2396} e conseqüentemente copurificada com os mutantes de YaeQ_{XAC2396}, YaeQ_{D68A} e YaeQ_{D83A}, pois os mutantes de YaeQ_{XAC2396} não deveriam conter atividade nucleásica, já que o sítio ativo putativo está corrompido.

Ma	+											
DNA	+	+	+	+	+	+	+	+	+	+	+	+
Mg⁺²	+	+	+	+								
EDTA					+	+	+	+				
Ca⁺²									+	+	+	+
YaeQ_{XAC2396}	+				+				+			
YaeQ_{D68A}				+				+				+
YaeQ_{D83A}				+				+				+

a ***b*** ***c***

Figura 36: Atividade nucleásica de YaeQ e dos mutantes YaeQ_{D68A} e YaeQ_{D83A} em diferentes tempos de reações.

As reações, mostradas na caixa de legenda, foram submetidas a diferentes tempos de reações, **a)** 1 hora de reação, **b)** 2 horas de reação e **c)** 16 horas de reação. Ma: Marcador de peso molecular, DNA: DNA do fago lambda contendo o tampão de reação.

4.3.7.4 Ensaio de Ligação ao DNA Marcado com α - ^{32} P-dATP por “Gel Shift”

YaeQ não interage com dupla fita nem simples fita de DNA. Ensaios *in vitro* de ligação de YaeQ_{XAC2396} à fragmentos de DNA marcados com α -dATP³² foram realizados utilizando a técnica de “Gel-Shift”. Os fragmentos de DNA foram marcados radioativamente durante a reação de PCR pela utilização de α - ^{32} P-dATP (descrito nos Materiais e Métodos).

Ao correr o fragmento de DNA no gel não desnaturante de acrilamida se observou um rastro (**Figura 37 poço 1**). Quando há interação da proteína ao fragmento de DNA esperaria observar uma banda de peso molecular maior, pois o DNA ligado à proteína migra mais lentamente que o DNA livre. No entanto, não foi observado nenhum efeito na migração do DNA radioativo dupla fita (**Figura 37**), simples fita (**Figura 37**) e nem com DNA dupla fita tratado com luz UV (**Figura 37**). Este resultado sugere que YaeQ_{XAC2396} e nem os mutantes de YaeQ_{XAC2396}, conseguem interagir com DNA nas condições testadas.

1μM α ³² -p-dATP-XAC0136 _{promotora} - DNA	+	+		+		+							
10nM α ³² -p-dATP-XAC0136 _{promotora} - DNA		+		+		+							
1μM α ³² -p-dATP-XAC4180 _{promotora} - DNA	+		+		+	+							
10nM α ³² -p-dATP-XAC4180 _{promotora} - DNA			+		+	+							
1μM α ³² -p-dATP-XAC4180 _{promotora} - UV						+	+	+	+				
10nM α ³² -p-dATP-XAC4180 _{promotora} - UV							+	+	+				
1μM α ³² -p-dATP-XAC4180 _{promotora} -ssDNA					+	+	+	+					
10nM α ³² -p-dATP-XAC4180 _{promotora} -ssDNA						+	+	+					
1μM YaeQ _{XAC2396}		+	+	+	+			+	+		+	+	
1μM YaeQ _{D68A}				+	+	+	+		+	+		+	+
1μM YaeQ _{D83A}						+	+	+	+		+	+	

Figura 37: Ensaio de “Gel shift” para verificar a interação de YaeQ, YaeQ_{D68A} e YaeQ_{D83A} com DNA, ssDNA e DNA tratado com UV.

Foram utilizadas diferentes amostras de DNA, DNA dupla fita (DNA), DNA simples fita (ssDNA) e DNA tratado com raios ultra violeta (UV) de regiões promotoras de XAC0136 (XAC0136_{promotora}) e de XAC4180 (XAC4180_{promotora}). Os ensaios foram feitos utilizando 1μM de proteína e duas concentrações de DNA, 1μM e 10 nM. As amostras de DNA marcado com α^{32} -p-dATP foram preparadas conforme descrito nos Materiais e Métodos. As condições utilizadas para realizar este ensaio estão descritas nos Materiais e Métodos.

4.3.8 Ensaio de 2-híbrido usando YaeQ como Isca

YaeQ_{XAC2396} foi utilizada como isca em ensaios de 2-híbrido contra a biblioteca de presas de Xac produzida por nosso grupo^{81,206}. Os resultados estão descritos na **Tabela 18**. Dezenove presas foram derivadas da proteína TolC, um conhecido falso positivo^{81,206}. As outras cinco presas seqüenciadas apenas 1 era um fragmento de DNA que codificava para uma proteína, XAC0129, com início no códon do aminoácido 463. XAC0129 tem 510 aminoácidos, portanto a região que interagiu com YaeQ_{XAC2396} são os últimos 47 aminoácidos. XAC0129 pertence à família AldA que possui atividade cloroacetaldeído dehidrogenase. Até o momento, estes resultados não foram confirmados por outros ensaios independentes e devem ser considerados inconclusivos.

Tabela 18: Resultados do ensaio de 2-híbrido utilizando YaeQ como isca.

Em parênteses se refere ao número do primeiro aminoácido da proteína que interagiu com a isca. TolC é um falso positivo⁸¹.

Isca	Interagiu	Número de clones independentes
YaeQ _{XAC2396}	TolC - XAC3463	19
	XAC0129(463)	1
	Plasmídio fechado (pOAD)	2
	Frame negativo	2

4.3.9 As Proteínas Codificadas pelo mesmo Operon de YaeQ: XAC2397 e XAC2398

Os genes XAC2396, XAC2397 e XAC2398 formam um possível *operon* encontrado em membros do gênero de *Xanthomonas* (vide Introdução). Com a hipótese de que estes genes possam ter alguma função em comum, clonamos, expressamos e fizemos ensaios de interação entre YaeQ_{XAC2396} e XAC2397_{his} (XAC2397 com cauda de histidina fusionada ao N-terminal da proteína) e entre YaeQ_{XAC2396} e FimX_{his-GGDEF-EAL(267-689)}. Até o momento, não obtivemos evidências de interação entre YaeQ_{XAC2396} e XAC2397_{his} ou entre YaeQ_{XAC2396} e FimX_{his-GGDEF-EAL(267-689)}. YaeQ_{XAC2396} não interagiu com XAC2397_{his} em ensaios de “overlay” (**Figura 38 a e b**) tão pouco com FimX_{his-GGDEF-EAL(267-689)} por “pulldown” (FimX_{his-GGDEF-EAL(267-689)} possui cauda de histidina e foi utilizado anticorpo contra cauda de his neste ensaio) (**Figura 38 c**).

Figura 38: YaeQ não interage com XAC2397 nem com FimX_{XAC2398}.

a) Membrana de nitrocelulose corada com “ponceau red” correspondendo a transferência de um gel de SDS-PAGE 18% contendo as seguintes amostras, 1: YaeQ_{XAC2396} pura, 2-3: amostras Induzidas de XAC0258, como controle negativo, 4-5: amostras induzidas de XAC2397, 6: amostra não induzida de XAC2397. **b)** Filme de raios-X da revelação da resina de nitrocelulose com anticorpo contra YaeQ_{XAC2396} em que foi utilizado anti IgG fusionado a Proteína A peroxidase. **c)** Gel de SDS-PAGE 18% correspondente ao experimento de “Pull-Down” entre XAC2398, com cauda de histidina e YaeQ_{XAC2396}, sem cauda de histidina. As frações de 1-4 correspondem a aplicação de apenas YaeQ_{XAC2396} na coluna de afinidade. 1-2: Frações da eluição da coluna de Níquel quando aplicado a amostra pura de YaeQ_{XAC2396}. 3-4: frações da eluição do gradiente da coluna de Níquel, mostrando que YaeQ_{XAC2396} não interage com a resina de níquel da coluna de afinidade. As frações de 5-11 corresponde ao ensaio de pulldown. 5-6 frações da eluição quando aplicado a amostra de YaeQ_{XAC2396} com FimX_{XAC2398} (fragmento de FimX contendo os domínios GGDEF e EAL, FimX_{his-GGDEF-EAL(267-689)}) e 7-11: são as frações do gradiente, mostrando que FimX_{his-GGDEF-EAL(267-689)} não interage com YaeQ_{XAC2396} nestas condições, ou que a interação é tão fraca que não foi possível verificar a banda de YaeQ_{XAC2396} no gel de SDS-PAGE. As setas indicam as proteínas de interesse.

4.3.10 Nocaute apolar do gene de *yaeQ* em *Xac*

O vetor pNPTS138 contém o gene de resistência a kanamicina e possui o gene de *sacB*, o que se torna letal para a célula quando há sacarose no meio de cultura⁵⁷. O gene *sacB* codifica a enzima secretada levansucrase que catalisa a hidrólise da sacarose e sintetiza o levans, que é um polímero de frutose de alta massa molecular²⁰⁷. Em bactérias Gram-negativas a enzima levansucrase é letal e em *E.coli* sua localização preferencial é no periplasma^{208,209}. Uma das explicações para a letalidade da levansucrase em bactérias Gram-negativas seria a obstrução do periplasma da bactéria pelo acúmulo de levans²¹⁰.

O método utilizado para obter células de *Xac* mutantes de *yaeQ*_{XAC2396} consiste basicamente na amplificação de dois fragmentos: a) o primeiro fragmento é a amplificação de 1000 pb que possui 960 pb antes do gene de *yaeQ*_{XAC2396} e 40 pb do gene de *yaeQ*_{XAC2396}; b) o segundo fragmento é a amplificação de 1000 pb que possui 40 pb do final do gene de *yaeQ*_{XAC2396} e mais 960 pb que o prosseguem. Posteriormente, os dois fragmentos de 1000pb são ligados e clonados no vetor pNPTS138 gerando o plasmídeo $\Delta YaeQ_{pNPTS}$.

A troca do gene de *yaeQ*_{XAC2396} cromossomal pelo gene truncado de *yaeQ*_{XAC2396} foi feito por dois eventos de recombinação: 1) Transformação de *Xac* com o vetor $\Delta YaeQ_{pNPTS}$ resultando na inserção por recombinação homóloga do vetor $\Delta YaeQ_{pNPTS}$ dentro do gene de *yaeQ*_{XAC2396} cromossomal. Neste momento, *Xac* transformada deve possuir resistência a kanamicina e sensibilidade a sacarose e conter duas cópias do gene de *yaeQ*_{XAC2396} – um tipo selvagem e outro com a deleção interna; 2) O segundo evento de recombinação é a excisão do vetor agora contendo o gene tipo selvagem de *yaeQ*_{XAC2396}. Este segundo evento de recombinação pode resultar na perda de qualquer uma das cópias do gene, necessitando, portanto, de posterior confirmação por PCR.

A **Figura 39 a** e **b** mostram a obtenção de colônias de *Xac* depois do primeiro evento de recombinação. Os clones que cresceram no meio contendo os antibióticos ampicilina e kanamicina (**Figura 39 a**) e não cresceram em meio contendo ampicilina e sacarose (**Figura 39 b**) foram crescidas na presença de sacarose e kanamicina e subsequente seleção para o segundo evento de recombinação. A **Figura 39 c-e** colônias que cresceram em meio contendo ampicilina e sacarose (**Figura 39 c**) e não

creceram em meio contendo ampicilina e kanamicina (**Figura 39 d**) tão pouco em meio contendo sacarose, kanamicina e ampicilina (**Figura 39 e**) foram selecionadas.

Foram identificados por PCR os clones que não possuíam o gene de *yaeQ*_{XAC2396} tipo selvagem. Para isto foram utilizados o “primer forward” do gene de *yaeQ*_{XAC2396} e “primer reverse” do gene XAC2397 (**Figura 40 a-d**). Aqueles que amplificaram bandas de 1723 pb continham o gene de *yaeQ*_{XAC2396} tipo selvagem enquanto que aqueles que amplificaram bandas de 1223 pb continham o gene de *yaeQ* com a deleção interna. Foram obtidos 28 clones de *Xac* mutantes do gene de *yaeQ*_{XAC2396}.

Figura 39: Obtenção de células de *Xac* nocaute de *yaeQ*.

Primeiro evento de recombinação, **a**) placa de petri contendo LB agar e os antibióticos kanamicina (seleciona células que tiveram a inserção do plasmídeo) e ampicilina (seleciona *Xac*) e **b**) placa de petri contendo LB agar com o antibiótico ampicilina e sacarose (seleção negativa). Nesta etapa, *Xac* deve crescer em meio contendo o antibiótico kanamicina e não em

meio contendo sacarose, pois o vetor pNPTS138 possui o gene de *sacB*, o que se torna letal na presença de sacarose. Os clones que cresceram em meio contendo o antibiótico kanamicina e não cresceram em meio contendo sacarose foram selecionados para o segundo evento de recombinação. *Segundo evento de recombinação.* Células crescidas em meio LB com ampicilina (sem sacarose e sem kanamicina) e depois plaqueadas em triplicata em: **c)** placa de petri contendo LB agar com o antibiótico ampicilina e sacarose, **d)** placa de petri contendo LB agar com os antibióticos kanamicina e ampicilina e **e)** placa de petri contendo LB agar com os antibióticos ampicilina e kanamicina, e sacarose. Nesta etapa, *Xac* deve crescer em meio contendo o antibiótico ampicilina e sacarose e não em meio contendo o antibiótico kanamicina. Os clones que cresceram apenas em meio LB agar com o antibiótico ampicilina e sacarose e não cresceram em meio contendo o antibiótico kanamicina foram selecionados para verificar, por PCR, dentre estes aqueles que eram nocautes ao gene de *yaeQ*_{XAC2396}.

Figura 40: Verificação das células nocautes de *yaeQ*.

a-d) Verificação por PCR das colônias selecionadas para verificar aquelas que eram nocautes ao gene de *yaeQ*. A amplificação foi feita utilizando o “primer forward” do gene de *yaeQ*_{XAC2396} e “primer reverse” do gene XAC2397. Aqueles que amplificaram bandas de 1723 pb contêm o gene de *yaeQ*_{XAC2396} tipo selvagem enquanto que aqueles que amplificaram bandas de 1223 pb não contêm o gene de *yaeQ*_{XAC2396} inteiro, apenas o gene de *yaeQ*_{XAC2396} com uma deleção interna. * clones positivos, ou seja, que contêm o gene de *yaeQ*_{XAC2396} com uma deleção interna.

4.3.11 Mutantes de *YaeQ* apresentam uma maior taxa espontânea de metagênese

Como descrito no item 4.3.6, vários membros da família PD-(D/E)XK estão envolvidos em vias de reparo de DNA. Alguns ensaios biológicos com células de *Xac* e células nocautes de *yaeQ*_{XAC2396} ($\Delta yaeQ$) foram feitos utilizando condições que ocasionam mutações no DNA. As diferentes cepas de *Xac* foram submetidas a raios UV e nenhuma diferencia no grau de sobrevivência foi verificada (dados não mostrados). Analisamos, também, a taxa de aparecimento espontânea de clones resistentes ao antibiótico rifampicina nas cepas *Xac* tipo selvagem e $\Delta yaeQ$. Rifampicina é inibidor da subunidade β da RNA polimerase bacteriana e mutações espontâneas no gene desta subunidade podem gerar cepas resistentes²¹¹. Estes ensaios tinham como objetivo verificar o papel de *YaeQ*_{XAC2396} no processo de reparo de DNA. A **Figura 41** apresenta os resultados de um experimento em que inoculamos meios de cultura com uma colônia isolada de *Xac* tipo selvagem e também com cinco colônias isoladas de *Xac* contendo o gene de *yaeQ*_{XAC2396} com deleção interna (Y1-Y5). Volumes equivalentes destas culturas foram então plaqueados em meio semi-sólido contendo quantidades crescentes de rifampicina. Como pode ser observado na **Figura 41** as culturas derivadas das colônias Y1 e Y4 apresentaram a mesma sensibilidade às células tipo selvagem: um pequeno número de colônias cresceram em 1 μ M rifampicina e nenhuma cresceu em concentrações maiores. Por outro lado, as culturas Y2, Y3 e Y5 geraram um número grande de células resistentes a concentrações elevadas (20 μ M) do antibiótico. Esta variabilidade de fenótipo entre as culturas Y1, Y4 com as culturas Y2, Y3 e Y5 é esperado já que o evento de mutação espontânea é aleatória, podendo ou não afetar o gene que codifica a subunidade β da RNA polimerase. Estes resultados preliminares sugerem que as células $\Delta yaeQ$ sofrem uma maior taxa de mutação do que as células tipo selvagem que levanta a hipótese de que *YaeQ* possa ter um papel em alguma via de reparo de DNA em *Xac* e em outras bactérias.

Figura 41: Crescimento de células de *Xac* e de $\Delta yaeQ$ em meio contendo rifampicina.

Células de *Xac* tipo selvagem (wt) e $\Delta yaeQ$ (Y1-Y5) foram plaqueadas em meio LB contendo 1μM (a) 10μM (b) e 20μM (c) de rifampicina. Como pode ser observado, células $\Delta yaeQ$ (Y2, Y3 e Y5) tiveram maior taxa de crescimento quando comparado com células tipo selvagem.

4.3.12 Determinação da estrutura de YaeQ em diferentes pHs

YaeQ_{XAC2396} foi cristalizada em diferentes pHs, desde pH ácido, pH 5.6, quanto em pH fisiológico, pH 7.1, e em pH básico, pH 8.5. Os dados estatísticos do processamento dos dados de difração de cada cristal de YaeQ_{XAC2396} crescidas em diferentes pHs estão mostrados na **Tabela 19**. As fases dos cristais de YaeQ_{XAC2396} crescidas em diferentes pHs foram calculadas pela técnica de substituição molecular usando YaeQ_{XAC2396} como modelo (PDBID: 3G3W). A estrutura de YaeQ_{XAC2396} para cada pH estão mostradas na **Figura 42**.

Interessantemente, os cristais de YaeQ_{XAC2396} produzidos em pH neutro e básico, pH 7.1 e 8.5, contêm duas moléculas de YaeQ_{XAC2396} na unidade assimétrica enquanto que, os cristais de YaeQ_{XAC2396} que cristalizaram em pH ácido, pH 5.6, contêm três moléculas de YaeQ_{XAC2396} na unidade assimétrica (**Figura 42**). Estas estruturas não foram refinadas, e portanto não foi feita uma análise mais profunda dos efeitos do pH na estrutura de YaeQ_{XAC2396}.

Tabela 19: Dados estatísticos do processamento de dados de difração de raios-X de YaeQ em diferentes pHs.

pHs	8.5	7.1	5.6
Grupo Espacial	P2 ₁	P2 ₁	P2 ₁
<i>a</i> (Å)	39.164	38.969	38.186
<i>b</i> (Å)	90.923	89.647	139.881
<i>c</i> (Å)	47.422	47.211	46.128
γ	109.044	108.702	110.014
Dados Estatísticos	Nativo	Nativo	Nativo
Resolução (Å)¹	40.00 - 2.20 (2.07 - 2.00)	40.00 - 2.95 (3.06 - 2.95)	40.00 - 2.15 (2.23 - 2.15)
N. reflexões observadas	41033	18529	33991
N. reflexões únicas	13135	6068	20005
<I/erro>¹	22.0 (2.0)	9.2 (2.4)	10.0 (1.2)
Multiplicidade¹	3.1 (1.5)	3.1 (2.0)	1.7 (1.6)
Completeza (%)¹	61.7 (7.9)	93.1 (56.3)	80.5 (63.0)
R (%)¹¹	0.119 (0.212)	0.108 (0.262)	0.058 (0.311)
N. imagens	200	158	99
λ (Å)	1.425	1.425	1.425

¹Os valores em parênteses se referem aos dados da última camada.

$$^1R = (\sum(I - \langle I \rangle)^{**2}) / \sum(I)^{**2}$$

Figura 42: Estrutura de YaeQ em diferentes pHs.

Estrutura de YaeQ_{XAC2396} cristalizada em pH 5.6 (**a**), em pH 7.1 (**b**) e em pH 8.5 (**c**). Há duas moléculas de YaeQ_{XAC2396} na unidade assimétrica quando YaeQ_{XAC2396} é cristalizada em pH 7.1 e em pH 8.5, mas há três moléculas de YaeQ_{XAC2396} na unidade assimétrica quando YaeQ_{XAC2396} é cristalizada em pH ácido, pH 5.6.

4.4 Discussão

YaeQ é altamente conservada entre proteobacterias e possui quatro motivos de sequência conservados. YaeQ_{XAC2396} é uma proteína altamente conservada em uma gama enorme de bactérias (222 espécies), e em alguns organismos eucarióticos (3 espécies). A presença de YaeQ em tantas bactérias sugere que YaeQ_{XAC2396} deva exercer alguma função importante e abrangente.

A análise de conservação de aminoácidos revelou que a família de YaeQ pode ser caracterizada pela presença de quatro motivos: 1- *Motivo AXHPSE*; 2- *Motivo ácido DEP(D/E)*; 3- *G85* pertencente ao loop $\beta 5$ - $\alpha 2$ e 4- *R148* (**Figura suplementar 2D-F do Anexo 4**).

Determinamos a estrutura tridimensional do primeiro membro da família de YaeQ, YaeQ_{XAC2396}, utilizando a técnica de difração de raios-X. Interessantemente todos os resíduos que compõem os motivos conservados na família de YaeQ estão localizados em uma face da estrutura (**Figura 21 e Figuras suplementares 1B do Anexo 4**). Esta observação sugere que esta seja a face de interação de YaeQ com outros alvos moleculares: proteína e/ou substrato.

YaeQ pertence a família PD-(D/E)XK de endonucleases. A comparação da estrutura tridimensional de YaeQ_{XAC2396} com outras estruturas depositadas no PDB revelou que YaeQ adota uma estrutura que é uma variação do motivo PD-(D/E)XK de nucleases dependentes de magnésio (**Tabela 1 suplementar, Figuras 3 do Anexo 4**). A família de homólogos de YaeQ não possui a lisina do motivo PD-(D/E)XK conservada, na sua posição há a glicina 85, absolutamente conservada. Além da G₈₅ os resíduos W₈₁, P₈₇ e a R₉₁ são conservados na família de YaeQ e fazem interações com o motivo P₆₇(D/E)₆₈-X(D/E)₈₃ (**Figuras 3 do Anexo 4**), e podem ser importantes para a sua atividade catalítica (ainda não caracterizada). Desta forma, o sítio ativo putativo de YaeQ_{XAC2396} é melhor descrito pelo seguinte motivo: P₆₇(D/E)₆₈-W₈₁X(D/E)₈₃XG₈₅XP₈₇X₃R₉₁.

Mecanismo de clivagem do DNA pela família de PD-(D/E)XK endonucleases – BamHI como exemplo. O sítio ativo de *BamHI* é composto pelos resíduos I₉₃D₉₄-E₁₁₁XE₁₁₃. O D₉₄, E₁₁₁ e o grupo carbonila da cadeia principal da F₁₁₂ estão envolvidos na interação aos cátions de magnésio (**Figura 3B do Anexo 4**) que por sua vez interagem com o fosfato que será clivado. O E₁₁₃, correspondendo a lisina do motivo PD-(D/E)XK, interage com uma molécula de água que irá atacar

nucleofílicamente o fosfato a ser clivado. Logo, o E₁₁₃ de *BamHI* poderia ser considerada como uma base geral.

A superposição do sítio ativo de *BamHI* com os correspondentes na estrutura de YaeQ_{XAC2396} mostrou que I₉₃D₉₄-E₁₁₁XE₁₁₃ de *BamHI* livre de DNA se sobrepõe muito bem com o sítio catalítico putativo de YaeQ_{XAC2396}, P₆₇(D/E)₆₈-(D/E)₈₃XG₈₅XR₉₁. A R₉₁ e R₃₉ de YaeQ_{XAC2396} se sobrepõe com as K₁₂₆ e K₆₁, respectivamente, de *BamHI* livre de DNA (**Figura 3C** do **Anexo 4**). Estas lisinas interagem com o fosfato anterior e posterior ao do fosfato que será clivado, provavelmente estabilizando-os por interações eletrostáticas.

A **Figura 43** mostra a superposição da estrutura de *BamHI* (ligada ao cálcio e em complexo com o DNA na conformação pré-clivado) com a estrutura de YaeQ_{XAC2396}. Esta análise mostra que o loop de YaeQ_{XAC2396} que contém o motivo AXHPSE (colorida em azul na **Figura 43**), está posicionado em direção a cavidade menor da dupla fita de DNA da estrutura de *BamHI* (discutido com mais detalhe a seguir).

Figura 43: Superposição de YaeQ com *BamHI* em complexo com DNA.

a) Superposição da estrutura de *BamHI* (PDBID: 2BAM) ligada ao DNA com a estrutura de YaeQ_{XAC2396} (PDBID: 2G3W). YaeQ_{XAC2396} está colorida em vermelho e *BamHI* em laranja, em violeta está colorido o loop de *BamHI* que corresponde ao loop ácido de YaeQ_{XAC2396} (verde). **b)** A estrutura de YaeQ_{XAC2396} com o DNA de *BamHI* e **c)** Superfície de YaeQ_{XAC2396} com o DNA de *BamHI*. Em branco está colorido o sítio catalítico de YaeQ_{XAC2396} (D₆₈, E₈₃, R₃₉ e R₉₁), em azul o motivo AXHPSE e em verde o loop que contém o motivo ácido DEPE. A barra colorida em preto representa o fosfato que é clivado por *BamHI*. **d)** Estrutura de YaeQ_{XAC2396}, colorida em verde, com o DNA de *BamHI* (PDBID: 2BAM), colorido em rosa, indicando o motivo AXHSPE em azul e os resíduos D₆₈, D₈₃, R₃₉ e R₉₁, pertencentes ao sítio ativo de YaeQ_{XAC2396}, em vermelho.

Possíveis funções biológicas para YaeQ. Membros da família PD-(D/E)XK possuem baixa ou quase nenhuma similaridade em seqüência primária, e novos membros só podem ser identificados a partir da comparação estrutural entre membros desta família. A similaridade estrutural ocorre apenas no sítio ativo, que é composto principalmente pelos dois resíduos acídicos do motivo PD-(D/E)XK envolvidos na coordenação a íons de magnésio que interagem com o fosfato do DNA ou RNA que serão clivados. A lisina do motivo PD-(D/E)XK ativa a molécula de água que irá atacar nucleofilicamente o fosfato da cadeia principal do nucleotídeo.

Desta forma, existe uma vasta gama de proteínas com diferentes domínios de interação a RNA ou a DNA que pertencem a família PD-(D/E)XK. A maioria tem como função biológica clivar a ligação fosfodiéster da cadeia principal do RNA ou DNA. A diferença funcional entre membros da família PD-(D/E)XK vem na natureza da estrutura de RNA ou DNA que estas enzimas interagem e clivam. Alguns exemplos de classes de proteínas pertencentes a esta família são as enzimas de restrição tipo II, proteínas envolvidas na via de reparo de DNA, “Holliday junction resolvases”, replicação e recombinação. Desta forma, é difícil prever em qual destas classes de funções YaeQ pertence.

Dentre as possíveis funções que YaeQ_{XAC2396} poderia desempenhar, com base nas funções descritas anteriormente para proteínas da família PD-(D/E)XK, acreditamos que YaeQ_{XAC2396} não é uma enzima de restrição, já que enzimas de restrição não são conservadas em diferentes espécies bacterianas.

XPF (xeroderma pigmentosa complementation group F, também conhecida como ERCC4), encontrada em organismos eucarióticos e em archaeas¹²⁰, liga ao ERCC1 (excision repair cross complementarity group 1) para formar um complexo com atividade endonucleásica que reconhece principalmente dupla fita de DNA adjacente a uma ponta 3' DNA simples fita²¹². XPF-ERCC1 tem papel importante no processo de reparo de DNA por excisão de nucleotídeo, recombinação e integridade do DNA genômico²¹². XPF é uma variação do motivo **PD-(D/E)XK** encontrado em endonucleases com sítio catalítico composto pelos resíduos G₃₇**D**₃₈-**E**₄₈R₄₉**K**₅₀X₃D₅₄, em que os resíduos destacados em negrito correspondem aos resíduos do motivo **PD-(D/E)XK** (vide Fig. 30 b). Os aminoácidos G₃₇D₃₈-E₄₈R₄₉K₅₀ estão presentes em uma fita β e o D₅₄ se localiza em uma hélice característica desta família. YaeQ_{XAC2396} não possui esta hélice e portanto o D₅₄ não se sobrepõe a nenhum aminoácido na

estrutura de YaeQ_{XAC2396} (**Figura 30 b**, proteína *Hef* com PDBID: 1J22). Desta forma, YaeQ_{XAC2396} tão pouco deve pertencer a família de XPF.

Alguns ensaios biológicos foram feitos a fim de identificar a função biológica de YaeQ, como interação de YaeQ com DNA dupla fita e simples fita e DNA tratada com raios UV. Nenhum destes ensaios forneceram resultados positivos, sugerindo que YaeQ não interage com DNA inespecífico nas condições testadas (**Figura 37**). Ensaios de degradação de DNA foram inconclusivos, já que YaeQ tipo selvagem e também os mutantes YaeQ_{D68A} e YaeQ_{D83A} clivam DNA somente quando está em altas concentrações. Este resultado sugere que a atividade DNAásica sendo observada seja de algum contaminante co-purificado com YaeQ e seus mutantes (**Figura 36**). Ensaios de 2-híbridos utilizando YaeQ_{XAC2396} como isca contra uma biblioteca genômica de *Xac* não forneceu nenhum resultado conclusivo (**Tabela 18**).

Os genes localizados próximos ao gene de *yaeQ* em *Xac*, XAC2397 e FimX_{XAC2398}, pertencentes a um possível *operon*, foram expressos e ensaios de interação proteína-proteína foram realizados. No entanto, YaeQ_{XAC2396} não interagiu com XAC2397 e nem com FimX_{XAC2398} nas condições testadas (**Figura 38**).

YaeQ pode estar envolvida na via de reparo de DNA. O resultado funcional mais interessante obtido foi do ensaio utilizando células nocautes de *yaeQ*_{XAC2396}. Este experimento mostrou que cepas de *Xac* $\Delta yaeQ_{XAC2396}$ possuem maior taxa de sobrevivência na presença do antibiótico rifampicina quando comparadas com a cepa de *Xac* tipo selvagem (**Figura 41**). O antibiótico rifampicina inibe a atividade da RNA polimerase (RNAP) por interagir com a subunidade β da RNAP²¹¹. A rifampicina liga em uma cavidade profunda da RNAP dentro do canal DNA/RNA, mas 12Å de distância do sítio ativo da RNAP, bloqueando o caminho de alongação do RNA quando a transcrição tem de 2 a 3 nucleotídeos de comprimento²¹¹. São conhecidos 23 posições na RNAP essenciais para que a rifampicina consiga inibir a RNAP, sendo que 12 delas fazem contatos diretos com o antibiótico²¹¹. Mutações que afetam estas posições resultam no surgimento de cepas resistentes.

Este resultado mostra que YaeQ_{XAC2396} provavelmente está envolvida em algum processo de reparo de DNA. Uma das hipóteses é que YaeQ_{XAC2396} esteja envolvida no processo de reparo de DNA por pareamento de base errada, esquematizado na **Figura 44** e descrita na legenda da figura.

Figura 44: Mecanismo de reparo de DNA por pareamento incorreto.

MutS reconhece o pareamento incorreto ou a inserção ou deleção de 1-4 nucleotídeos na fita filha do DNA, MutH liga a fita hemimetilada na sequência GATC e é ativada pela interação com MutS que ocorre apenas na presença de MutL. MutH cliva a fita filha ocasionando uma quebra. Posteriormente, a DNA helicase separa as duas fitas de DNA (fita metilada da fita não metilada (fita filha)) a partir da ação de exonucleases que degradam o DNA na direção do pareamento incorreto (notar que exonucleases diferentes (5'→3' ou 3'→5') são utilizadas de forma dependente da posição relativa de MutH e MutSL. A DNA polimerase III junto com SSB ("Single Strand Binding" protein) preenchem a região do DNA filha removida e a ligase liga a fita recém sintetizada com a fita filha antiga^{196,213-216}.

A proteína MutH também pertencente a família PD-(D/E)XK^{196,215} e está envolvida no reconhecimento e na clivagem do DNA hemimetilado durante o processo de reparo de DNA por pareamento de base errada¹⁹⁶. A clivagem do DNA por MutH é ativada pela interação com MutL do complexo MutL-MutS. Interessantemente, o genoma de *Xac* não codifica para um ortólogo de MutH (**Tabela 20**). Uma hipótese a ser testada é que YaeQ_{XAC2396} possa desempenhar uma função correlacionada á de MutH, pois vários organismo, incluindo *Xac*, possuem MutL e MutS mas não codificam a um homólogo de MutH (**Tabela 20**).

O banco de dados Pfam lista 272 cepas correspondendo a 142 espécies que codificam a um MutH enquanto YaeQ é encontrada 370 cepas de 222 espécies¹²⁰. A **Tabela 20** resume uma análise da distribuição dos genes *yaeQ* e *mutH* nos genomas seqüenciados até agora.

Tabela 20: Árvore de espécies mostrando a ocorrência dos genes *yaeQ*, *mutH* e *vsr* em diferentes espécies¹²⁰.

	YaeQ	MutH	Vsr
Total espécies	222	142	156
Total cepas/Isoladas	370	272	296
EUCARIOTOS	3	0	0
BACTÉRIA	367	272	296
-Actinobacteria	8	5	45
---Mycobacteria	7	0	4
---Nocardiaceae	1	0	1
---Coriobacteridae	0	1	5
---Bifidobacteriaceae	0	4	14
---Streptomyetaceae	0	0	3
---Micrococcineae	0	0	5
-Tenericutes	0	1	0
-Proteobacteria	355	233	187
---alphaproteobacteria	14	0	39
-----Rhodobacterales	12	0	11
-----Caulobacteracea	0	0	5
---Deltaproteobacteria	20	1	8
---Betaproteobacteria	84	1	29
-----Burkholderiales	75	0	15
-----Neisseriales	3	1	10
---Zetaproteobacteria	1	0	0
---Epsilonproteobacteria	1	1	1
---Gamaproteobacteria	229	230	109
-----Aeromonadales	3	3	0
-----Alteromonadales	35	31	3
-----Vibionales	22	40	0
-----Xanthomonadales	11	0	1
-----Enterobacterales	118	121	91
-----Pseudomonales	28	0	3
-----Oceanospirillales	7	0	2

----- <i>Pateurellales</i>	0	25	2
----- <i>Legionellales</i>	0	4	3
-Firmicutes	0	32	28
-Cyanobacteria	0	0	6
-Acidobacteria	0	0	2
-Verrucomicrobia	0	0	2
-Chlorobi	0	0	3
ARCHAEA	0	1	3

O sítio de interação de MthH de *E.coli* ao DNA é metilado pela enzima Dam metilase²¹⁷, na posição N⁶ da adenina da sequência GATC²¹⁸⁻²²⁰. O genoma de *Xac* não codifica a nenhum ortólogo da Dam metilase. Isso significa que se YaeQ desempenhasse uma função similar ao MthH o sítio de interação de YaeQ_{XAC2396} ao DNA seria outro.

Outra enzima envolvida na metilação do DNA é o Dcm²²¹, que metila a posição C⁵ da citosina interna da sequência CC(A/T)GG²¹⁸⁻²²⁰. A metilação da citosina pelo DCM aumenta a espontaneidade da deaminação da ^{5me}C que se converte na base timina. Desta forma, o par de base ^{5me}C:G se converte espontaneamente para T:G²²². Este pareamento de base errado é corrigido pela endonuclease vsr (Very-short-patch repair) que reconhece a sequência, CT(T/A)GG pareada com a GG(A/T)CC, e cliva a posição 5' da timina do pareamento errado, T:G²²³. Aparentemente MutL e MutS são as proteínas que reconhecem o pareamento de base errado e apresenta para o vsr, pois MutL aumenta a afinidade do vsr ao DNA²²⁴.

A maioria dos organismos que contem Dcm também contem a endonuclease vsr de tal forma que, o códon de terminação do gene dcm se sobrepõe ao códon iniciador do vsr²²⁵. Provavelmente os dois genes são transcritos no mesmo RNAm. O papel do Dcm em *E. coli* é desconhecido, mas homólogos do *dcm-vsr* são encontrados em várias espécies de bactérias. Em algumas delas, a metilação por Dcm está envolvida na proteção da auto-clivagem por enzimas de restrição²²⁶.

A endonuclease Vsr também pertence à família PD-(D/E)XK²²⁷. A análise genômica de *Xac* mostrou que XAC2213, de 487 aa, está anotada como Dcm, mas sua identidade com Dcm de *E.coli*, que possui 472 aa, é de apenas 28%. Além do mais, apenas no aminoácido 65 de XAC2213 começa a ter similaridade com o

aminoácido 158 do Dcm de *E.coli*. Mas é interessante o fato de XAC2213 conter na região N-terminal (11-208) um domínio anotado pelo banco de dados Pfam como DNA metilase de C-5 citosina¹²⁰. Uma outra observação interessante é que o gene de XAC2213 tem no seu lado genes para um topoisomerase do tipo III (XAC2212) e uma proteína SSB (single-strand binding protein, XAC2211), ambos podem participar em processos de reparo de DNA e recombinação. Sem ensaios funcionais uma dúvida existe sobre a função desempenhada pela XAC2213 e se ela é um verdadeiro ortólogo de Dcm.

Caso XAC2213 desempenhe uma função correlacionada ao Dcm, YaeQ poderia desempenhar uma função parecida ao Vsr, pois ambas pertencem a família PD-(D/E)XK^{227,228}. Além disso, assim como em *Xac* vários outros organismos contem o gene de *yaeQ* mas não possuem o gene *vsr* (**Tabela 20**).

Existem outras vias de reparo de DNA essenciais para as bactérias, como a via de reparo de DNA ocasionada pela deaminação espontânea dos nucleotídeos adenina e citosina, a perda das purinas que também ocorre espontaneamente, alquilação e oxidação dos nucleotídeos ocasionados por fatores endógenos, entre outros²²⁹. Todos estes tipos de modificações na estrutura do DNA quando não corrigidos podem levar a mutações no DNA e desta forma, necessitam ser reparados. YaeQ pode portanto, pertencer a qualquer uma destas vias de reparo. Mais experimentos precisam ser desenhados para desvendar a função desta nova família de proteínas.

O motivo conservado AXHPSE pode estar envolvido na interação com a cavidade menor do DNA. As estruturas de MutH, Vsr e *BamHI* já foram resolvidas em complexo com dupla fita de DNA. Todas estas proteínas pertencem a família de endonucleases com motivo PD-(D/E)XK. A sobreposição da estrutura de YaeQ_{XAC2396} com MutH (PDBID: 2AOR), *BamHI* (PDBID: 2BAM) e com Vsr (PDBID: 1ODG), estão apresentadas na **Figura 45**. Uma observação interessante é que a alça que contem os resíduos do motivo conservado AXHPSE na estrutura de YaeQ_{XAC2396} coincidentemente entra no sulco menor do DNA quando sobreposto com as estruturas de MutH, Vsr e *BamHI* (**Figura 45**). Esta análise sugere que o motivo altamente conservado AXHPSE pode desempenhar a função de contatar o sulco menor do complexo YaeQ-DNA. Assim sendo, estes resíduos poderiam, também, estar envolvidos com a especificidade de YaeQ_{XAC2396} com o seu substrato, ainda desconhecido.

Figura 45: Sobreposição estrutural de proteínas pertencentes a família PD-(D/E)XK, MutH, vsr e BamHI.

Sobreposição da estrutura de YaeQ (PDBID: 2G3W), colorida em rosa claro, com as estruturas das proteínas MutH (PDBID: 2AOR) (**a**), vsr (PDBID: 1ODG) (**b**) e BamHI (PDBID: 2BAM) (**c**), coloridas em azul claro. As estruturas de DNA estão coloridas em laranja, e o loop AXHPSE da estrutura de YaeQ_{XAC2396} está colorida em azul escuro. O loop da estrutura de MutH, composto pelos resíduos D₄₃, L₄₄, K₄₅, R₄₆, D₄₇ e K₄₈, envolvido na interação a cavidade menor da estrutura do DNA está colorido em verde escuro no item **a**. Em todos os casos, o loop AXHPSE da estrutura de YaeQ coincidentemente contacta a cavidade menor das estruturas de DNA das proteínas MutH (**a**), vsr (**b**) e BamHI (**c**). Esta análise sugere que o loop AXHPSE na estrutura de YaeQ esteja envolvido na interação e na especificidade de YaeQ ao seu substrato. A seta destaca a localização do loop AXHPSE, colorido em azul, na estrutura de YaeQ_{XAC2396}.

4.5 Conclusões

Determinamos a estrutura tridimensional do primeiro membro da família de YaeQ, YaeQ_{XAC2396}. Desde então, outros grupos resolveram as estruturas de homólogos em *P. syringae* (PDBID: 2OT9) e em *E. coli* (PDBD: 3C0U). A estrutura de YaeQ é composta por uma hélice central envolvida por duas folhas betas que formam quase um barril fechado. Há duas hélices externas a esse barril que formam uma topologia similar ao Rossmann-fold, β - α - β - α - β .

A comparação da estrutura de YaeQ_{XAC2396} com outras estruturas protéicas depositadas no PDB mostrou que YaeQ é um novo padrão de enovelamento. Esta conclusão foi feita com base na sobreposição da estrutura de YaeQ_{XAC2396} com todas as estruturas que possuíam similaridade estrutural com YaeQ_{XAC2396}. Apenas metade da estrutura de YaeQ, a primeira folha beta e a hélice central, se sobrepõe com estas estruturas.

Interessantemente, todas as proteínas com similaridade estrutural com YaeQ_{XAC2396} pertencem a família de endonuclease PD-(D/E)XK. Membros desta família possuem apenas o sítio catalítico conservado espacialmente. Desta forma, novos membros só podem ser identificados através de informações estruturais. As funções desses membros variam de forma dependente da natureza do nucleotídeo que estas proteínas interagem. A única similaridade funcional de membros desta família é a atividade nucleásica. Verificamos que YaeQ_{XAC2396} é um novo membro da família PD-(D/E)XK em que a K é substituída pela G₈₅, absolutamente conservada na família de YaeQ_{XAC2396}. O papel desta glicina, aparentemente, é de manter a estrutura do loop que conecta vários resíduos do sítio putativo catalítico de YaeQ, P₆₇(D/E)₆₈-W₈₁X(D/E)₈₃XG₈₅XP₈₇X₃R₉₁.

A família de YaeQ é altamente conservada entre proteobactérias, mas também pode ser encontrada em alguns organismos eucarióticos. Membros da família de YaeQ são caracterizados pela presença de quatro motivos de sequência conservados: 1- *Motivo AXHPSE*; 2- *Motivo ácido DEP(D/E)*; 3- G₈₅ pertencente ao loop β 5- α 2 e 4- R₁₄₈. Todos estes motivos localizam-se em uma face da estrutura de YaeQ. Isto sugere que, estes resíduos participem da interface de interação de YaeQ com outras proteínas ou com o seu substrato, ainda desconhecido.

A sobreposição da estrutura de YaeQ com outros membros da família PD-(D/E)XK, mostrou que o motivo altamente conservado AXHPSE pode desempenhar a função de contatar a cavidade menor do DNA. Assim sendo, estes resíduos poderiam, também, estar envolvidos com a especificidade de YaeQ_{XAC2396} com o seu substrato.

Células $\Delta yaeQ_{XAC2396}$ sofrem maior taxa de mutação espontânea. Esta observação levanta a hipótese de que YaeQ possa desempenhar um papel importante na via de reparo de DNA, consistente com as funções de outros membros da família PD-(D/E)XK. Dentre as vias de reparo de DNA que envolve membros da família PD-(D/E)XK tem vias de reparo por pareamento de base errada²²⁹. Duas destas vias utilizam as enzimas MutL e MutS em conjunto com uma nuclease da família PD-(D/E)XK envolvidas na discriminação da fita mãe e filha: MutH ou Vsr^{214,224}. O genoma de *Xac* codifica para as proteínas MutL e MutS, mas não para MutH ou para Vsr. Esta observação levanta as hipóteses, a serem testadas, de que YaeQ poderia atuar de forma similar a MutH ou Vsr em *Xac*.

Vale a pena ressaltar, que apesar do genoma de *Xac* não codificar as enzimas Dam e Dcm metilases, *Xac* codifica três proteínas, XAC2639, XAC2213 e XAC2436, anotadas como possíveis proteínas envolvidas na metilação do DNA. Desta forma, existe a possibilidade de que YaeQ participe de uma nova via de reparo de DNA de tal forma que, o sítio de interação de YaeQ ao DNA é metilado por uma enzima ainda desconhecida.

CAPÍTULO 5
Outros Alvos Estruturais:
SufE está envolvida na Biossíntese de Centros de Ferro-
Enxofre (Fe-S)
&
Sistema de “Quorum Sensing” em Xanthomonas
axonopodis pv citri

5.1 *SufE: componente do sistema cisteína dessulfarase envolvida na Biossíntese de Centros de Ferro-Enxofre*

5.1.1 *Introdução*

Função de centros de Fe-S: Centros de ferro-enxofre (Fe-S) estão envolvidos em vários processos biológicos importantes incluindo respiração, fotossíntese, regulação de transcrição, reparo de DNA, fixação de nitrogênio, ciclo do ácido cítrico, e na biossíntese de aminoácidos e de diferentes cofatores como os grupos haem, biotina e molibdênio^{230-233, 234}. O centro de Fe-S nestes processos pode assumir diferentes funções, que vai desde um excelente doador e receptor de elétrons nos processos de respiração e fotossíntese, até como ácido de Lewis no processo de catálise enzimática. Centros de Fe-S podem também servir como doadores de íons de enxofre como ocorre com as enzimas biotina sintetase e lipoato sintetase²³⁵ e como sensores das condições ambientais ou intracelulares em processos de regulação gênica, como ocorre com os fatores de transcrição FNR, IscR e SoxR²³⁶. Recentemente foi levantada a hipótese do centro de Fe-S ter um papel puramente estrutural, como foi sugerido para a proteína DNA helicase dependente de ATP envolvida em reparo de DNA^{234,237}. Portanto, a importância do papel biológico destes centros é inquestionável e conseqüentemente, são encontrados em todos os seres vivos²³⁸.

Os centros de Fe-S mais comumente encontrados estão na forma rômboico [2Fe–2S] e cubano [4Fe–4S], mas podem ser encontrados também na forma [3Fe–4S]. Os íons de ferro podem estar em dois estados de oxidação, na forma reduzida, Fe⁺² ou na forma oxidada, Fe⁺³. As proteínas coordenam estes centros normalmente com as cadeias laterais de resíduos de cisteínas ou histidinas e em casos raros, pela cadeia lateral dos resíduos de aspartato, arginina e serina entre outros^{238,239}.

Sistemas de síntese de centros de Fe-S: Existem três sistemas de síntese de centros de Fe-S encontrados em organismos procarióticos: *NIF*, *ISC* e *SUF* (**Figura 46**). O sistema *NIF* “Nitrogen Fixation System” foi considerado ter um papel exclusivo na maturação de enzimas bacterianas envolvidas na fixação de nitrogênio, mas também pode ser encontrado em alguns organismos anaeróbicos sem atividade nitrogenase²⁴⁰. O sistema *ISC* “Iron Sulfur Cluster” foi inicialmente identificado em *Azotobacter vinelandii* pela sua similaridade com o sistema *NIF* e posteriormente

caracterizado como um novo sistema. O sistema *SUF* “Sulfur Assimilation” foi inicialmente identificado por conter genes envolvidos na virulência do fitopatógeno *Erwinia chrysanthemi* e, posteriormente como participante na manutenção dos centros Fe-S sob estresse oxidativo e em condições de limitações de ferro em células e *E. coli*^{241,242}.

Os três sistemas de síntese de centros de Fe-S compartilham do mesmo mecanismo de captação de átomos de enxofre pelas proteínas cisteínas dessulfarases, SufS, IscS, NifS. As cisteínas dessulfarases transferem os átomos de enxofre para a proteína receptora de enxofre que pode montar o centro de Fe-S, no caso dos sistemas *ISC* e *NIF*, pelas proteínas IscU e NifU, ou não, como no caso do sistema *SUF* pela proteína SufE. SufE transfere o enxofre para o sistema ATPásico SufBCD onde o centro de Fe-S é formado (**Figura 46**). Todos estes sistemas também possuem a proteína montadora de centros de Fe-S SufA, IscA ou NifA. A seguir há uma descrição mais detalhada do pouco que conhece sobre o mecanismo de montagem de centros de Fe-S pelo sistema *SUF*.

Figura 46: Organização gênica dos sistemas *ISC* e *SUF* de *E. coli* e do sistema *NIF* da bactéria fixadora de nitrogênio *A. vinelandii* e sua comparação com proteínas homólogas encontradas em *Xac*.

O sistema *NIF* é encontrado principalmente em bactérias de fixação de nitrogênio. *E. coli* possui tanto o sistema *ISC* quanto o *SUF*. No entanto, nesta bactéria o sistema *SUF* é apenas expresso em condições de estresse oxidativo e em condições de limitação de íons de ferro. Proteínas com a mesma cor provavelmente possuem a mesma função, mostrando a relação evolutiva entre estes sistemas. Diagrama do manuscrito publicado por Goldsmith-Fischman e colaboradores²⁴³.

Sistema *SUF*: O sistema *SUF* em *E. coli* é codificado por seis genes: *sufABCDSE* (**Figura 46**). SufS assim como IscS e CsdA, é uma cisteína dessulfarase dependente de piridoxal-fosfato, que catalisa a decomposição de L-cisteína à L-alanina. SufS sozinha tem alta atividade selenocisteína deselenase e baixa atividade como cisteína dessulfarase *in vitro*²⁴⁴. No entanto, a atividade de cisteína dessulfarase é aumentada drasticamente na presença de SufE (superando a atividade selenocisteína deselenase) e aumentada ainda mais na presença concomitante de SufE e SufBCD²⁴⁴. SufBCD formam um complexo ATPase que pode estar atuando como um fornecedor de energia para a construção do centro de Fe-S²⁴⁵.

SufA tem similaridade em sequência primária com NifA e IscA e possivelmente regula a homeostase de ferro atuando como sensora ao estresse oxidativo como descrito em *cyanobacteria*²⁴⁶. SufA de *E. coli* recebe íons de enxofre (S) de SufE para formar os centros de Fe-S, [2Fe-2S] e [4Fe-4S], e o mecanismo de montagem sugerido é diferente do descrito para IscA, em que SufA necessita primeiro da transferência de íons de S de SufE para depois coordenar os íons de Fe²⁴⁷.

As ferropoteínas SufR e IscR são repressores de transcrição do sistema *SUF* e *ISC*, respectivamente²³⁴. IscR, na forma apo, em condições de deficiência de íons de ferro e em estresse oxidativo, atua como um fator de transcrição do *operon SUF*²³⁴, demonstrando uma coordenação funcional entre o sistema *ISC* com o sistema *SUF*.

Mecanismo de formação de centros de Fe-S pelo sistema *SUF*: O primeiro etapa na formação do centro Fe-S é a produção de íons S²⁻ por uma cisteína dessulfarase. No sistema *SUF*, esta função é desempenhada pelas proteínas SufS e SufE, esquematizada na **Figura 47**. O primeiro passo é a retirada de íons de enxofre da cadeia lateral de aminoácidos de cisteína. A interação de SufS com SufE permite a transferência do enxofre de SufS para SufE. Por isso, o complexo SufS-SufE é conhecida como um sistema de dois componentes com atividade cisteína dessulfarase. Este processo envolve a formação de grupos persulfetos nas cadeias laterais de cisteínas altamente conservadas nas duas enzimas (**Figura 47**).

Figura 47: Sistema de cisteína dessulfarase catalisada por SufS-SufE. N-CH₂-SH corresponde a uma cisteína e R-CH₃ a uma alanina²⁴⁴.

Após SufE receber o grupo tiol de SufS, SufE o transfere para SufB, apenas na presença de SufC²⁴⁸. SufB pode montar o centros de Fe-S e utilizá-lo como cofator no processo de redox ou transferi-lo para outras proteínas que necessitam deste centro (**Figura 48**)²⁴⁸. Aparentemente SufC, do complexo SufBCD, facilita a transferência do centro de Fe-S de SufB para apoproteínas, via consumo de ATP²³⁴.

Figura 48: Modelo de biogênese de centros de Fe-S pelo sistema Suf^{248,234}.

SufS transfere grupos sulfídricos de L-cisteína para SufE. SufE, por sua vez, transfere o enxofre para o complexo ternário SufBCD, em que SufB pode montar o centro de Fe-S e utilizá-lo ou pode transferi-lo para apoenzimas com o auxílio de SufC, via consumo de ATP^{248,234}. As linhas pontilhadas são os possíveis caminhos para o íon de enxofre (S²⁻), a linha em negrito mostra os caminhos que o centro de Fe-S pode seguir²⁴⁸.

Devido a grande importância destes centros Fe-S nos seres vivos, a comunidade científica tem se desempenhado em entender cada vez mais os detalhes destas maquinarias de biogênese de centros de Fe-S e as enzimas que recebem estes centros. Muitas outras proteínas estão sendo descobertas, nos últimos três anos, por

participarem no processo de biogênese destes centros, algumas mencionamos brevemente a seguir. HscA/HscB são chaperoninas que interagem com IscU facilitando a transferência de centros [Fe-S] de IscU para apo-ferrodoxinas de forma dependente de ATP, mas não facilita a transferência de [4Fe-4S] de SufU para aconitase^{249,250}. CyaY doa íons de ferro para IscU de forma dependente de IscS, assim como frataxin, que estoca de seis a sete íons de ferro e os doa para o sistema *ISU*, homólogo ao sistema *ISC*, em células eucarióticas^{250,251}. Como *Xac* não possui o sistema *ISC*, também não possui homólogos dos demais componentes desta via de biogênese, como CyaY, HscA e HscB.

Por muito tempo, se propôs que os sistemas *ISC* e *SUF* fossem sistemas redundantes e possuíam a mesma função²⁵²⁻²⁵⁴. No entanto, um trabalho recentemente publicado por Vinella e colaboradores, analisando o efeito de diferentes células nocautes de *E. coli* aos genes pertencentes ao sistema *ISC* e *SUF*, sugere que há uma comunicação entre estes sistemas²⁵⁵. A **Figura 49** usa como exemplo a via de síntese de centro de Fe-S da proteína IspG/H de *E. coli*. Em condições de estresse oxidativo (**Fig. 49d**) e em ausência de ferro o sistema *SUF* é prioritário. O centro Fe-S é sintetizado por SufS/E/B/C/D e o transfere para SufA e posteriormente para IspG/H. Em condições anaeróbicas (**Fig. 49b**) o sistema *ISC* é prioritário. IscU transfere o centro Fe-S para IscA ou para ErpA e por sua vez o transfere para IspG/H. Em condições aeróbicas (**Fig. 49c**) o sistema *ISC* também é prioritário em que IscU transfere o centro de Fe-S para IscA e IscA o transfere para ErpA e esta o transfere para IspG/H. No entanto, todos os sistemas podem ser sintetizados simultaneamente em diferentes quantidades e o caminho que o centro Fe-S irá seguir depende das condições do meio e das regulações internas (**Fig. 49a**). Um resumo sobre os diferentes caminhos disponíveis para sintetizar centros de Fe-S pelo sistema *SUF* e *ISC* estão apresentados na **Figura 50** e na legenda da legenda.

Figura 49: Modelo do tráfego de centros de Fe-S que podem levar a maturação de IspG/H em *E. coli*.

O diagrama mostrado no item **a** mostra todas as possibilidades de caminho de Fe-S, em que os números representam os caminhos dos centros de Fe-S. IscU pode transferir o centro de Fe-S para IscA (1) ou para SufA (2) ou para ErpA (3). Os diagramas mostrados nos itens **b**, **c** e **d** representam os modelos para a maturação de IspG/H em diferentes condições de crescimento celular descrito abaixo de cada diagrama²⁵⁵.

Figura 50: Modelo de biossíntese de centros de Fe-S em bactérias, maquinaria ISC e SUF²³⁴.

Para a maturação das apoenzimas (Apo) as bactérias podem usar um sistema ou ambos em paralelo. O sistema *ISC*, ilustrado no lado esquerdo, consiste da proteína cisteína dessulfarase, IscS, que adquire íons de enxofre (esferas amarelas) a partir de resíduos de cisteína, convertendo-os a alanina. Estes enxofres formam ligações persulfeto (-SSH) com a cadeia lateral do resíduo de cisteína de IscS formando IscS-SSH. Posteriormente, proteínas montadoras de centros de Fe-S, *scaffold*, como IscA e IscU, formam centros de Fe-S transientes, que são formados pela transferência de íons de ferro (esferas vermelhas) por proteína ligadora de ferro CyaY (ou por outras proteínas ligadoras de ferro, todavia desconhecidas) e pelos íons de enxofre fornecidos por IscS-SSH que teve o enxofre reduzido pela ferredoxina Fdx. Este centro de Fe-S transiente é transferido para apoenzimas (Apo) que coordena estes centros através de resíduos específicos (cisteínas ou histidinas). A transferência do centro de Fe-S transiente de IscU para apoenzimas é mediado pelo sistema de chaperoninas HscA-HscB que ligam ao motivo conservado LPPVK. O sistema *ISC* funciona em condições normais de crescimento bacteriano enquanto que o sistema *SUF* (lado direito da figura) é ativado em condições de estresse oxidativo ou em condições de ausência de ferro em *E. coli*. A biossíntese inicia pela cisteína dessulfarase SufS, que atua da mesma forma que IscS, com exceção de que SufS transfere seu enxofre para SufE que forma a ligação persulfeto (-SSH), SufE-SSH. Este enxofre de SufE pode ser transferido para três proteínas montadoras de centros de Fe-S, SufA e SufU, que atuam da mesma forma que IscA e IscU, respectivamente, e SufB, que forma uma complexo estável com SufC-SufD. SufC é uma ATPase que poderia facilitar a dissociação do centro de Fe-S de SufB para apoproteínas. No caso do sistema *SUF*, não é conhecido as proteínas doadoras de íons de ferro e as proteínas envolvidas na redução do enxofre de SufE-SSH para SufE-SH²³⁴.

Sistema *SUF* em *Xac*: Em *Xac* os genes *sufRBCDS* formam um operon (XAC2934, XAC2935, XAC2936, XAC2937 e XAC2938 respectivamente) mas, *sufA* (XAC1619) assim como *sufE* (XAC2355) se encontram em outra região do genoma de *Xac* (**Figura 46**). *Xac* codifica mais dois homólogos de SufA, ErpA e NifA^{Like} codificadas pelos genes XAC2117 e XAC0499, respectivamente. Portanto, os genes pertencentes ao sistema *SUF* não formam um operon único como ocorre em *E. coli*.

Além do mais, em *E. coli* há três sistemas de dois componentes de cisteína dessulfarases (IscS, CsdA e SufS) e aceptores de enxofre (IscU, YgdK e SufE), enquanto que em *Xac* há apenas uma cisteína dessulfarase, SufS_{XAC2938} (anotada originalmente como NifS) e um aceptor de enxofre, SufE_{XAC2355}. Isto sugere que SufE_{XAC2355} e SufS_{XAC2938} possam ser a única via de geração de íons S⁻² destinados à formação de centros de Fe-S em *Xac*.

Genes próximos ao operon *SUF* codificam proteínas correlacionados ao metabolismo de íons de ferro. Interessantemente, ao lado do operon *SUF*, em *Xac*, há genes que codificam proteínas correlacionadas ao metabolismo de ferro e a enzimas que utilizam centros de Fe-S (**Fig. 51**). Dentre estas proteínas há FhuA, receptor dependente de TonB, está envolvida no transporte de Fe⁺³, na forma de ferricromo, e de antibióticos como albomicina e rifamicina²⁵⁶. Portanto, FhuA poderia estar envolvida na internalização de íons ferro que seriam posteriormente utilizados por SufB assim como por ErpA, IscA e NifA^{like} para a montagem de centros de Fe-S. Ao lado do gene *fhuA* há um gene que codifica a uma oxigenase dependente de ferro, *bedB* e mais adiante se encontra o gene *cysJ*. Tanto BedB, benzeno 1,2-dioxigenase ferredoxina, assim como CysJ, sulfito reductase, utilizam centros de Fe-S para sua catalise (**Figura 51**). Interessantemente também há neste locus o gene que codifica a ApbE, envolvida na biogênese de tiamina, que está envolvida de forma tangencial no metabolismo de centros de Fe-S^{257,258}. A significância da proximidade destes genes com o operon *SUF* em termos de correção de expressão gênica deve ser estudada através de novos ensaios biológicos.

5.1. 2 Objetivos

O objetivo inicial desta parte do projeto era determinar a estrutura tridimensional de SufE_{XAC2355} por cristalografia de raios-X, a fim de identificar sua função, que na época da sua escolha como alvo, era desconhecida¹. Durante este trabalho a função de SufE_{XAC2355} foi descrita por outros grupos^{88,243,244,247,248,259}, mas detalhes de seu mecanismo de ação se mantiveram desconhecidos.

Por este motivo novos objetivos foram adicionados como o estudo estrutural e funcional de outros componentes do sistema *SUF*, que inclui SufR_{XAC2934}, SufS_{XAC2938}, SufA_{XAC1619}, NifA_{XAC2117} e ErpA_{XAC0499}. Pretendíamos confirmar o aumento da atividade cisteína dessulfarease de SufS_{XAC2938} na presença de SufE_{XAC2355} e entender como ocorre a regulação gênica do sistema *SUF* por SufR_{XAC2934}.

5.1. 3 Resultados e Discussão

5.1.3. 1 Cristalização de SufE

A clonagem do gene de *sufE*_{XAC2355}, sua purificação e cristalização foram feitas em colaboração com a Dra. Leonor Galvão de Botton e a Dra. Lucicleide Ribeiro¹. A clonagem, expressão, purificação e cristalização de SufE_{XAC2355} (**Figura 1 do Anexo 5**) assim como as condições de cristalização estão descritas nos Materiais e Métodos do **Anexo 5**.

O conjunto de dados de difração de raios-X dos cristais de SufE_{XAC2355} foram coletados na linha MX1 do laboratório Nacional de Luz Sincrotron em Campinas. Os detalhes do processamento dos dados de difração estão descritos no **Anexo 5**, a seguir.

Anexo 5 – Expressão, cristalização e análise cristalográfica preliminar de SufE (XAC2355) de *Xanthomonas axonopodis* pv *citri*.

– Expression, crystallization and preliminary crystallographic analysis of SufE (XAC2355) from *Xanthomonas axonopodis* pv *citri*

crystallization communications

Acta Crystallographica Section F
Structural Biology
and Crystallization
Communications

ISSN 1744-3091

Cristiane R. Guzzo,^a
Lucicleide R. Silva,^a Leonor M. P.
Galvão-Botton,^a João A. R. G.
Barbosa^b and Chuck S. Farah^{a*}

^aDepartamento de Bioquímica, Instituto de
Química, Universidade de São Paulo,
CEP 05508-000, São Paulo, SP, Brazil, and

^bCentro de Biologia Molecular Estrutural, LNLS,
CEP 13084-971, Campinas, SP, Brazil

Correspondence e-mail: chsfarah@iq.usp.br

Received 3 January 2006
Accepted 9 February 2006

Expression, crystallization and preliminary crystallographic analysis of SufE (XAC2355) from *Xanthomonas axonopodis* pv. *citri*

Xanthomonas axonopodis pv. *citri* (*Xac*) SufE (XAC2355) is a member of a family of bacterial proteins that are conserved in several pathogens and phytopathogens. The *Escherichia coli* *suf* operon is involved in iron-sulfur cluster biosynthesis under iron-limitation and stress conditions. It has recently been demonstrated that SufE and SufS form a novel two-component cysteine desulfurase in which SufS catalyses the conversion of L-cysteine to L-alanine, forming a protein-bound persulfide intermediate. The S atom is then transferred to SufE, from which it is subsequently transferred to target molecules or reduced to sulfide in solution. Here, the cloning, expression, crystallization and phase determination of *Xac* SufE crystals are described. Recombinant SufE was crystallized in space group $P2_12_12_1$ and diffracted to 1.9 Å resolution at a synchrotron source. The unit-cell parameters are $a = 45.837$, $b = 58.507$, $c = 98.951$ Å, $\alpha = \beta = \gamma = 90^\circ$. The calculated Matthews coefficient indicated the presence of two molecules in the asymmetric unit. Phasing was performed by molecular-replacement using *E. coli* SufE as a model (PDB code 1mzg) and an interpretable map was obtained.

1. Introduction

Xanthomonas axonopodis pv. *citri* (*Xac*) SufE (XAC2355) is a member of a family of proteins that are conserved in many prokaryotic pathogens and phytopathogens, including *Salmonella* spp., *Escherichia coli*, *Pseudomonas* spp., *Shigella* spp., *Yersinia* spp., *Leptospira interrogans*, *Erwinia carotovora*, *Vibrio* spp., *Azotobacter vinelandii*, *Shewanella baltica*, *Xanthomonas* spp. and *Xylella* spp. DNA sequences coding for SufE homologs have also been found in several eukaryotic genomes: *Plasmodium*, *Anopheles gambiae* and *Arabidopsis thaliana*. This protein was originally selected for structural analysis owing to its well folded and non-aggregated state in solution at a time when its function was unknown (Galvão Botton *et al.*, 2003).

In *E. coli*, the *sufABCDSE* cluster is expressed under conditions of iron limitation and oxidative stress, one of its functions being the restoration of damaged Fe-S clusters (Nachin *et al.*, 2001, 2003). In *E. coli*, SufA is a scaffold protein for the assembly of Fe-S clusters. SufB and SufD form a complex with SufC, a soluble cytoplasmic ABC-ATPase. SufBCD stimulates the cysteine desulfurase activity of the SufS-SufE complex (Outten *et al.*, 2003). In *Xac*, the cluster of genes *XAC2935-XAC2936-XAC2937-XAC2938* codes for homologues of SufB, SufC, SufD and SufS, respectively, although they were not annotated as such in the *Xac* genome (da Silva *et al.*, 2002). Genes coding for homologues of SufA (*XAC1619*) and SufE (*XAC2355*) are found elsewhere in the *Xac* genome.

Cysteine desulfurases are important pyridoxal 5'-phosphate (PLP) dependent enzymes that are involved in sulfur mobilization in many living organisms (Marquet, 2001; Mihara *et al.*, 2002; Fontecave *et al.*, 2003; Zheng *et al.*, 1994). The chemical steps involved in the cysteine desulfurase activity of the SufS-SufE complex consist of two stages. The first stage consists of the PLP-dependent transfer of sulfur from L-cysteine to a Cys residue on SufS, forming a persulfide linkage. The sulfur is then passed from SufS to a Cys residue on SufE, where it also forms a persulfide linkage (Loiseau *et al.*, 2003; Ollagnier-de-Choudens *et al.*, 2003). The cysteine desulfurase activity of SufS on its

© 2006 International Union of Crystallography
All rights reserved

own is approximately 30-fold to 50-fold less than that of the SufS–SufE complex (Loiseau *et al.*, 2003).

Recently, the structures of five SufE homologues have been determined: the NMR structures of *E. coli* YjdK (PDB code 1ni7; Liu *et al.*, 2005), *Mus musculus* IscU (PDB code 1wfz; unpublished data) and *Haemophilus influenzae* IscU (PDB codes 1r9p and 1q48; Ramelot *et al.*, 2004) and the crystal structures of *E. coli* SufE (PDB code 1mzg; Goldsmith-Fischman *et al.*, 2004) and *Bacillus subtilis* IscU (PDB code 1xjs; unpublished data). In this report, we describe the cloning, expression, purification and crystallization of recombinant *Xac* SufE. An initial electron-density map of the *Xac* SufE crystal structure has been obtained by molecular replacement using the *E. coli* SufE structure as a model (PDB code 1mzg; Goldsmith-Fischman *et al.*, 2004).

2. Cloning and expression of XAC2355

The gene coding for SufE (XAC2355) was amplified by PCR from *Xac* genomic DNA using the following primers designed based on the published *Xac* genome sequence (da Silva *et al.*, 2002): forward, 5'-CATGCCATGGCTCATATGACACCTCCCCCTTCC-3'; reverse, 5'-GGAATTCAAGCTTTCACCTGCTGCGCGCG-3'. The PCR product was digested with *Hind*III and *Nde*I and subcloned into the pET-3a vector (Studier *et al.*, 1990) previously digested with the same endonucleases. SufE was expressed in *E. coli* strain BL21(DE3)pLysS (Studier *et al.*, 1990). The protein was produced by growing a culture in 2×TY medium to an optical density (600 nm) of 0.8, at which point heterologous protein expression was induced by the addition of 1 mM isopropyl β -D-thiogalactopyranoside and the cells were harvested after growing for 4 h and stored at 203 K.

3. Protein purification

Cells from 1 l culture were resuspended in 25 ml 50 mM Tris–HCl pH 8.0, 25% sucrose, 1 mM EDTA, 1 mM PMSF and lysed using a French press. The soluble fraction was applied onto a Q-Sepharose Fast Flow (FF) Hi-Load 16/10 column (Amersham Pharmacia) previously equilibrated with 50 mM Tris–HCl pH 7.0, 1 mM EDTA and 14 mM β -mercaptoethanol. Bound proteins were eluted using a 12 column volume 0–300 mM NaCl gradient. Fractions containing SufE were concentrated using an Amicon system with a 10 kDa molecular-weight cutoff membrane and then purified further by gel filtration on a Superdex 75 prep-grade column (Amersham Pharmacia) equi-

Figure 1
Crystals of SufE. The largest plates are approximately 400 μ m in the longest dimension.

Table 1

Crystal parameters and data-reduction statistics.

The data sets were measured from a single crystal. Values in parentheses refer to the highest resolution shell.

Space group	$P2_12_12_1$
Unit-cell parameters	
a (Å)	45.837
b (Å)	58.507
c (Å)	98.951
No. of images	212
Resolution range (Å)	50.00–1.90 (1.97–1.90)
Wavelength (Å)	1.438
No. of observed reflections	128189
No. of unique reflections	20431 (2002)
$\langle I/\sigma(I) \rangle$	25.6 (4.4)
Multiplicity	6.3 (5.8)
Completeness (%)	93.5 (93.3)
R_{int} (%)	5.8 (37.3)

$$\dagger R = \sum |I(h) - \langle I(h) \rangle| / \sum I(h).$$

brated with 5 mM Tris–HCl pH 7.0. The purity of the protein was judged by visualization of Coomassie-stained SDS–PAGE.

4. Crystallization

Xac SufE crystals (Fig. 1) were grown using the hanging-drop vapour-diffusion technique at 291 K. Initial crystallization conditions that produced needle-like crystals were obtained by the sparse-matrix sampling approach using the Index crystallization screen kit from Hampton Research. Optimization was then pursued by varying the precipitant concentration and the buffer pH. Suitable crystals for diffraction experiments were obtained by mixing 1 μ l of a 9.8 mg ml^{−1} protein solution (5 mM Tris–HCl pH 7.0) with 1 μ l reservoir solution consisting of 17–21% (w/v) PEG 5000 and 0.1 M Tris–HCl pH 7.0, followed by equilibration against 0.4 M reservoir solution. Rectangular crystals appeared within a few days and grew to mature size within a few weeks.

5. Data collection and preliminary structure analysis

X-ray diffraction data were collected at the D03B beamline of the Laboratório Nacional de Luz Síncrotron, Campinas, Brazil using a MAR CCD detector. Crystals equilibrated against 19% PEG 5000 and 0.1 M Tris–HCl pH 7.0 were transferred to solution containing

Figure 2
A section of the current SufE crystal $2F_o - F_c$ electron-density map contoured at 1.0σ .

crystallization communications

26% PEG 400 for cryoprotection. The crystal was then flash-frozen in liquid nitrogen and maintained at 100 K in a nitrogen-gas stream during data acquisition. A single data set was collected at 1.438 Å using 1 min exposure and 1° oscillation from 1 to 212°. The diffraction patterns from this data set extended to approximately 1.9 Å resolution (Table 1). The crystal belongs to space group $P2_12_12_1$, with unit-cell parameters $a = 45.837$, $b = 58.507$, $c = 98.951$ Å, $\alpha = \beta = \gamma = 90^\circ$. There are two protein molecules per asymmetric unit, as suggested by the Matthews coefficient ($V_M = 2.7$ Å³ Da⁻¹; solvent content 40.4%).

Diffraction data were indexed, integrated, scaled and merged using the *HKL2000* package (Otwinowski & Minor, 1997). Details of data acquisition and data-processing statistics are shown in Table 1. Initial phases of the *Xac* SufE crystal were calculated by molecular replacement using *Phaser* (Storoni *et al.*, 2004; Read, 2001) from the *CCP4* suite (Collaborative Computational Project, Number 4, 1994) with the *E. coli* SufE structure 1mzg (Goldsmith-Fischman *et al.*, 2004) as the search model. The *E. coli* and *Xanthomonas* proteins share 34% sequence identity. The phases were improved by density-modification protocols using the programs *SOLOMON* (Abrahams, 1997) and *DM* (Collaborative Computational Project, Number 4, 1994; Cowtan, 1994). Interpretation of electron-density maps (Fig. 2) and model refinement are currently under way using the programs *COOT* (Emsley & Cowtan, 2004), *REFMAC* (Murshudov *et al.*, 1997) and *CNS* (Brünger *et al.*, 1998).

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (Brazil), Associação Brasileira de Tecnologia de Luz Síncrotron (ABTLuS) and FCT/MCT (Portugal).

References

- Abrahams, J. P. (1997). *Acta Cryst.* **D53**, 371–376.
- Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* **D54**, 905–921.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Cowtan, K. (1994). *Int. CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **31**, 34–38.
- da Silva, A. C. *et al.* (2002). *Nature (London)*, **417**, 459–463.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Fontecave, M., Ollagnier-de-Choudens, S. & Mulliez, E. (2003). *Chem. Rev.* **103**, 2149–2166.
- Galvão Botton, L. M. P., Katsuyama, A. M., Guzzo, C. R., Almeida, F. C., Farah, C. S. & Valente, A. P. (2003). *FEBS Lett.* **552**, 207–213.
- Goldsmith-Fischman, S., Kuzin, A., Edstrom, W. C., Benach, J., Shastry, R., Xiao, R., Acton, T. B., Honig, B., Montelione, G. T. & Hunt, J. F. (2004). *J. Mol. Biol.* **344**, 549–565.
- Liu, G., Li, Z., Chiang, Y., Acton, T., Montelione, G. T., Murray, D. & Szyperki, T. (2005). *Protein Sci.* **14**, 1597–1608.
- Loiseau, L., Ollagnier-de-Choudens, S., Nachin, L., Fontecave, M. & Barras, F. (2003). *J. Biol. Chem.* **278**, 38352–38359.
- Marquet, A. (2001). *Curr. Opin. Chem. Biol.* **5**, 541–549.
- Mihara, H., Fujii, T., Kato, S., Kurihara, T., Hata, Y. & Esaki, N. (2002). *J. Biochem. (Tokyo)*, **131**, 679–685.
- Murshudov, G. N., Vagin, A. A. & Dodson, E. I. (1997). *Acta Cryst.* **D53**, 240–255.
- Nachin, L., El Hassouni, M., Loiseau, L., Expert, D. & Barras, F. (2001). *Mol. Microbiol.* **39**, 960–972.
- Nachin, L., Loiseau, L., Expert, D. & Barras, F. (2003). *EMBO J.* **22**, 427–437.
- Ollagnier-de-Choudens, S., Lascoux, D., Loiseau, L., Barras, F., Forest, E. & Fontecave, M. (2003). *FEBS Lett.* **555**, 263–267.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Ottens, F. W., Wood, M. I., Munoz, F. M. & Storz, G. (2003). *J. Biol. Chem.* **278**, 45713–45719.
- Ramelot, T. A., Cort, J. R., Goldsmith-Fischman, S., Kornhaber, G. J., Xiao, R., Shastry, R., Acton, T. B., Honig, B., Montelione, G. T. & Kennedy, M. A. (2004). *J. Mol. Biol.* **344**, 567–583.
- Read, R. J. (2001). *Acta Cryst.* **D57**, 1373–1382.
- Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). *Acta Cryst.* **D60**, 432–438.
- Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). *Methods Enzymol.* **185**, 60–89.
- Zheng, L., White, R. H., Cash, V. L. & Dean, D. R. (1994). *Biochemistry*, **33**, 4714–4720.

5.1.3. 2 Determinação da Estrutura de SufE

Determinação da estrutura de SufE: As fases iniciais para as reflexões do conjunto de dados do cristal de SufE_{XAC2355} foram determinadas usando a técnica de substituição molecular pelo programa *PHASER*⁶⁸ do pacote do *CCP4*²⁶⁰. Utilizamos como modelo SufE de *E. coli* (PDBID:1MZG)²⁴³. Foi obtido um z-score de 9.67 e ll-gain de 73.64. O refinamento do modelo da estrutura de SufE_{XAC2355} está inacabada, sendo os valores atuais de R_{free} e R_{factor} , 0,29 e 0,20, respectivamente (**Tabela 21**).

Há duas moléculas de SufE_{XAC2355} na unidade assimétrica (**Figura 52 a**) e a superfície enterrada entre as duas moléculas é de 722 Å², sugerindo que este dímero não é biologicamente relevante¹³⁹. A estrutura de SufE_{XAC2355} é formada por 3 fitas e 7 alfa-hélices com topologia, α - α - α - β - β - α - α - α , que forma um domínio compacto (**Figura 52 b**).

Tabela 21: Dados do refinamento da estrutura de SufE.

Estes são os dados estatísticos da última estrutura de SufE_{XAC2355} refinada. O valor de R_{free} é de 0,29 e sua diferença com o valor R_{factor} é de 0.09, mostram que o refinamento ainda não está terminado.

Resolution (Å)		25.2 -1.9
R_{free}		0.29
R_{factor}		0.20
FOM		0.80
RMS	Angle (°)	1.613
	Bond (Å)	0.014

R_{free} were calculated with 5.1 % of the data

a *b*

Figura 52: Estrutura de SufE¹⁶⁴.

O dímero de SufE_{XAC2355} na unidade assimétrica (**a**) e a estrutura do monômero de SufE_{XAC2355}, destacando os números das alfa hélices (α) e as fitas betas (β) (**b**). As estruturas de SufE_{XAC2355}, formato de cartoon, foram feitas utilizando a interface gráfica do programa PYMOL²⁶¹.

Motivos conservados na família de SufE. A família de SufE é altamente conservada, sendo encontrada em organismos eucarióticos e procarióticos. Até o momento foram anotados 844 ortólogos de SufE encontrados em 419 espécies¹²⁰. O domínio de SufE geralmente é encontrado sozinho, representando 815 órtologos, ou pode ser encontrado fusionado a outros domínios, por exemplo: i. BolA, envolvida na divisão celular e também é induzida em condições de estresse²⁶², ii. NadA, quinolato sintetase, que catalisa a condensação do dihidroxiacetona fosfato e iminoaspartato para formar o ácido quinólico, da via de síntese de piridina²⁶³, iii. Aminotran_5, aminotransferase encontrada em cisteína dessulfarase¹²⁰, e iv. tRNA_ME_trans, é uma tRNA(5-metilaminometil-2-tioridina)-metiltransferase que está envolvida na biosíntese do 5-metilaminometil-2-tioridina encontrada em posições de “wobble” do tRNA²⁶⁴.

A família de SufE possui quatro motivos de sequência altamente conservados (**Figura 53**): a) G₅₉C₆₀XS₆₂, b) S₈₁D₈₂, c) G₈₈ e d) R₁₂₆XXG₁₂₉, em que os números correspondem aos aminoácidos da proteína de SufE de *Xac* e a C₆₀ é o sítio de formação do grupo persulfeto. Os aminoácidos destes motivos estão todos localizados em uma única região na superfície da estrutura de SufE_{XAC2355} (**Figura 54 a**) e formam uma rede de interação entre eles (**Figura 54 b**), descritos a seguir. Além dos aminoácidos conservados mencionados acima, os resíduos hidrofóbicos F₂₃, L₄₃, V₆₄, W₆₅, F₇₇, L₁₁₈ e L₁₁₂ são também conservados e devem contribuir para manter a estabilidade do cerne hidrofóbico da proteína (**Figura 54 a**). A S₈₁ e D₈₂ fazem pontes de hidrogênio com a S₆₂ que por sua vez interage com a C₆₀ e a R₁₂₆. A G₅₉ provavelmente é importante para a estabilização da conformação da C₆₀ no core hidrofóbico, envolvida na formação da ligação tiosulfeto com o enxofre proveniente de SufS^{243,244,265}. A cadeia lateral da C₆₀ está localizada em um sulco profundo e parcialmente hidrofóbico na estrutura de SufE_{XAC2355} (**Figura 54 a**). Neste sulco, os átomos mais próximos ao S_γ da C₆₀ (menos de 5Å) são NH₂ da R₁₂₆, C_α da G₁₂₉, C_γ da V₈₆, C_γ da V₆₄ e o O_γ da S₆₂. Aparentemente, os aminoácidos dos motivos de sequência conservados estão envolvidos na estabilização da C₆₀ neste core hidrofóbico, protegendo-a de ser oxidada por componentes em solução (**Figura 55**). A C₆₀ faz pontes de hidrogênio com três moléculas de água, w₁₁₀, w₅₉ e w₆₂ encontrados no sulco (**Figura 55 a**) e w₁₁₀ pode estar ocupando o espaço da ligação tiosulfeto que SufE irá formar quando receber o enxofre de SufS (**Figura 55 b**).

Figura 53: Alinhamento múltiplo de ortólogos da família de SufE..

As caixas em preto mostram os quatro motivos de aminoácidos mais conservados na família de SufE e os asteriscos em rosa destacam outros resíduos conservados. Foram utilizados 844 seqüências de proteínas ortólogos a SufE. Estas seqüências foram obtidas no site do Pfam¹²⁰ e posteriormente utilizadas para o desenho gráfico pelo programa WebLogo¹⁶⁵. A seqüência de aminoácidos abaixo de cada linha corresponde a seqüência de aminoácidos de SufE de Xac iniciando no aminoácido 20 e terminando no aminoácido 142. A representação é proveniente do programa WebLogo^{164, 165, 166}.

Figura 54: Os aminoácidos conservados na família de SufE se localizam em uma face da estrutura e formam uma rede de interações entre eles.

Os aminoácidos conservados na família de SufE (Figura 53) estão todos localizados em uma face da estrutura de SufE_{XAC2355} (a), e fazem uma rede de interação entre si (b). As moléculas de água que fazem ponte de hidrogênio com a C₆₀ estão mostradas em forma de esferas coloridas em amarelo (w₁₁₀), laranja (w₅₉) e cyan (w₆₂). A cadeia lateral da cisteína 60 (amarelo) está localizada em uma cavidade parcialmente hidrofóbica na estrutura de SufE_{XAC2355}. Os resíduos pertencentes aos motivos conservados na família de SufE estão coloridos em azul, e os demais resíduos conservados na família de SufE estão coloridos em rosa e a C₆₀ em amarelo, os demais aminoácidos estão coloridos em vermelho. A representação das estruturas foi produzida utilizando o programa PYMOL²⁶¹.

Figura 55: Os aminoácidos conservados na família de SufE se localizam em uma face da estrutura e formam uma rede de interações entre eles.

A C_{60} está enterrada em um sulco parcialmente hidrofóbico e faz ponte de hidrogênio com três moléculas de água, w59, w62 e w110 (menos de 4,5Å) (**a**) sendo que uma delas, neste caso w110, pode estar ocupando o espaço da ligação tiosulfeto (**b**). Os mapa de densidade eletrônicas foram desenhados utilizando a diferença entre $2F_o - F_c$ e sigma de 1.0, programa COOT⁷⁴.

5.1.3. 3 *Proposta de possíveis mecanismos de ação de SufE*

A observação que a C₆₀ se encontra dentro de um sulco parcialmente hidrofóbico e que faz contatos com vários resíduos conservados na família de SufE põe restrições nos possíveis mecanismos para o processo de oxidação e redução de SufE. Em termos gerais, podemos imaginar dois possíveis mecanismos de transferência do grupo tiol de SufS para SufE. O primeiro mecanismo envolve poucas mudanças conformacionais no sítio ativo de SufE, descrito nas **Figuras 54 e 55**, e a ativação da C₆₀ que irá atacar nucleofilicamente a ligação tiosulfeto, S-SH, de SufS. Uma segunda possibilidade (mecanismo 2) é o rearranjo do loop que contém a C₆₀, durante o processo de transferência do grupo tiol de SufS para SufE.

Mecanismo 1. A análise do sítio ativo de SufE_{XAC2355} de *Xac*, **Figura 54-55**, mostrou uma rede de interações de aminoácidos conservados na família de SufE que envolve, D₈₂, S₈₁, S₆₂, R₁₂₆ e C₆₀. Com base nesta rede de interações podemos propor um mecanismo de ativação da C₆₀ no sítio ativo e posterior oxidação desta por SufS-SH (**Figura 56**). O D₈₂ e S₈₁ fazem ponte de hidrogênio com a S₆₂ de tal forma que o grupo hidroxil da S₆₂ deve estar parcialmente desprotonada. Esta carga parcial é estabilizada pela R₁₂₆ e pela ponte de hidrogênio com o grupo tiol da cadeia lateral da C₆₀, que também deve estar parcialmente desprotonada e sendo estabilizada pela R₁₂₆ e pelas águas W₅₉, W₆₂, W₁₁₀ (W₁₁₀ na **Figura 56**).

A interação de SufS-SH (SufS oxidada) poderia ocasionar mudanças conformacionais sutis no sítio ativo de SufE_{XAC2355}, trazendo o D₈₂ e a S₈₁ mais próximo da S₆₂. Esta mudança conformacional desencadearia uma cascata de protonação e desprotonação (**Figura 56 b**), em que o D₈₂ desprotona a S₆₂ com a ajuda da S₈₁. A S₆₂ desprotonada retira o próton do grupo tiol da C₆₀ (**Figura 56 c**). O enxofre reduzido da C₆₀ é estabilizado pela R₁₂₆ e pela esfera de hidratação envolvendo as moléculas de água, W₅₉, W₆₂, W₁₁₀. Desta forma a C₆₀ reduzida está ativada para reduzir a ponte tiosulfeto (-S-SH) de SufS-SH.

Podemos imaginar duas possibilidades para o ataque nucleofílico do íon tiolato da C₆₀ na ligação tiosulfeto de SufS-SH. Uma delas é a partir da mudança conformacional da cadeia lateral da C₆₀, que é facilitado pela presença da G₅₉, que gira para poder ter acesso a ligação tiosulfeto de SufS-SH, exatamente no sítio ativo de SufS.

Outra possibilidade seria que a ligação tiosulfeto de SufS-SH entraria no sítio ativo de SufE e neste caso não seria necessário mudanças conformacionais no sítio ativo de SufE. Este processo faria com que uma ou mais moléculas de água na esfera de hidratação da C₆₀ fosse expulsa do sítio ativo. Este buraco seria ocupado pelo grupo –SH transferido para SufE_{XAC2355} (**Figura 56 d**). Desta forma, todos os passos da reação de transferência do grupo tiosulfeto acontecem dentro deste core hidrofóbico protegido de ser reduzido de forma inespecífica pelo solvente.

Este mecanismo de ativação da C₆₀ é condizente com o mecanismo mais aceito pela comunidade científica de formação de um íon tiolato que irá atacar nucleofilicamente a ligação tiosulfeto de SufS-SH, que por sua vez é estabilizado pela carga positiva da R₁₂₆ ou por algum resíduo positivo de SufS^{88,243,244}. Assim como mudanças conformacionais próximas a C₆₀, que permitirá a transferência do grupo persulfídico de SufS para SufE^{88,243,244}.

Figura 56: Possível mecanismo de ação de SufE (Mecanismo 1).

O sítio ativo de SufE_{XAC2355} é composto por uma rede de interações entre aminoácidos conservados na família de SufE (a), em que a interação de SufS-SH com SufE ocasiona a aproximação da D₈₂ e da S₈₁ com a S₆₂ desencadeando uma cascata de protonação e desprotonação (b). D₈₂ com a ajuda da S₈₁ desprotona a S₆₂ que por sua vez desprotona a C₆₀. O próton da S₆₂ está colorido em verde e da C₆₀ está colorido em vermelho. Com isso há um relay de prótons da C₆₀ para S₆₂ para D₈₂. (c). A R₁₂₆ assim como a camada de solvatação composta pelas águas w₅₉, w₆₂, w₁₁₀, (está ilustrado na figura apenas a água w₁₁₀) estabilizam o grupo tiol da C₆₀ desprotonada. A C₆₀ desprotonada ataca nucleofílicamente a ponte tiosulfeto de SufS-SH e volta para o sítio ativo de SufE oxidada, expulsando pelo uma das moléculas de água da camada de solvatação (w₁₁₀ na figura) (d). Este diagrama foi produzida utilizando o programa ACD/ChemSketch²⁶⁶.

Mecanismo 2. Uma outra possibilidade é que o loop $\alpha 3\text{-}\beta 1$, que contém a C₆₀, sofra mudanças conformacionais de tal forma que a cadeia lateral da C₆₀ fique mais exposta ao solvente e acessível para reagir com SufS–SSH. Esta hipótese é baseada na observação de que a C₆₀ faz ponte de hidrogênio com 3 moléculas de água em um sulco parcialmente hidrofóbico (**Figura 55 a**). Isto poderia indicar que a C₆₀ poderia sofrer mudanças conformacionais de tal forma que ora estaria localizada neste sulco e ora estaria exposta ao solvente. A quebra das pontes de hidrogênio com as águas no sulco parcialmente hidrofóbico seriam refeitas com moléculas de água do solvente.

5.1.3. 4   *Comparação da Estrutura de SufE com outras estruturas depositadas no PDB.*

SufE possui similaridade estrutural com IscU/NifU. A comparação da estrutura de SufE_{XAC2355} com outras estruturas depositadas no PDB usando o programa Dali server²⁶⁷ identificou três classes de proteínas: proteínas classificadas como SufE ou YgdK, proteínas classificadas como IscU ou NifU, IscU/NifU e LplA (Lipoato ligase A) envolvida na síntese de lipoil-AMP a partir de lipoato e ATP²⁶⁸ (**Tabela 22**). As duas últimas famílias de proteínas possuem baixa similaridade de sequência primária com SufE de Xac (**Tabela 22**).

Tabela 22: Comparação da estrutura de SufE com outras estruturas depositadas no PDB.

As estruturas de proteínas que possuem similaridade estrutural com a estrutura de SufE de Xac²⁶⁷. Membros da família de SufE estão coloridas em verde escuro, membros da família IscU/NifU em verde claro e membros da família de Lipoato ligase A em azul.

PDBID (cadeia)	Z-score	% identidade com SufE _{XAC2355}	Nome da Proteína	Organismo	Referência
3g0m (A)	19.2	32	SufE	<i>S. Typhimurium</i>	265
1mzg (A)	18.5	32	SufE	<i>E. coli</i>	243
1ni7 (A)	16.1	27	Proteína Hipotética YGDK	<i>E. coli</i>	259
1wlo (A)	13.5	20	SufE	<i>T. thermophilus</i>	Structural Genomics 269
2qq4 (H)	7.3	13	IscU	<i>T. thermophilus</i>	269
1su0 (B)	7.1	8	IscU	<i>S. pyogenes</i>	Structural Genomics 270
1xjs (A)	6.8	8	NifU like protein	<i>B. subtilis</i>	271
2z7e (B)	6.0	10	NifU like protein	<i>A. aeolicus</i>	271
2azh (A)	5.7	9	IscU	<i>B. subtilis</i>	272
1wfz (A)	5.1	8	Cluter de fixação de Nitrogênio	<i>Mus musculus</i>	273
1x2g (B)	4.2	8	Lipoate-protein ligase A	<i>E. coli</i>	268
1x2h (B)	4.2	8	Lipoate-protein ligase A	<i>E. coli</i>	268

Sobreposição estrutural de SufE com outras proteínas pertencentes a família de SufE. Existem quatro estruturas no PDB de proteínas com similaridade de sequência entre 20-30% com SufE_{XAC2355}. A sobreposição da estrutura de SufE_{XAC2355} com estas proteínas mostrou alta similaridade estrutural (**Figura 58 a**). Todas possuem a mesma topologia α - α - α - β - β - α - α - α , com exceção das estruturas de PDBID 1mzg, 1ni7 e 1wlo, em que estas proteínas possuem uma mini fita beta, β^* , (composta por apenas dois resíduos) entre a α 3 e a β 1, formando a topologia α - α - α - β^* - β - β - α - α - α (**Figura 58 a e e**). No entanto, estas pequenas mudanças estruturais não afetaram a posição dos resíduos dos motivos de sequência conservados, G^C-S, SD, G e RxxG da família de SufE (**Figura 57, 58 e e 59 a**), que se sobrepõe muito bem entre eles (**Figura 59 a**). Isso sugere, que SufE interaja com SufS da mesma forma como IscU interage com IscS, como já tinha sido sugerido por Goldsmith-Fischman²⁴³.

No entanto, estes resíduos de sequência conservados não são encontrados nas estruturas da família IscU/NifU nem tão pouco na LplA (**Figura 57**). A comparação da estrutura de SufE_{XAC2355} com LplA mostrou que apenas as duas fitas e as quatro últimas hélices da estrutura de SufE_{XAC2355} se sobrepõe com a estrutura de LplA (**Figura 58 d e e**), sendo que as três ultimas hélices sofrem um deslocamento

estrutural em relação a estrutura de SufE_{XAC2355}. No entanto, não há nenhuma correlação do sítio ativo de LplA com a estrutura de SufE_{XAC2355}.

Sobreposição estrutural de SufE com membros da família NifU/IscU. A comparação da estrutura de SufE_{XAC2355} com as estruturas das proteínas pertencentes a família de NifU/IscU mostrou que as estruturas pertencentes a família IscU/NifU possuem a mini fita β^* maior. Nem todas possuem as três hélices iniciais (**Figura 58 e**) e as três últimas hélices sofrem um deslocamento estrutural quando comparadas com a estrutura de SufE_{XAC2355} (**Figura 58 b, c e e**). Apesar da grande similaridade estrutural entre estas proteínas, estas não possuem os aminoácidos dos motivos conservados na família de SufE.

Figura 57: Alinhamento das seqüências das proteínas com similaridade estrutural com SufE de Xac²⁷⁴.

As caixas em vermelho destacam os motivos de seqüência conservados na família de SufE, e a linha pontilhada em vermelho mostra que apenas as estruturas com PDBID acima da linha contem estes motivos de seqüência conservados.

Família	Figura	PDBID	Topology	Ligante	Aminoácidos envolvidos na ligação ou motivos conservados
SufE	A , C e D	SufE	α - α - α - β - β - α - α - α	-	G ₅₉ C ₆₀ -S ₆₂ S ₈₁ D ₈₂ R ₁₂₆ XXG ₁₂₉
	A	3g0m (A)	α - α - α - β - β - α - α - α	-	G ₅₀ C ₅₁ -S ₅₃ S ₇₃ D ₇₄ R ₁₁₉ XXG ₁₂₁
	A	1mzg (A)	α - α - α - β^* - β - α - α - α	-	G ₅₀ C ₅₁ -S ₅₃ S ₇₃ D ₇₄ R ₁₁₉ XXG ₁₂₁
	A	1ni7 (A)	α - α - α - β^* - β - α - α - α	-	G ₆₀ C ₆₁ -X S ₈₃ E ₈₄ R ₁₂₉ XXQ
	A	1wlo (A)	α - α - β^* - β - α - α - α	-	XC ₄₇ -X X D ₇₀ R ₁₁₅ XXG ₁₁₈
IscU/NifU	B	2qq4 (H)	- α - β - β - β - α - α - α	Zn	C ₃₉ GD ₄₁ C ₆₄ X ₂ S ₆₇ C ₁₂₅
	B	1su0 (B)	- α - β - β - β - α - α - α	Zn	C ₄₀ GD ₄₂ C ₆₅ X ₂ S ₆₈ C ₁₂₇
	B	1xjs (A)	- α - β - β - β - α - α - α	Zn	C ₄₁ GD ₄₃ C ₆₆ X ₂ S ₆₉ C ₁₂₈
	B	2azh (A)	- α - β - β - β - α - α - α	Zn	C ₄₁ GD ₄₃ C ₆₆ X ₂ S ₆₉ C ₁₂₈
	C	2z7e (B)	- α - β - β - β - α - α - α	[Fe ₂ -S ₂]	C ₃₆ C ₆₃ XS ₆₅ H ₁₀₆ C ₁₀₇
	C	1wfz (A)	- - - β - β - β - α - α - α	Zn	C ₃₃ GD ₃₅ C ₅₉ XS ₆₁ H ₁₀₁ C ₁₀₂
Ligase	D	1x2g (B)	outroDomínio- β - β - β - α - α - α	-	-----
	D	1x2h (B)	outroDomínio- β - β - β - α - α - α	-	-----

Figura 58: Sobreposição da estrutura de SufE com outras estruturas depositadas no PDB com similaridade estrutural²⁷⁴.

Sobreposição da estrutura de SufE de *Xac* (vermelho) com SufE de PDBID 3g0m (laranja), 1mzg (verde claro), 1ni7 (cyan) e 1wlo (rosa claro) (**a**) mostrando alta similaridade estrutural entre eles. Sobreposição estrutural de proteínas pertencentes à família IscU/NifU com PDBID 2qq4 (preto), 1su0 (violeta), 1xjs (laranja) e 2azh (verde escuro) (**b**), mostrando alta similaridade estrutural entre eles. As proteínas com PDBID 2z7e (pink) e 1wfz (amarelo) possui divergências estruturais com as demais estruturas mostradas no item **b**, principalmente nas últimas três hélices e estão sobrepostas com a estrutura de SufE de *Xac* (vermelho) (**c**). Sobreposição da estrutura de SufE de *Xac* (vermelho) com as LplA, PDBID 1x2g em cinza e 1z2h em azul, mostrando que apenas a parte C-terminal de SufE se sobrepõe com estas proteínas (**d**). O item **e** é uma tabela que resume os itens **a-d**, em que classifica as proteínas em três grupos: SufE, IscU/SufU e Ligase. O PDBID está colorido da mesma forma que a estrutura está colorida e mostra em que item este PDB está sendo mostrado, descreve a topologia de cada estrutura em que os elementos de estrutura secundária estão colocados de forma alinhada com a topologia de SufE de *Xac*, em que o β^1 representa uma fita com apenas dois resíduos, e as hélices coloridas em marrom representam mudanças conformacionais quando comparada com as respectivas hélices na estrutura de SufE de *Xac*. A tabela destaca o ligante que cada proteína possui, e os aminoácidos conservados na família no caso de membros da família de SufE, ou os aminoácidos envolvidos na coordenação ao ligante.

SufE provavelmente não interage com íons de Zinco como ocorre com membros da família de NifU/IscU. Interessantemente, cinco das seis estruturas pertencentes à família de NifU/IscU possuem um íon de zinco, e os aminoácidos envolvidos na sua coordenação são: CGD, CX₂S e C (**Figura 58 e e 59 b**). A sobreposição deste sítio com a estrutura de SufE_{XAC2355} revelou que se trata da mesma região em que SufE_{XAC2355} contem os aminoácidos dos motivos conservados na família de SufE_{XAC2355}, GC-S, SD e RxxG, (**Figura 59 c**). No entanto, a ligação de zinco por membros da família de SufE não seria possível devido ao fato de que o sítio de interação com Zn⁺² é composto por 3 cisteínas e uma serina, e o sítio ativo de SufE_{XAC2355} contem apenas uma cisteína do motivo GC-S.

Os resíduos de NifU/IscU envolvidos na coordenação ao íon de Zn⁺² são: CGD e C, e estes se sobrepõem aos resíduos dos motivos conservados na família de SufE, GC-S e a arginina do motivo RxxG, respectivamente. Os resíduos CX₂S de IscU/NifU não se sobrepõem a nenhum resíduo na estrutura de SufE_{XAC2355} (**Figura 59 c**). Fica, portanto, evidente que SufE_{XAC2355} não poderia interagir com íons de Zn⁺², mas é intrigante a similaridade estrutural das regiões envolvidas na ligação ao metal por IscU/NifU com o sítio envolvido com a atividade enzimática de SufE.

IscU de PDBID 2z7e coordena centros de Fe-S. A proteína IscU de *A. aeolicus* (PDBID 2z7e), bactéria termofílica²⁷⁵, coordena os centros de Fe-S, [Fe₂-S₂] (**Figura 58 e e 59 d**). Os aminoácidos envolvidos na sua coordenação, C, CXS e HC,

são praticamente os mesmos resíduos em que membros da família de IscU/NifU coordenam o íon de Zn: CGD, CX₂S e C (**Figura 58 e e 59 e**). A histidina do motivo HC, envolvida na coordenação do centros Fe-S, não está presente nas demais proteínas da família IscU/NifU (**Figura 58 e**), com exceção da proteína com PDBID 1wfz, de camundongo. A ausência desta histidina no sítio de interação ao centro de Fe-S aparentemente não evitaria com que estas proteínas sintetizassem os centros de Fe-S.

A similaridade estrutural entre os membros da família SufE com membros da família NifU/IscU revelam uma correlação evolutiva entre estes sistemas. O sistema NifU/IscU interagem com NifS/IscS para formar o sistema de dois componentes, cisteína dessulfarase, assim como ocorre com SufE-SufS. No entanto, NifU/IscU são compazes de incorporar o enxofre obtido de IscS/NifS em centros Fe-S, o que não ocorre com SufE, que transfere seu enxofre para outras proteínas, como SufB do complexo ternário SufB/SufC/SufD²⁴⁸.

A montagem dos centros de Fe-S por IscU aparentemente está ligada a mudanças na estrutura quaternária. A proteína IscU se mantém na forma monomérica quando está reduzida. Quando ela é oxidada por IscS ela forma uma estrutura trimérica, em que há diferenças estruturais entre as moléculas do trímero: apenas um monômero contem o centro Fe-S²⁷¹. Isso sugere que a ligação de IscU/NifU ao zinco estabiliza sua estrutura monomérica que permite a sua interação com IscS/NifS. Apesar de SufE não interagir com moléculas de zinco há evidências que ela se comporta como monômero em solução. SufE_{XAC2355} elui na coluna de gel filtração na forma monomérica, e a superfície enterrada do dímero na unidade assimétrica do cristal de SufE_{XAC2355} é baixo demais para estabilizar representar um dímero biologicamente relevante.

Figura 59: Sobreposição estrutural dos aminoácidos pertencentes ao sítio ativo.

Sobreposição dos aminoácidos do motivo conservado na família de SufE, G₅₉C₆₀-S₆₂ S₈₁D₈₂, R₁₂₆XXG₁₂₉, com os correspondentes nas demais proteínas da família de SufE, PDBID 3g0m (laranja), 1mzg (verde claro), 1ni7 (cyan) e 1wlo (rosa claro) (**a**). Aminoácidos envolvidos na coordenação ao íon cálcio pela proteína IscU/NifU de PDBID 1xjs, em laranja (**b**) e sua sobreposição com o sítio ativo de SufE de Xac (vermelho) (**c**). Aminoácidos envolvidos na coordenação ao centros [Fe₂-S₂] pela proteína IscU/NifU de PDBID 2z7e, em pink (**d**) e sua sobreposição com o sítio ativo de IscU/NifU de PDBID 1xjs (laranja) e 1wfz (amarelo) (**e**)²⁷⁴.

5.1.3. 5 *SufE aumenta atividade cisteína dessulfarase de SufS*

SufS foi inicialmente descrita como uma proteína responsável pela mobilização de íons de selênio provenientes de seleniocisteína^{276,277} e posteriormente foi observado que sua atividade basal de cisteína dessulfarase era drasticamente aumentada na presença de SufE²⁴⁴. Esta atividade cisteína dessulfarase, em que aminoácidos de cisteína são convertidos em alaninas, é dependente de piridoxal fosfato (PLP)²⁷⁸.

As enzimas cisteína dessulfarases são classificadas em dois grupos: Classe I que inclui NifS/IscS/NFS e possui o sítio ativo composto pelos seguintes aminoácidos: SSGSAC(T/S)S e Classe II que inclui SufS/CsdA/CpNifS que possui o sítio ativo composto pelos seguintes aminoácidos: RxGHCA²⁷⁸. Em geral, as bactérias possuem duas cisteínas dessulfarases, uma pertencente ao sistema ISC que é expresso constitutivamente e outro pertencente ao sistema SUF, que é ativado sob estresse oxidativo e em condições de ausência de ferro^{240,241}. No entanto, o genoma de Xac codifica apenas para uma cisteína dessulfarase, SufS_{XAC2938} (**Figura 60 e 46**).

SufE de *E. coli* e de *E. chrysanthemi* aumenta a atividade cisteína dessulfarase de SufS em torno de 10 vezes^{244 88,279}. Resolvemos testar se SufE_{XAC2355} era capaz de estimular a atividade cisteína dessulfarase de SufS_{XAC2398}. As condições de purificação de SufS estão descritas nos Materiais e Métodos. O ensaio de cisteína dessulfarase foi feito a 30 °C durante 3 horas e foram utilizadas diferentes proporções de SufE/SufS, 0:1 até 20:1. A reação foi feita em tampão contendo piridoxal 5'-fosfato e DTT e é iniciada pela adição de cisteína ao meio e terminada pela adição de ácido que reduz e libera os átomos de enxofre, que na presença de N,N-dimetil-p-fenilenediamina e FeCl₃ forma azul de metileno que é medido pela absorbância no comprimento de onda de 670nm^{89,90}. A **Figura 61** demonstra que a atividade de SufS_{XAC2398} é aumentada de 10 a 12 vezes na presença de SufE_{XAC2355}.

Figura 60: Alinhamento múltiplo das seqüências primárias das cisteínas dessulfarases encontradas em *E. coli*: SufS, CsdA e IscS com SufS de *Xac*¹⁶⁴.
E.coli codifica três cisteínas dessulfarases, mas *Xac* codifica apenas para uma, SufS_{XAC2938}.

Atividade Cisteína Desulfarase de SufE/SufS

Figura 61: Atividade cisteína dessulfarase de SufS na presença de SufE.

SufE aumenta de 10 a 12 vezes a atividade cisteína dessulfarase de SufS na proporção de 1:1. A reação foi feita em 160 μ L contendo Tris pH 7,4, NaCl, piridoxal 5'-fosfato, DTT, cisteína (Sigma) e 2,5 μ M de SufE e SufS puras (proporção de 1:1). Posteriormente é adicionado N,N-dimetil-p-fenilenediamina e FeCl_3 (diluído em 1.2M HCl) que forma azul de metileno ($\text{C}_{16}\text{H}_{18}\text{ClN}_3\text{S}$). A adição de ácido ao meio faz com que os átomos de enxofre sejam reduzidos e liberados ao meio, que na presença de N,N-dimetil-p-fenilenediamina e FeCl_3 forma azul de metileno que é medido pela absorbância no comprimento de onda de 670 nm^{89,90}.

5.1.3. 6 Possível regulação gênica do sistema *SUF* por SufR e os possíveis scaffolds para a construção de centros de Fe-S em *Xac*.

Como foi mencionado anteriormente, *Xac* possui apenas uma cisteína dessulfarase, SufS, sugerindo que esta faria também o papel constitutivo de IscS de *E. coli*. No entanto, é possível que há em *Xac* outra forma de regular o seu nível de expressão ou atividade em condições de estresse oxidativo ou na ausência de ferro, já que são nestas condições em que o sistema *SUF* é ativado em *E. coli*²⁴¹.

SufR de *Synechocystis* sp. coordena $[4\text{Fe-4S}]^{2+,1+}$ e atua como um repressor de transcrição. Recentemente foi demonstrado que SufR é um homodímero que coordena dois centros $[4\text{Fe-4S}]^{2+, 1+}$ e atua como um repressor de transcrição do operon *SUF* em *Synechocystis* sp.. Neste organismo, SufR possui alta afinidade ao promotor que regula a transcrição do operon *SUF* e baixa afinidade ao sítio que regula a transcrição de *sufR*^{91,280}. Interessantemente, neste organismo há três cisteína dessulfarases, IscS e SufS e por ser uma bactéria fixadora de nitrogênio também contem NifS. A deleção de *iscS* não é letal para a bactéria, mas a deleção de

sufS é letal^{281,280}, sugerindo que o operon *SUF* seja, nesta bactéria, importante para a biogênese e manutenção dos centros de Fe-S, assim como deve ocorrer em *Xac*.

Shen e colaboradores sugerem que o sítio de interação de SufR ao DNA seja CAAC-N₆-GTTG²⁸⁰. No entanto, há dois sítios de interação na região promotora do operon *SUF*, sugerindo que SufR atue como um tetrâmero²⁸⁰. *Xac* não contém um gene ortólogo ao *sufR*_{SLL0088} de *Synechocystis* sp. PCC 6803 e a análise das regiões promotoras (300 pb antes do códon iniciador de cada gene) dos genes de *sufR*, *sufB* e *sufC* não contém o sítio CAAC-N₆-GTTG, sugerindo que SufR de *Xac* reconheça outra sequência.

Apesar de não haver nenhuma similaridade de sequência primária entre SufR_{SLL0088} de *Synechocystis* sp., que possui 240 aa, com SufR_{XAC2934} de *Xac*, que possui 142 aa, o gene *sufR*_{XAC2934} se localiza no início do operon do sistema *SUF* (**Figura 46**) e codifica a uma proteína que possui o domínio Rrf2, fator de transcrição¹²⁰, e possui 4 cisteínas conservadas, que poderiam coordenar centros de Fe-S, assim como ocorre com SufR_{SLL0088} de *Synechocystis* sp. Provavelmente o mecanismo de ação destas proteínas devem ser correlacionados.

SufR não está envolvida em patogenicidade em Xac. Ensaios funcionais de SufR_{XAC2934} mostrou que a deleção do gene de *sufR*_{XAC2934} em *Xac* (Δ *sufR*_{XAC2934}) não afetou sua patogenicidade (**Figura 62**). Ensaios de sobrevivência de células Δ *sufR*_{XAC2934} em meios contendo peróxido de hidrogênio e quelante de ferro, não tiveram nenhuma diferença quando comparadas com as células tipo selvagem (dados não mostrados).

Purificamos SufR_{XAC2934} expressa em *E. coli* (**Figura 63**) para posteriores ensaios de interação com regiões promotoras dos genes *sufR*, *sufB* e *sufC*. Estes ensaios de gel shift para verificar a sua interação com DNA ainda não foram realizados.

Figura 62: Ensaio de patogenicidade em planta de cepas de *Xac* que contem o gene de *sufR* truncado.

wt, são células de *Xac* tipo selvagem e ΔsufR são células de *Xac* com o gene de *sufR* truncado. Amostras de células de *Xac* tipo selvagem e ΔsufR foram crescidas em 2xTY até D.O_{600nm} de 0.3 e foram infiltradas utilizando seringas descartáveis nas folhas de laranjeiras do tipo Baia doce (*Citrus sinensis* (L.) Osbeck). As plantas foram mantidas na estufa a 28°C. Após 4 dias pode-se observar os primeiros sintomas da doença e após 10 dias a foto foi tirada das folhas inoculadas.

Figura 63: Purificação de SufR.

a) Mini indução de SufR_{XAC2934}. Amostra não induzida (1,2), amostra induzida de SufR (3) **b)** Amostras de SufR eluídas da coluna de gel filtração, mostrando a pureza de SufR. A seta destaca a banda de indução de SufR_{XAC2934}.

5.2 Proteínas de *Xanthomonas axonopodis* pv *citri* envolvidas em *Quorum-sensing* e sinalização por c-diGMP

5.2.1 Introdução

5.2.1.1 *Quorum-Sensing* Mediado por DSF em *Xanthomonas*.

Comunicação célula-célula: *Quorum-sensing* (Q.S.) é o termo utilizado para descrever a habilidade de células bacterianas de detectar a presença de outras ao seu redor e integrar esta informação em redes de controle de diferentes fenótipos bacterianos, tais como motilidade, bioluminescência, transferência de plasmídeo, virulência e formação de biofilme²⁸²⁻²⁸⁴. Portanto, se trata de um sistema intercelular de comunicação entre as bactérias a fim de facilitar a adaptação de uma população em diferentes condições do meio ambiente. Desta forma, os patógenos e fitopatógenos empregam o Q.S. para coordenar sua colonização e infecção na célula hospedeira^{282-284,285,286}.

Q.S. é mediada por sinais químicos (moléculas difusíveis) chamadas de autoindutores (AI). Estes sinais podem ser classificados em duas categorias principais: a) aminoácidos e polipeptídios pequenos, que são usualmente utilizados por bactérias Gram-positivas, e b) os derivados de ácidos graxos, que são usualmente utilizados pelas bactérias Gram-negativas²⁸⁵. Nesta segunda categoria, o que mais predomina é o N-acyl-homoserina lactona (acil HSLs). Apesar de não existir moléculas sinalizadoras universais, os auto-indutores-2 são produzidos por bactérias Gram-positivas e por bactérias Gram-negativas²⁸⁵.

Sistemas de Q.S. Em bactérias, membros da família de LuxI estão envolvidos na formação dos autoindutores da família da acil HSLs, que são específicos para cada organismo. Este AI difunde livremente pela membrana bacteriana até chegar no meio extra-celular, ocasionando o aumento de concentração local do AI de forma proporcional ao número de células. Membros da família de LuxR estão envolvidos na ligação a este AI que provoca a ativação da transcrição de diferentes genes alvos (**Figura 64 a**)²⁸⁷⁻²⁸⁹. Este é apenas um exemplo de um circuito de Q.S, mas existem diversos sistemas de Q.S descritos^{282-284,285}.

Algumas bactérias possuem mais de um sistema de Q.S que a permite distinguir e responder de forma diferente ou equivalente a diferentes AI. Um exemplo deste sistema múltiplo é o que ocorre com *V. harveyi*. Esta bactéria possui dois sistemas de

Q.S e conseqüentemente dois AI, AI-1 e AI-2, produzidos pelas enzimas sintetases LuxLM e LuxS, respectivamente (**Figura 64 b**). AI-1 e AI-2 difundem para o meio extracelular e ligam a suas proteínas sensoras, LuxN²⁹⁰ e LuxP^{291,292}, respectivamente. LuxP está complexada com a proteína quimérica LuxQ e forma, portanto, o complexo LuxPQ-AI-2^{291,292}. Estas interações desencadeiam uma cascata de fosforilação que envolve além das proteínas sensoras, LuxN e LuxPQ, as proteínas LuxU e LuxO^{293,294}. LuxO controla a expressão dos genes alvos, *luxCDABE*^{295,296}.

Figura 64: Sistemas de Q.S²⁹⁷.

a) Sistema LuxRI encontrado em bactérias Gram-Negativas. LuxI está envolvida na síntese do AI da família de N-acyl-homoserina lactona, losango azul escuro. Este AI se liga ao LuxR que está envolvido no controle da transcrição de diferentes genes alvos. **b)** Sistemas de Q.S. encontrado em *V. harveyi*. Os AI, AI-1 (losango verde claro) e AI-2 (esfera verde escura) são sintetizados por LuxLM e LuxS, respectivamente. O AI-1 e o AI-2 ligam a suas proteínas sensoras, LuxN e LuxP (do complexo LuxPQ), respectivamente. Estas interações com os AI desencadeiam uma cascata de fosforilação indicada com a letra P e com setas em vermelho. Esta fosforilação envolve as proteínas sensoras, LuxN e LuxPQ, e as proteínas LuxU e LuxO. LuxO controla a transcrição de diferentes genes alvos²⁹⁷.

Lócus *rpf* e Q.S em *Xac*. A resposta ao Q.S. em *Xanthomonas* é controlada pelas proteínas codificadas pelo lócus *rpf* (“regulatory of pathogenicity factor”)²⁹⁸ e pelo auto indutor DSF. Em *Xcc*, este sistema controla a produção de diferentes fatores de patogenicidade como proteases e glucanases extracelulares e genes envolvidos na biossíntese de EPS (“Extracellular Polysacchacaride” ou goma Xanthana), que são essenciais para a infecção e colonização da bactéria na célula hospedeira.

O lócus *rpf* de *Xac* é muito similar ao de *Xcc* e é constituído por sete genes denominados *rpfA-G*, sendo que os principais componentes formam um cluster que contém *rpfB/rpfF/rpfC/rpfG*. A RpfC é uma proteína quimérica com 4 domínios: a) Receptor ou sensor, localizada na membrana; b) Histidina quinase (HK); c) regulador de resposta (RR); e d) o Histidina fosfotransferase (HPT). A proteína RpfG contém dois domínios, RR e HD-GYP. O domínio HD-GYP atua como uma fosfodiesterase que hidrolisa o segundo mensageiro bacteriano c-diGMP⁴¹, como descrito no **Capítulo 3** item **3.4**.

RpfF e RpfB estão envolvidos na síntese do auto-indutor DSF que se difunde livremente pela membrana bacteriana e se liga ao domínio sensor da RpfC, **Figura 65**. Foi demonstrado que o domínio RR de RpfC interage com RpfF⁷ e que esta interação inibe a produção de DSF²⁹⁹. A ligação do DSF no domínio sensor da RpfC desencadeia a autofosforilação do domínio ATPásico da RpfC, e uma cascata de fosforilação que envolve os resíduos His, Asp e His dos domínios HK, RR e HPT, respectivamente, da RpfC. Este sinal é transferido para o resíduo Asp do domínio RR da RpfG que acarretará na transcrição dos genes envolvidos na patogenicidade^{7,297,299}. A mudança no estado de fosforilação da RpfC provavelmente ocasiona mudanças conformacionais no seu domínio RR que diminui a sua afinidade pela proteína RpfF (que na forma livre sintetiza mais DSF) e aumenta a sua afinidade pela proteína RpfG²⁹⁷⁻³⁰⁰.

O sistema de Q.S. dependente do DSF e difere dos sistemas de Q.S. bem caracterizados, como o sistema LuxIR, no sentido de que a produção do AI é autoregulada, ou seja, os genes envolvidos na síntese do AI são induzidos pela presença do AI. No caso do sistema *rpf* a síntese de RpfF aumenta de forma dependente a densidade celular^{301,302}. Outra importante diferença é a regulação da atividade da RpfF pela RpfC, RpfF ligada a RpfC é inativa e na ausência da RpfC é ativa (**Figura 65**)²⁹⁹.

Envolvimento do segundo mensageiro bacteriano c-diGMP com quorum sensing em *Xanthomonas*. Nosso grupo demonstrou que o domínio HD-GYP da RpfG interage com nove das trinta proteínas codificadas pelo genoma de *Xac*⁷. Interessantemente, todas estas proteínas tem domínios provavelmente ativos enzimaticamente (como descrito na **Figura 18** e **Tabela 14**) e também se destacam por não possuírem domínios sinalizadores encontrados nas outras proteínas como: RR, PAS, GAF, CHASE e MHYT (**Figura 66**). A interação de dois domínios com atividades antagônicas é intrigante e a atividade destes domínios devem ser reguladas de tal forma que apenas um deles esteja ativo de forma dependente das condições ambientais. Conseqüentemente, a concentração local do segundo mensageiro bacteriano c-diGMP deve variar de forma dependente de qual domínio está ativo. Esta variação de concentração de c-diGMP local poderia regular outras proteínas com domínios receptoras ao c-diGMP. Portanto, a atividade de outras proteínas receptoras a esse segundo mensageiro poderiam ser reguladas de forma indireta pelo DSF.

Figura 65. Modelo do funcionamento do sistema *rpf* por DSF na regulação da produção de fatores de patogenicidade em *Xanthomonads*.

(baseado no modelo proposto por Slater *et al.*²⁹⁸, Andrade *et al.*⁷ e He *et al.*²⁹⁹). A autofosforilação da RpfC, ativada pelo DSF, ocasiona uma cascata de fosforilação no sentido His-Asp-His que tem como finalidade transferir seu fosfato para o domínio RR da RpfG. O domínio HD-GYP da RpfG degrada c-diGMP mas os domínios GGDEF que HD-GYP interage sintetiza c-diGMP a partir de duas moléculas de GTP. Os resíduos de His e Asp envolvidos são indicados com as letras H e D, respectivamente. O grupo fosforil é indicado com a letra P. A forma fosforilada da RpfG regula positivamente a expressão dos genes requeridos na infecção. Não sabe se é de forma direta ou indireta, como endoglucanases (engXCA), proteases (prtA) e o *cluster gum*. MI, membrana interna, ME, membrana externa da bactéria.

Figura 66. Esquema dos domínios encontrados nas proteínas de *Xac* que possuem o domínio GGDEF.

As proteínas numeradas em rosa representam aquelas que interagir com o domínio HD-GYP da RpfG⁷.

Xanthomonas sp. respondem diferentemente aos sinais de Q.S. Trabalhos recentemente publicados sugerem que diferentes gêneros de *Xanthomonas* respondem de forma diferente aos sinais de Q.S.. Isto sugere que cada gênero de *Xanthomonas* se adaptou as necessidades de colonização e proliferação nas células hospedeiras³⁰³. Em *Xcc* assim como em *Xac*, células nocautes de *rpfC* ou *rpfF* ocasionam a diminuição na virulência pela diminuição da síntese de fatores da patogenicidade incluindo proteases, endoglucanases, endomananase e goma Xanthana^{298,301,304}. Células nocautes de *rpfC* possui aumento nos níveis de DSF, mas não na transcrição de *rpfF*, sugerindo que a RpfC é um inibidor da síntese de DSF pela interação com RpfF, como proposto por nosso grupo. A interação do DSF a RpfC pode ocasionar mudanças estruturais no complexo, RpfC-RpfF, que permita que RpfF sintetize DSF^{7,298}.

No entanto, em *Xylella fastidiosa* células nocautes de *rpfC*, aumenta a transcrição do gene *rpfF* e conseqüentemente eleva os níveis de DSF, provocando diminuição na patogenicidade. Além de ativar a síntese de *rpfF* também ocorre um aumento da expressão de adesinas, como FimA, HxfA, HxfB, aumentando a adesão das células e inibindo-as de migrarem no xilema de vinheiras³⁰⁵⁻³⁰⁸. De forma inesperada, células de *Xylella fastidiosa* nocautes de *rpfF* são mais patogênicas que as células tipo selvagem³⁰⁵⁻³⁰⁸. Esta diferença de fenótipo ocorre devido ao fato de que em altos níveis de DSF ocorre a produção de adesinas em *Xylella fastidiosa*, o que não deve ocorrer em *Xcc* e nem em *Xac*. Em compensação, em baixos níveis de DSF as células de *Xylella fastidiosa* não devem produzir adesinas e, portanto, devem se mover mais rapidamente facilitando seu espalhamento no xilema da planta hospedeira.

Células de *Xylella fastidiosa* duplamente nocauteadas por *rpfC* e *rpfF* tem o mesmo fenótipo que células tipo selvagem. Como a síntese de adesinas foi observada por células nocautes de *rpfC* mas não pelas células duplamente nocauteadas, *rpfC* e *rpfG*, Chatterjee e colaboradores sugerem que haja outro receptor de DSF na superfície da bactéria que ative e fosforile RpfG para que ocorra a síntese de biofilme de forma independente da RpfC³⁰⁵⁻³⁰⁸.

5.2.2 Resultados e Discussões

5.2.2.1 Purificação das proteínas RpfF, RpfG, CMF e dos Domínios RR e HPT da RpfC

As proteína RpfF_{XAC1879}, RpfG_{XAC1877}, CMF_{XAC0330-his-(99-654)} e os domínios RR (RpfC_{his-RR(415-613)}) e HPT (RpfC_{his-HPT(584-725)}) da RpfC_{XAC1878} foram purificados seguindo os protocolos descritos nos Materiais e Métodos. A **Figura 67** mostra o gel de SDS-PAGE 15% da etapa final de purificação de cada proteína. Todas as proteínas foram submetidas a ensaios de cristalização e apenas o domínio HPT da RpfC (RpfC_{his-HPT(584-699)}) cristalizou, como descrito a seguir.

RpfC_{his}-HPT(584-725)

Ma 33 35 37 39 41 43 45 47 49 51 53 55 57

RpfC_{his}-RR(415-613)

69 67 65 I 64 63 61 59 57 54 50 47

RpfF_{XAC1879}

NI I GF1 GF2

RpfG_{XAC1877}

24 23 22 21 19 18

CMF_{XAC0330-his(99-654)}

Ma 25 27 29 31 33 35 37 39 48 51

RpfG_{XAC1877}, RpfF_{XAC1879} e Diferentes
construções da RpfC_{XAC1878}

1 2 3 4 5

Figura 67: Gel de 15%SDS-PAGE referente a última etapa de purificação das proteínas RpfF, RpfG, RpfC e CMF e dos domínios RR e HPT da RpfC.

NI: amostra não induzida; **I:** amostra induzida; **Ma**, proteína pura. Os números se referem aos tubos da eluição na coluna de gel filtração. Os números em verde são as frações que foram selecionadas para testes de cristalização. **GF1** e **GF2** do gel de SDS-PAGE da RpfF_{XAC1879} são as amostras puras eluídas na primeira e na segunda aplicação na coluna de gel filtração. O gel de SDS-PAGE com título “RpfG_{XAC1877}, RpfF_{XAC1879} e Diferentes construções da RpfC_{XAC1878}” são amostras puras das proteínas: RpfG_{XAC1877} (2), RpfF_{XAC1877} (3), RpfC_{his}-(173-725) (4) e RpfC_{his}-(276-725) (5) e 1, Marcador de peso molecular. A massa molar da RpfG_{XAC1877} é de 42,2KDa, da RpfF_{XAC1877} é de 32 KDa, CMF_{XAC0330-his(99-557)} é de 50 KDa, RpfC_{his}-(173-725) é de 61 KDa, RpfC_{his}-(276-725) é de 51 KDa, RpfC_{his}-HPT(584-725) é de 17,4 KDa e RpfC_{his}-RR(415-613) é de 16,7 KDa.

5.2.2.2 Cristalização do Domínio HPT da RpfC.

Cristais pequenos do domínio HPT da RpfC (RpfC_{his-HPT(584-699)}) cresceram na concentração de 6,32 g/L na solução do poço 9 do Kit Crystal Screen I da empresa Hampton (2M sulfato de amônio, 0,1M citrato de sódio pH 5,6 e 200 mM tartarato de sódio potássio) (**Figura 68**). A otimização das condições de cristalização foram feitas variando a concentração do agente precipitante e do pH. As melhores condições de cristalização foram em 2M sulfato de amônio, 0,1M citrato de sódio pH 6,4, 200 mM tartarato de sódio potássio e 1mM sulfato de zinco (**Figura 68**).

O conjunto de dados de difração de raios-X dos cristais da RpfC_{his-HPT(584-729)} foram coletados na linha MX1 do LNLS, utilizando o detector MARCCD. Não foi necessária a utilização de crioprotetante. O cristal foi rapidamente congelado em um fluxo de nitrogênio a 100K e mantido nesta temperatura durante a coleta. Um destes cristais difratou até 2,6 Å. Na tentativa de obter cristais ainda melhores do domínio HPT da RpfC tratamos a proteína com trombina para a retirada da cauda de histidina que estava fusionada ao N-terminal da proteína. Refizemos os testes de cristalização com a proteína RpfC_{HPT(584-699)} sem cauda de histidina e observamos cristais apenas na condição onde a proteína havia cristalizado (**Figura 69 a e b**). O tempo necessário para observamos os primeiros cristais da proteína não mudou, que é de aproximadamente três meses. A adição de sulfato de zinco, que antes fazia os cristais de RpfC_{his-HPT(584-729)} serem melhores, resultou na precipitação da proteína.

Os cristais da RpfC_{HPT(584-699)} sem cauda de histidina são maiores e um cristal difratou até 2.2 Å. Os dados do processamento de dados das imagens de difração de raios-X estão mostrados na **Tabela 23**. Para este conjunto de dados, foram coletados apenas 78 imagens porque havia acabado o nosso tempo de linha no LNLS. Este conjunto de dados foi o terceiro conjunto de dados obtido do mesmo cristal depois de seguidos tratamentos com *annealing* (descongelamento e congelamento rápido do cristal na linha de raios-X). Este procedimento melhorou muito a qualidade do padrão de difração assim como a resolução da difração.

Nós tentamos inicialmente resolver as fases do cristal utilizando a técnica de substituição molecular a partir da estrutura do domínio HPT da proteína ArcB de *E. coli* depositadas no PDB (estrutura cristalina, PDBID: 2A0B³⁰⁹, e estrutura resolvida por RMN, PDBID: 1FR0³¹⁰). Várias tentativas foram feitas sem êxito, possivelmente devido a baixa similaridade de seqüência primária, 16% de identidade, entre estes modelos

estruturais com o domínio HPT da RpfC (**Figura 70**). Apesar de haver baixa similaridade de sequência primária entre estas proteínas, a estrutura secundária predita da RpfC_{his-HPT(584-725)} é bastante similar com a estrutura secundária da estrutura do domínio HPT de ArcB, ArcB_{HPT} (**Figura 70**), mostrando que ambas as proteínas devem ter a mesma topologia.

A estrutura de ArcB_{HPT} é composta por 6 hélices, $\alpha 1$ - $\alpha 6$, em que as quatro últimas hélices formam o motivo “four helix bundle”^{309,310} (**Figura 71 a**). A estrutura de ArcB_{HPT} em solução não contem a $\alpha 1$ sugerindo que esta hélice seja móvel³¹⁰ (**Figura 71 b**). A histidina envolvida na transferência de grupos fosforil para seus parceiros, domínios RR, está localizada na $\alpha 4$ e está exposta ao solvente na estrutura de ArcB_{HPT} (**Figura 71 a-b**).

Outras estruturas de domínios HPT foram resolvidas nos últimos anos. Por exemplo, a estrutura do domínio HPT de Ypd1 em complexo com o domínio RR de Sdn1 de *S. cerevisiae*, assim como as estruturas dos domínios HPT das proteínas YojN e BarA de *E. coli*. Apesar de haver baixa similaridade de sequência primária entre estes domínios (menos de 20% de identidade), todos possuem o motivo “four helix bundle” conservados (**Figura 71 c-e**). Por isso, é altamente provável que o domínio HPT da RpfC também adote esta topologia.

Na tentativa de resolver as fases coletamos dados de difração dos cristais do domínio HPT (RpfC_{his-HPT(584-725)}) dopado com európio na linha MX1 do LNLS, que possui comprimento de onda de 1,425Å, mas não foi observado nenhum sinal anômalo deste átomo. Tentamos, também, observar o sinal anômalo de zinco, mas o seu sinal no comprimento de onda que os dados foram coletados é muito baixo e também não foi possível observar nenhum sinal anômalo dos conjuntos de dados coletados para os cristais de RpfC_{his-HPT(584-725)} crescidos na presença de zinco.

Desta forma, expressamos o RpfC_{his-HPT(584-725)}, HPT-SeMet, removemos sua cauda de histidina e tentamos cristalizá-la (**Figura 69 c e d**). No entanto, os cristais de HPT-SeMet são ainda menores do que os cristais tipo selvagem e não difrataram. Devido ao seu tamanho, pedimos tempo para utilizar a linha de microfoco, PROXIMA 1, do Sincrotron Soleil na França. Infelizmente, no transporte dos cristais do Brasil para a França, nossos cristais foram destruídos devido à quebra do Dry Shipper. Não tivemos tempo para refazer os cristais do domínio HPT-SeMet devido ao seu longo tempo de cristalização.

Figura 68: Cristais do domínio HPT da RpfC.

Cristais do domínio HPT ($RpfC_{\text{His-HPT}(584-725)}$), **a)** cristais iniciais ; **b)** após otimização das condições de cristalização com sulfato de zinco (2M sulfato de amônio, 0,1M citrato de sódio pH 6,4, 200 mM tartarato de sódio potássio e 1mM sulfato de zinco).

Figura 69: Cristais do domínio HPT da RpfC sem cauda de histidina e incorporada com selênio metionina.

a e b) Cristais do domínio HPT da RpfC ($RpfC_{\text{HPT}(584-725)}$) sem cauda de histidina. **c e d)** Cristais do domínio HPT da RpfC incorporada com selênio metionina sem cauda de histidina. As condições de cristalização são 2M sulfato de amônio, 0,1M citrato de sódio pH 6,4 e 200 mM tartarato de sódio potássio.

Tabela 23: Dados estatísticos do processamento dos dados de difração de raios-X do domínio HPT da RpfC, (RpfC_{HPT(584-725)}).

Grupo Espacial	R3
<i>a</i> (Å)	52.185
<i>b</i> (Å)	52.185
<i>c</i> (Å)	90.000
Dados Estatísticos	Nativo
Resolução (Å)¹	35 - 2.2 (2.28 - 2.20)
N. reflexões observadas	8185
N. reflexões únicas	3947
<I/erro>¹	15.1 (4.37)
Multiplicidade¹	2.1 (1.3)
Completeza (%)¹	73.5 (24.7)
R (%)^{1†}	0.047 (0.110)
N. imagens	78
λ (Å)	1.425

¹Os valores em parênteses se referem aos dados da última camada.

$$^{\dagger}R=(\sum(I-\langle I \rangle)^{**2})/\sum(I)^{**2}$$

```

RpfChis-HPT_584-725_   VRQTDAAIGKRGPATNGAVLDPVLDLAEELGMGEFFEREFIQQCLSDATACVVELGKAA 60
2A0B_HPT_ArcB_Ecoli   -----TTEENSKSEALLDIPMLEQYLEL-VGPKLITDGLAVFEKMMPGYVSVLESNL 51
                        .. :.. *** :.*** ** :* : : : . . . * * .

RpfChis-HPT_584-725_   ADEDWGRLEHHAIVKGVSSNLGLVILAEQGGELMRLPEAQLRSEWRQRLEQLKTGLAQQ 120
2A0B_HPT_ArcB_Ecoli   LAQDKKGIVEECETIKGAAGSVGLRHLQQLGQQIQSPDLPAWEDNVGEWIEEMKEEWRHD 111
                        : : * : * . * : ** : . : ** * : * : : . . . : : *** * : .

RpfChis-HPT_584-725_   REALAQRARRKTEGGASGNGDP 142
2A0B_HPT_ArcB_Ecoli   VEVLKAWVAKATKK----- 125
                        * . * . : * :

```

Figura 70: Alinhamento da sequência da RpfC_{his-HPT(584-725)} com a ArcB_{HPT} de *E. coli*.

RpfC_{his-HPT(584-725)} e ArcB_{HPT} possuem 16% de identidade. A estrutura secundária do domínio HPT de ArcB (PDBID: 2A0B) e a predição da estrutura secundária da RpfC_{his-HPT(584-725)} estão mostrados de tal forma que os resíduos presentes em alfa-hélices estão coloridos em vermelho e os presentes em fita-betas estão coloridos em amarelo³¹¹. A caixa em azul destaca a histidina que é fosforilada nos domínios HPT³⁰⁹. Alinhamento feito utilizando o programa ClustalW¹⁶⁴.

Figura 71: Estruturas de domínios HPT das proteínas ArcB, BarA e YojN de *E.coli* e Ypd1 de *S. cerevisiae*.

a) Estrutura cristalina do domínio HPT de ArcB de *E.coli* (PDBID: 2A0B, cadeia A) e sua sobreposição com a estrutura resolvida em solução (PDBID: 1FR0, modelo 1) (**b**). As alfa-hélices estão coloridas em vermelho e os loops em verde, a histidina que é fosforilada está colorida em amarelo, His 115 na estrutura de PDBID: 2A0B. Sobreposição da estrutura do domínio HPT de ArcB de *E.coli* (PDBID: 2A0B), colorida em vermelho, com as estruturas dos domínios HPT das proteínas YojN (PDBID: 1SR2, cadeia A) colorida em rosa (**c**), Ypd1 (PDBID: 2R25, cadeia A) colorida em azul escuro, e a inserção entre as $\alpha 5$ e $\alpha 6$ (com base na estrutura de ArcB, item **a**) está colorida em azul claro (**d**), e BarA (PDBID: 3IQT, cadeia A) colorida em salmão (**e**). Nota-se, itens de **c-e**, que todos os domínios HPT possuem o motivo “four helix bundle” conservados.

5.2.3 Conclusões

Dentre todas as proteínas pertencentes ao sistema *rpf* estudadas, apenas o domínio HPT da RpfC (RpfC_{his-HPT(584-725)} e RpfC_{HPT(584-725)}) cristalizou, mas não obtivemos êxito na determinação de sua estrutura tridimensional utilizando a técnica MAD, devido ao pequeno tamanho dos cristais.

CONCLUSÕES FINAIS

Tivemos como objetivo central a determinação estrutural de várias proteínas do fitopatógeno *Xanthomonas axonopodis* pv citri utilizando como técnica principal difração de raios-X. Utilizamos RMN, ensaio de 2-híbridos, Western- e Far-Western, mutação sítio dirigida, nocaute gênico, a fim de entender mais detalhadamente suas funções biológicas.

Interações entre PilZ, FimX e PilB, proteínas reguladoras da biogêneses do pilus tipo IV. A comparação da estrutura de PilZ_{XAC1133} com outras estruturas de PilZ complexadas ao segundo mensageiro bacteriano c-diGMP revelou que PilZ_{XAC1133} exibe diferenças na sua estrutura secundária e terciária que abolem o sítio de interação a este dinucleotídeo e explica porque membros da família PA2960/PilZ_{XAC1133} não interagem com c-diGMP⁶. O assinalamento da cadeia principal de PilZ por RMN possibilitou a predição da estrutura secundária da proteína em solução, que condiz perfeitamente com aquela determinada por cristalografia de raios-X⁶.

Ensaio de 2-híbridos revelaram dois parceiros de PilZ_{XAC1133} envolvidos na biogênese do T4P: PilB_{XAC3239} e FimX_{XAC2398}. Estes resultados são a primeira evidência que PilZ_{XAC1133} interage fisicamente com a maquinaria de biogênese do T4P. Ensaio de interação proteína-proteína mostraram que PilZ_{XAC1133} interage especificamente com o domínio EAL de FimX_{XAC2398}. Ensaio de mutação sítio dirigida com o objetivo de identificar a interface de interação de PilZ_{XAC1133} com PilB_{XAC3239} e com FimX_{XAC2398}, revelou que o C-terminal de PilZ_{XAC1133} faz parte da interface de interação com PilB_{XAC3239}, mas não com FimX_{XAC2398}, e a Y22 provavelmente tem o papel de regular a interação de PilZ_{XAC1133} com FimX_{XAC2398} e com PilB_{XAC3239}⁶.

PilZ_{XAC1133} forma um complexo ternário com FimX_{XAC2398} e com PilB_{XAC3239} de forma independente da interação de FimX_{XAC2398} ao c-diGMP⁶. Este complexo ternário contém um conjunto de domínios degenerados, ou seja, o domínio RR de FimX_{XAC2398} não contém o aspartato conservado para ser fosforilado, os domínios GGDEF e EAL de FimX_{XAC2398} são enzimaticamente inativos e PilZ_{XAC1133} não é um receptor de c-diGMP. Estas observações sugerem que a regulação da atividade destas proteínas não devem ser mais via o segundo mensageiro bacteriano, mas sim, via interações proteína-proteína. Com base em todos estes resultados propusemos um mecanismo de ação de PilZ_{XAC1133} na biogênese do T4P, com foco no complexo ternário FimX-PilZ-PilB.

Tanto a estrutura de PilZ_{XAC1133} quanto a sua interação com dois componentes com papéis na biogênese do T4P são inéditas. Estas informações podem contribuir para o conhecimento do mecanismo de formação do T4P, importante em vários processos bacterianos, como motilidade *twitching*, transformação natural, formação de biofilme, secreção de proteínas e quimitaxia^{3,4,115,116,122}.

YaeQ pertence à família de endonucleases PD-(D/E)XK. Membros da família de YaeQ são altamente conservados em proteobactérias Gram-negativas, mas podem ser encontrados em alguns organismos eucarióticos. Com a resolução da estrutura de YaeQ_{XAC2396} demonstramos que a família de YaeQ é um membro da família de PD-(D/E)XK encontrado em endonucleases dependentes de magnésio. O sítio putativo catalítico de YaeQ é P₆₇(D/E)₆₈-W₈₁X(D/E)₈₃XG₈₅XP₈₇X₃R₉₁ em que os dois resíduos acídicos são responsáveis pela coordenação a dois cátions de magnésio envolvidos na ativação da molécula de água que irá atacar nucleofílicamente o fosfato a ser clivado. A G₈₅ e a P₈₇ provavelmente tem o papel na estruturação do loop $\beta 5$ - $\alpha 2$, que conecta vários resíduos do sítio catalítico. A R₉₁ e a R₃₉ possivelmente tem o papel de interagir com o fosfato anterior e posterior ao do fosfato que será clivado.

Ortólogos da família de YaeQ são caracterizados pela presença de quatro motivos de sequência conservados: 1- *Motivo AXHPSE*; 2- *Motivo acídico DEP(D/E)*, pertencente ao sítio catalítico de YaeQ; 3- G₈₅ e 4- R₁₄₈. Todos estes resíduos se localizam em uma face da estrutura de YaeQ, sugerindo que estes devam participar da interface de interação de YaeQ com outras proteínas ou com o seu substrato, ainda desconhecido. O motivo AXHPSE pode desempenhar a função de contactar a cavidade menor do DNA e assim, estar envolvido na especificidade de YaeQ ao seu substrato.

A determinação da estrutura de YaeQ foi importante para apontar hipóteses sobre sua função. Dentre os resultados funcionais realizados para YaeQ, obtivemos fortes evidências de que YaeQ está envolvida em alguma via de reparo de DNA em Xac. Novos ensaios biológicos devem ser feitos para desvendar o substrato de YaeQ e em que processo de reparo de DNA YaeQ pertence.

Outros alvos estruturais. Além dos resultados acima descritos, determinamos a estrutura de SufE_{XAC2355} e mostramos que esta aumenta a atividade cisteína desulfarese de seu parceiro, SufS_{XAC2938}, em torno de 10-12 vezes, como ocorre com SufE-SufS de *E.coli*^{244,248}. Membros da família de SufE possuem quatro motivos de

seqüência conservados: 1- G₅₉C₆₀XS₆₂; 2- S₈₁D₈₂; 3- G₈₈ e 4- R₁₂₆XXG₁₂₉, em que os números correspondem aos aminoácidos da proteína de SufE_{XAC235}. Os aminoácidos destes motivos compõe o sítio catalítico de SufE_{XAC235} e formam uma rede de interações que envolve principalmente os resíduos D₈₂, S₈₁, S₆₂, R₁₂₆ e C₆₀. Com base nesta rede de interações propusemos alguns mecanismos de ação de redução da C₆₀ no sítio ativo de SufE_{XAC235} e posterior oxidação desta por SufS-SH.

Clonamos, expressamos, purificamos e fizemos ensaios de cristalização de várias proteínas envolvidas em sinalização por c-diGMP e Q.S. em *Xac*. Conseguimos apenas cristalizar o domínio HPT da RpfC, sensora-histidina quinase chave da via de Q.S. em *Xac*. No entanto, não determinamos a sua estrutura tridimensional.

Ensaio de 2-híbridos do domínio GGDEF de XAC0258 mostrou que este domínio interage com os domínios de sinalização das proteínas MCPs, XAC1892 e XAC3768. No entanto, não obtivemos confirmações destas interações através de ensaios de interação proteína-proteína *in vitro*. Desta forma, novos ensaios devem ser realizados para comprovar estas interações.

REFERÊNCIAS

1. Galvao-Botton, L.M. et al. High-throughput screening of structural proteomics targets using NMR. *FEBS Lett* **552**, 207-13 (2003).
2. Guzzo, C.R., Silva, L.R., Galvao-Botton, L.M., Barbosa, J.A. & Farah, C.S. Expression, crystallization and preliminary crystallographic analysis of SufE (XAC2355) from *Xanthomonas axonopodis* pv. citri. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **62**, 268-70 (2006).
3. Alm, R.A., Boder, A.J., Free, P.D. & Mattick, J.S. Identification of a novel gene, pilZ, essential for type 4 fimbrial biogenesis in *Pseudomonas aeruginosa*. *J Bacteriol* **178**, 46-53 (1996).
4. Ryjenkov, D.A., Simm, R., Romling, U. & Gomelsky, M. The PilZ domain is a receptor for the second messenger c-di-GMP: the PilZ domain protein YcgR controls motility in enterobacteria. *J Biol Chem* **281**, 30310-4 (2006).
5. Huang, B., Whitchurch, C.B. & Mattick, J.S. FimX, a multidomain protein connecting environmental signals to twitching motility in *Pseudomonas aeruginosa*. *J Bacteriol* **185**, 7068-76 (2003).
6. Guzzo, C.R., Salinas, R.K., Andrade, M.O. & Farah, C.S. PILZ protein structure and interactions with PILB and the FIMX EAL domain: implications for control of type IV pilus biogenesis. *J Mol Biol* **393**, 848-66 (2009).
7. Andrade, M.O. et al. The HD-GYP domain of RpfG mediates a direct linkage between the Rpf quorum-sensing pathway and a subset of diguanylate cyclase proteins in the phytopathogen *Xanthomonas axonopodis* pv citri. *Mol Microbiol* **62**, 537-51 (2006).
8. Hauk, P., Guzzo, C.R., Ho, P.L. & Farah, C.S. Crystallization and preliminary X-ray analysis of LipL32 from *Leptospira interrogans* serovar Copenhageni. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **65**, 307-9 (2009).
9. Hauk, P., Guzzo, C.R., Roman Ramos, H., Ho, P.L. & Farah, C.S. Structure and calcium-binding activity of LipL32, the major surface antigen of pathogenic *Leptospira* sp. *J Mol Biol* **390**, 722-36 (2009).
10. Hasse, G. *A Laranja no Brasil*, (1987).
11. Abecitrus. http://www.abecitrus.com.br/historia_br.html.
12. Bevan, M. Plant pathology. The bugs from Brazil. *Nature* **406**, 140-1 (2000).
13. NEVES, E.M., DAYOUB, M., DRAGONE, D.S. & NEVES, M.F. CITRICULTURA BRASILEIRA: EFEITOS ECONÔMICOFINANCEIROS, 1996 – 20001.
14. Page, I.H. <http://www.iapar.br/>.
15. Page, F.H. <http://www.fundecitrus.com.br/>.
16. Schubert, T.S. et al. Meeting the challenge of eradicating citrus canker in Florida - Again. *Plant Disease* **85**, 340-356 (2001).
17. Gottwald, T.R., Granam, J.H. & Schubert, T.S. Citrus Canker: The Pathogen and Its Impact. *Plant Health Progress* (2002).
18. Rossetti, V. Citrus Canker in Latin America: A review. *Proceedings of the International Society of Citriculture* **3**, 918-924 (1997).
19. Nascimento, J.F., Rodrigues, N.J., Alves, J.M.A., Rego, M. & Araujo, A.E.S. Ocorrência de cancro cítrico no estado de Rondônia *Anais do congresso Paulista de Fitopatologia*, 26 (2003).
20. Schubert, T.S. & Sun, X. Bacterial Citrus Canker. *Plant Pathology Circular* **No. 377**(2003).

21. Brunings, A.M. & Gabriel, D.W. *Xanthomonas citri*: breaking the surface. *Molecular Plant Pathology* **4**, 141-154 (2003).
22. Lowe, d. CURRENT SITUATION, MANAGEMENT AND ECONOMIC IMPACT OF CITRUS CANCER IN FLORIDA. <http://www.calcitrusquality.org/wp-content/uploads/2009/05/current-situation2.pdf> (2009).
23. M.M. Dewdney & J.H. Graham. 2010 Florida Citrus Pest Management Guide: Citrus Canker. (ed. <http://edis.ifas.ufl.edu/pdf/CG/CG04000.pdf>) This document is PP-182 (Plant Pathology Department, Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida. , 2009).
24. L. W. Timmer, J. H. Graham & Chamberlain, H.L. Fundamentals of Citrus Canker Management. in <http://edis.ifas.ufl.edu/pdf/PP/PP15300.pdf> (Plant Pathology Department, Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida., 2006).
25. da Silva, A.C. et al. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature* **417**, 459-63 (2002).
26. Buttner, D. & Bonas, U. Getting across--bacterial type III effector proteins on their way to the plant cell. *EMBO J* **21**, 5313-22 (2002).
27. Buttner, D. & Bonas, U. Port of entry--the type III secretion translocon. *Trends Microbiol* **10**, 186-92 (2002).
28. Christie, P.J. Type IV secretion: intercellular transfer of macromolecules by systems ancestrally related to conjugation machines. *Mol Microbiol* **40**, 294-305 (2001).
29. Hueck, C.J. Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol Mol Biol Rev* **62**, 379-433 (1998).
30. Plano, G.V., Day, J.B. & Ferracci, F. Type III export: new uses for an old pathway. *Mol Microbiol* **40**, 284-93 (2001).
31. Baron, C., D, O.C. & Lanka, E. Bacterial secrets of secretion: EuroConference on the biology of type IV secretion processes. *Mol Microbiol* **43**, 1359-65 (2002).
32. Burns, D.L. Type IV transporters of pathogenic bacteria. *Curr Opin Microbiol* **6**, 29-34 (2003).
33. Alegria, M.C. et al. Identification of new protein-protein interactions involving the products of the chromosome- and plasmid-encoded type IV secretion loci of the phytopathogen *Xanthomonas axonopodis* pv. *citri*. *J Bacteriol* **187**, 2315-25 (2005).
34. Aldon, D., Brito, B., Boucher, C. & Genin, S. A bacterial sensor of plant cell contact controls the transcriptional induction of *Ralstonia solanacearum* pathogenicity genes. *EMBO J* **19**, 2304-14 (2000).
35. Brito, B., Aldon, D., Barberis, P., Boucher, C. & Genin, S. A signal transfer system through three compartments transduces the plant cell contact-dependent signal controlling *Ralstonia solanacearum* hrp genes. *Mol Plant Microbe Interact* **15**, 109-19 (2002).
36. Wang, T.T., Wu, X.H., Zhang, S.L. & Chan, J.S. Molecular mechanism(s) of action of norepinephrine on the expression of the angiotensinogen gene in opossum kidney cells. *Kidney Int* **54**, 785-95 (1998).
37. Buttner, D. & Bonas, U. Regulation and secretion of *Xanthomonas* virulence factors. *FEMS Microbiol Rev* (2009).
38. Stock, A.M., Robinson, V.L. & Goudreau, P.N. Two-component signal transduction. *Annu Rev Biochem* **69**, 183-215 (2000).

39. Galperin, M.Y. Bacterial signal transduction network in a genomic perspective. *Environ Microbiol* **6**, 552-67 (2004).
40. Docena, C. Identificação das Interações envolvendo proteínas relacionadas aos Sistemas de Dois-Componentes e os Sistemas Secretórios do Tipo III e Tipo IV do fitopatógeno *Xanthomonas axonopodis* pv citri. Tese (2006).
41. Ryan, R.P. et al. Cell-cell signaling in *Xanthomonas campestris* involves an HD-GYP domain protein that functions in cyclic di-GMP turnover. *Proc Natl Acad Sci U S A* **103**, 6712-7 (2006).
42. Chan, J.W. & Goodwin, P.H. The molecular genetics of virulence of *Xanthomonas campestris*. *Biotechnol Adv* **17**, 489-508 (1999).
43. Vitkup, D., Melamud, E., Moulton, J. & Sander, C. Completeness in structural genomics. *Nat Struct Biol* **8**, 559-66 (2001).
44. PDB.
http://www.rcsb.org/pdb/static.do?p=general_information/pdb_statistics/index.html. (2009).
45. Lundstrom, K. Structural genomics and drug discovery. *J Cell Mol Med* **11**, 224-38 (2007).
46. Sakakibara, D. et al. Protein structure determination in living cells by in-cell NMR spectroscopy. *Nature* **458**, 102-5 (2009).
47. Inomata, K. et al. High-resolution multi-dimensional NMR spectroscopy of proteins in human cells. *Nature* **458**, 106-9 (2009).
48. Sakai, T. et al. In-cell NMR spectroscopy of proteins inside *Xenopus laevis* oocytes. *J Biomol NMR* **36**, 179-88 (2006).
49. Selenko, P. et al. In situ observation of protein phosphorylation by high-resolution NMR spectroscopy. *Nat Struct Mol Biol* **15**, 321-9 (2008).
50. Stoll, V. et al. X-ray crystallographic structure of ABT-378 (lopinavir) bound to HIV-1 protease. *Bioorg Med Chem* **10**, 2803-6 (2002).
51. Varghese, J.N., Hrmova, M. & Fincher, G.B. Three-dimensional structure of a barley beta-D-glucan exohydrolase, a family 3 glycosyl hydrolase. *Structure* **7**, 179-90 (1999).
52. Schindler, T. et al. Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. *Science* **289**, 1938-42 (2000).
53. Campbell, S.F. Science, art and drug discovery: a personal perspective. *Clin Sci (Lond)* **99**, 255-60 (2000).
54. Sambrook, J. & Russel, D.W. "Molecular Cloning a laboratory manual" 1.116., (2001).
55. Sanger, F., Nicklen, S. & Coulson, A.R. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463-7 (1977).
56. Roberts, R.C., Mohr, C.D. & Shapiro, L. Developmental programs in bacteria. *Curr Top Dev Biol* **34**, 207-57 (1996).
57. Dickon, A. To be publish.
58. De Feyter, R., Yang, Y. & Gabriel, D.W. Gene-for-genes interactions between cotton R genes and *Xanthomonas campestris* pv. malvacearum avr genes. *Mol Plant Microbe Interact* **6**, 225-37 (1993).
59. Studier, F.W. & Moffatt, B.A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189**, 113-30 (1986).
60. Studier, F.W., Rosenberg, A.H., Dunn, J.J. & Dubendorff, J.W. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* **185**, 60-89 (1990).

61. Berne, P.F., Doublié, S. & Carter, C.W.J. *Crystallization of Nucleic Acids and Proteins: A Practical Approach*, (Oxford University Press., 1999).
62. Gasteiger, E. et al. Protein Identification and Analysis Tools on the ExPASy Server; John M. Walker (ed): *The Proteomics Protocols Handbook*, Humana Press, 571-607 (2005).
63. Evans, G.P., R. F. . CHOOCH: a program for deriving anomalous-scattering factors from X-ray fluorescence spectra. *J. Appl. Cryst.* **34**, , 82-86. (2001).
64. Blu-Ice.
http://images.google.com.br/imgres?imgurl=http://smb.slac.stanford.edu/facilities/software/blu-ice/images/scan_tab/scan_pic.jpg&imgrefurl=http://smb.slac.stanford.edu/facilities/software/blu-ice/scan_tab.html&usg=__Zm61AkHPpnZ_CumOnOOwZXbdkNs=&h=840&w=1116&sz=409&hl=pt-BR&start=29&um=1&tbnid=lefhlGI0tgkeuM:&tbnh=113&tbnw=150&prev=/image%3Fq%3DBlu-Ice%26ndsp%3D18%26hl%3Dpt-BR%26sa%3DN%26start%3D18%26um%3D1
65. Otwinowski, Z. & Minor, W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. in *Methods in Enzymology* Vol. 276 307-326 (C.W. Carter, Jr. & R. M. Sweet, Eds., (New York), 1997).
66. Leslie, A.G.W. Joint CCP4 + ESF-EAMCB *Newsletter on Protein Crystallography* **26**(1992).
67. *The CCP4 suite: programs for protein crystallography*, 760-763 (Collaborative Computational Project, 1994).
68. McCoy A. J., G.-K.R.W., Adams P. D, Winn M. D. , Storoni L.C. and Read R.J. . . Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658-674 (2007).
69. Schneider, T.R. & Sheldrick, G.M. Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr* **58**, 1772-9 (2002).
70. La Fortelle, E.D., Bricogne, G. . Maximum-Likelihood Heavy-Atom Parameter Refinement for Multiple Isomorphous Replacement and Multiwavelength Anomalous Diffraction Methods *Methods in Enzymology* **276**, 472-494 (1997).
71. Perrakis, A., Morris, R. & Lamzin, V.S. Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* **6**, 458-63 (1999).
72. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-55 (1997).
73. Brunger, A.T. et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**, 905-21 (1998).
74. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32 (2004).
75. Painter, J. & Merritt, E.A. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* **62**, 439-50 (2006).
76. Laskowski R. A., M.M.W., Moss D. S. and Thornton J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* **26**, 283-291 (1993).
77. James, P., Halladay, J. & Craig, E.A. Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144**, 1425-36 (1996).

78. Gietz, R.D. Yeast two-hybrid system screening. *Methods Mol Biol* **313**, 345-71 (2006).
79. Gietz, R.D. & Woods, R.A. Yeast transformation by the LiAc/SS Carrier DNA/PEG method. *Methods Mol Biol* **313**, 107-20 (2006).
80. Uetz, P. et al. A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* **403**, 623-7 (2000).
81. Alegria, M.C. et al. New protein-protein interactions identified for the regulatory and structural components and substrates of the type III Secretion system of the phytopathogen *Xanthomonas axonopodis* Pathovar citri. *J Bacteriol* **186**, 6186-97 (2004).
82. Hall, R.A. Studying protein-protein interactions via blot overlay or Far Western blot. *Methods Mol Biol* **261**, 167-74 (2004).
83. Guzzo, C.R. et al. Expression, purification, crystallization and preliminary X-ray analysis of YaeQ (XAC2396) from *Xanthomonas axonopodis* pv. citri. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **61**, 493-5 (2005).
84. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-93 (1995).
85. Vranken, W.F. et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687-96 (2005).
86. Sattler M., S.J., Griesinger C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients *Progress in Nuclear Magnetic Resonance Spectroscopy* **34**, 93-158 (1999).
87. Pace, C.N. & Sholtz, J.M. Determining the conformational stability of a protein. in *Protein Structure, A Practical Approach* (ed. Creighton, T.E.) 299-321 (IRL Press, Oxford, 1997).
88. Outten, F.W., Wood, M.J., Munoz, F.M. & Storz, G. The SufE protein and the SufBCD complex enhance SufS cysteine desulfurase activity as part of a sulfur transfer pathway for Fe-S cluster assembly in *Escherichia coli*. *J Biol Chem* **278**, 45713-9 (2003).
89. Ye, H., Abdel-Ghany, S.E., Anderson, T.D., Pilon-Smits, E.A. & Pilon, M. CpSufE activates the cysteine desulfurase CpNifS for chloroplastic Fe-S cluster formation. *J Biol Chem* **281**, 8958-69 (2006).
90. Siegel, L.M. A Direct Microdetermination for Sulfide. *Anal Biochem* **11**, 126-32 (1965).
91. Wang, T. et al. The sufR gene (sll0088 in *Synechocystis* sp. strain PCC 6803) functions as a repressor of the sufBCDS operon in iron-sulfur cluster biogenesis in cyanobacteria. *J Bacteriol* **186**, 956-67 (2004).
92. Craig, L. et al. Type IV pilus structure by cryo-electron microscopy and crystallography: implications for pilus assembly and functions. *Mol Cell* **23**, 651-62 (2006).
93. Mattick, J.S. Type IV pili and twitching motility. *Annu Rev Microbiol* **56**, 289-314 (2002).
94. Craig, L., Pique, M.E. & Tainer, J.A. Type IV pilus structure and bacterial pathogenicity. *Nat Rev Microbiol* **2**, 363-78 (2004).
95. Han, X., Kennan, R.M., Parker, D., Davies, J.K. & Rood, J.I. Type IV fimbrial biogenesis is required for protease secretion and natural transformation in *Dichelobacter nodosus*. *J Bacteriol* **189**, 5022-33 (2007).
96. Craig, L. & Li, J. Type IV pili: paradoxes in form and function. *Curr Opin Struct Biol* **18**, 267-77 (2008).

97. Miller, R.M. et al. *Pseudomonas aeruginosa* twitching motility-mediated chemotaxis towards phospholipids and fatty acids: specificity and metabolic requirements. *J Bacteriol* **190**, 4038-49 (2008).
98. Pelicic, V. Type IV pili: e pluribus unum? *Mol Microbiol* **68**, 827-37 (2008).
99. Darzins, A. The pilG gene product, required for *Pseudomonas aeruginosa* pilus production and twitching motility, is homologous to the enteric, single-domain response regulator CheY. *J Bacteriol* **175**, 5934-44 (1993).
100. Darzins, A. Characterization of a *Pseudomonas aeruginosa* gene cluster involved in pilus biosynthesis and twitching motility: sequence similarity to the chemotaxis proteins of enterics and the gliding bacterium *Myxococcus xanthus*. *Mol Microbiol* **11**, 137-53 (1994).
101. Darzins, A. & Russell, M.A. Molecular genetic analysis of type-4 pilus biogenesis and twitching motility using *Pseudomonas aeruginosa* as a model system--a review. *Gene* **192**, 109-15 (1997).
102. Li, Y. et al. Type I and type IV pili of *Xylella fastidiosa* affect twitching motility, biofilm formation and cell-cell aggregation. *Microbiology* **153**, 719-26 (2007).
103. Ohno, S. *Evolution by gene duplication* (Springer-Verlag. ISBN 0-04-575015-7, 1970).
104. Varshavsky, A. The N-end rule pathway of protein degradation. *Genes Cells* **2**, 13-28 (1997).
105. Craig, L. et al. Type IV pilin structure and assembly: X-ray and EM analyses of *Vibrio cholerae* toxin-coregulated pilus and *Pseudomonas aeruginosa* PAK pilin. *Mol Cell* **11**, 1139-50 (2003).
106. Lory, S. & Strom, M.S. Structure-function relationship of type-IV prepilin peptidase of *Pseudomonas aeruginosa*--a review. *Gene* **192**, 117-21 (1997).
107. Chiang, P., Habash, M. & Burrows, L.L. Disparate subcellular localization patterns of *Pseudomonas aeruginosa* Type IV pilus ATPases involved in twitching motility. *J Bacteriol* **187**, 829-39 (2005).
108. Bitter, W., Koster, M., Latijnhouwers, M., de Cock, H. & Tommassen, J. Formation of oligomeric rings by XcpQ and PilQ, which are involved in protein transport across the outer membrane of *Pseudomonas aeruginosa*. *Mol Microbiol* **27**, 209-19 (1998).
109. Winther-Larsen, H.C. et al. A conserved set of pilin-like molecules controls type IV pilus dynamics and organelle-associated functions in *Neisseria gonorrhoeae*. *Mol Microbiol* **56**, 903-17 (2005).
110. Merz, A.J., So, M. & Sheetz, M.P. Pilus retraction powers bacterial twitching motility. *Nature* **407**, 98-102 (2000).
111. Kazmierczak, B.I., Lebron, M.B. & Murray, T.S. Analysis of FimX, a phosphodiesterase that governs twitching motility in *Pseudomonas aeruginosa*. *Mol Microbiol* **60**, 1026-43 (2006).
112. page, K.K.E.o.G.a.G.h. http://www.genome.ad.jp/dbget-bin/show_pathway?xac03090+XAC3239. (2008).
113. Carbonnelle, E., Helaine, S., Prouvensier, L., Nassif, X. & Pelicic, V. Type IV pilus biogenesis in *Neisseria meningitidis*: PilW is involved in a step occurring after pilus assembly, essential for fibre stability and function. *Mol Microbiol* **55**, 54-64 (2005).
114. McCarthy, Y. et al. The role of PilZ domain proteins in the virulence of *Xanthomonas campestris* pv. *campestris*. *Mol Plant Pathol* **9**, 819-24 (2008).
115. Amikam, D. & Galperin, M.Y. PilZ domain is part of the bacterial c-di-GMP binding protein. *Bioinformatics* **22**, 3-6 (2006).

116. Benach, J. et al. The structural basis of cyclic diguanylate signal transduction by PilZ domains. *EMBO J* **26**, 5153-66 (2007).
117. Pratt, J.T., Tamayo, R., Tischler, A.D. & Camilli, A. PilZ domain proteins bind cyclic diguanylate and regulate diverse processes in *Vibrio cholerae*. *J Biol Chem* **282**, 12860-70 (2007).
118. Ramelot, T.A. et al. NMR structure and binding studies confirm that PA4608 from *Pseudomonas aeruginosa* is a PilZ domain and a c-di-GMP binding protein. *Proteins* **66**, 266-71 (2007).
119. Merighi, M., Lee, V.T., Hyodo, M., Hayakawa, Y. & Lory, S. The second messenger bis-(3'-5')-cyclic-GMP and its PilZ domain-containing receptor Alg44 are required for alginate biosynthesis in *Pseudomonas aeruginosa*. *Mol Microbiol* **65**, 876-95 (2007).
120. Finn, R.D. et al. The Pfam protein families database. *Nucleic Acids Res* **36**, D281-8 (2008).
121. Crossman, L. & Dow, J.M. Biofilm formation and dispersal in *Xanthomonas campestris*. *Microbes Infect* **6**, 623-9 (2004).
122. Cotter, P.A. & Stibitz, S. c-di-GMP-mediated regulation of virulence and biofilm formation. *Curr Opin Microbiol* **10**, 17-23 (2007).
123. Simm, R., Lusch, A., Kader, A., Andersson, M. & Romling, U. Role of EAL-containing proteins in multicellular behavior of *Salmonella enterica* serovar Typhimurium. *J Bacteriol* **189**, 3613-23 (2007).
124. Simm, R., Morr, M., Kader, A., Nimtz, M. & Romling, U. GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Mol Microbiol* **53**, 1123-34 (2004).
125. Yan, H. & Humes, E. Convenient synthesis of 3',5'-cyclic diguanylic acid (cdiGMP). *Nucleic Acids Symp Ser (Oxf)*, 5-6 (2006).
126. Chan, C. et al. Structural basis of activity and allosteric control of diguanylate cyclase. *Proc Natl Acad Sci U S A* **101**, 17084-9 (2004).
127. Romling, U. & Amikam, D. Cyclic di-GMP as a second messenger. *Curr Opin Microbiol* **9**, 218-28 (2006).
128. Hengge, R. Principles of c-di-GMP signalling in bacteria. *Nat Rev Microbiol* **7**, 263-73 (2009).
129. Jenal, U. & Malone, J. Mechanisms of cyclic-di-GMP signaling in bacteria. *Annu Rev Genet* **40**, 385-407 (2006).
130. Romling, U., Gomelsky, M. & Galperin, M.Y. C-di-GMP: the dawning of a novel bacterial signalling system. *Mol Microbiol* **57**, 629-39 (2005).
131. Aldridge, P., Paul, R., Goymer, P., Rainey, P. & Jenal, U. Role of the GGDEF regulator PleD in polar development of *Caulobacter crescentus*. *Mol Microbiol* **47**, 1695-708 (2003).
132. Bobrov, A.G., Kirillina, O. & Perry, R.D. The phosphodiesterase activity of the HmsP EAL domain is required for negative regulation of biofilm formation in *Yersinia pestis*. *FEMS Microbiol Lett* **247**, 123-30 (2005).
133. Sudarsan, N. et al. Riboswitches in eubacteria sense the second messenger cyclic di-GMP. *Science* **321**, 411-3 (2008).
134. ncbi, p.d. bis(3',5')-cyclic diguanylic acid - Compound Summary (CID 6323195). (http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=6323195&loc=ec_r_cs).
135. Yeates, T.O. Detecting and overcoming crystal twinning (web page: <http://nihserver.mbi.ucla.edu/Twinning/>). in *Methods Enzymol* 1997/01/01 edn Vol. 276 344-58 (1997).

136. Guzzo, C.R. & Farah, C.S. Expression, crystallization and preliminary crystallographic analysis of PilZ(XAC1133) from *Xanthomonas axonopodis* pv. citri. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **65**, 304-6 (2009).
137. McPhillips T. M., M.S.E., Chiu H.-J., Cohen A. E., Deacon A. M., Ellis P. J., Garman E., Gonzalez A., Sauter N. K., Phizackerley R. P., Soltis S. M. and Kuhn P. Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. *J. Synchrotron Rad.* **9**, 401-406 (2001).
138. M., U.S.G. Advances in direct methods for protein crystallography". *Curr. Op. Struct. Biol.* 1998, 643-648 (1998).
139. Krissinel, E. PISA, Protein Interfaces, Surfaces and Assemblies. http://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver edn (EMBL-EBI, 2009).
140. Li, T.N., Chin, K.H., Liu, J.H., Wang, A.H. & Chou, S.H. XC1028 from *Xanthomonas campestris* adopts a PilZ domain-like structure without a c-di-GMP switch. *Proteins* (2008).
141. Harshey, R.M. Bacterial motility on a surface: many ways to a common goal. *Annu Rev Microbiol* **57**, 249-73 (2003).
142. Brown, II & Hase, C.C. Flagellum-independent surface migration of *Vibrio cholerae* and *Escherichia coli*. *J Bacteriol* **183**, 3784-90 (2001).
143. Henrichsen, J. Twitching motility. *Annu Rev Microbiol* **37**, 81-93 (1983).
144. Ottow, J.C. Ecology, physiology, and genetics of fimbriae and pili. *Annu Rev Microbiol* **29**, 79-108 (1975).
145. Shrout, J.D. et al. The impact of quorum sensing and swarming motility on *Pseudomonas aeruginosa* biofilm formation is nutritionally conditional. *Mol Microbiol* **62**, 1264-77 (2006).
146. Murray, T.S. & Kazmierczak, B.I. *Pseudomonas aeruginosa* exhibits sliding motility in the absence of type IV pili and flagella. *J Bacteriol* **190**, 2700-8 (2008).
147. Henrichsen, J. Bacterial surface translocation: a survey and a classification. *Bacteriol Rev* **36**, 478-503 (1972).
148. Astua-Monge, G. et al. Resistance of tomato and pepper to T3 strains of *Xanthomonas campestris* pv. vesicatoria is specified by a plant-inducible avirulence gene. *Mol Plant Microbe Interact* **13**, 911-21 (2000).
149. Schulte, R. & Bonas, U. Expression of the *Xanthomonas campestris* pv. vesicatoria hrp gene cluster, which determines pathogenicity and hypersensitivity on pepper and tomato, is plant inducible. *J Bacteriol* **174**, 815-23 (1992).
150. Tung, S.Y. & Kuo, T.T. Requirement for phosphoglucose isomerase of *Xanthomonas campestris* in pathogenesis of citrus canker. *Appl Environ Microbiol* **65**, 5564-70 (1999).
151. Wengelnik, K. & Bonas, U. HrpXv, an AraC-type regulator, activates expression of five of the six loci in the hrp cluster of *Xanthomonas campestris* pv. vesicatoria. *J Bacteriol* **178**, 3462-9 (1996).
152. Leduc, J.L. & Roberts, G.P. Cyclic di-GMP Allosterically Inhibits the CRP-like Protein (Clp) of *Xanthomonas axonopodis* pv. citri. *J Bacteriol* (2009).
153. Hickman, J.W. & Harwood, C.S. Identification of FleQ from *Pseudomonas aeruginosa* as a c-di-GMP-responsive transcription factor. *Mol Microbiol* **69**, 376-89 (2008).
154. Lee, V.T. et al. A cyclic-di-GMP receptor required for bacterial exopolysaccharide production. *Mol Microbiol* **65**, 1474-84 (2007).

155. 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. *J Bacteriol* **180**, 4416-25 (1998).
156. Kulasakara, H. et al. Analysis of *Pseudomonas aeruginosa* diguanylate cyclases and phosphodiesterases reveals a role for bis-(3'-5')-cyclic-GMP in virulence. *Proc Natl Acad Sci U S A* **103**, 2839-44 (2006).
157. Christen, M., Christen, B., Folcher, M., Schauerte, A. & Jenal, U. Identification and characterization of a cyclic di-GMP-specific phosphodiesterase and its allosteric control by GTP. *J Biol Chem* **280**, 30829-37 (2005).
158. Kumar, M. & Chatterji, D. Cyclic di-GMP: a second messenger required for long-term survival, but not for biofilm formation, in *Mycobacterium smegmatis*. *Microbiology* **154**, 2942-55 (2008).
159. Ferreira, R.B., Antunes, L.C., Greenberg, E.P. & McCarter, L.L. *Vibrio parahaemolyticus* ScrC modulates cyclic dimeric GMP regulation of gene expression relevant to growth on surfaces. *J Bacteriol* **190**, 851-60 (2008).
160. Navarro, M.V., De, N., Bae, N., Wang, Q. & Sondermann, H. Structural analysis of the GGDEF-EAL domain-containing c-di-GMP receptor FimX. *Structure* **17**, 1104-16 (2009).
161. Ross, P. et al. The cyclic diguanylic acid regulatory system of cellulose synthesis in *Acetobacter xylinum*. Chemical synthesis and biological activity of cyclic nucleotide dimer, trimer, and phosphothioate derivatives. *J Biol Chem* **265**, 18933-43 (1990).
162. Romling, U. Rationalizing the evolution of EAL domain-based cyclic di-GMP-specific phosphodiesterases. *J Bacteriol* **191**, 4697-700 (2009).
163. Christen, M. et al. DgrA is a member of a new family of cyclic diguanosine monophosphate receptors and controls flagellar motor function in *Caulobacter crescentus*. *Proc Natl Acad Sci U S A* **104**, 4112-7 (2007).
164. Larkin, M.A. et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-8 (2007).
165. Crooks, G.E., Hon, G., Chandonia, J.M. & Brenner, S.E. WebLogo: a sequence logo generator. *Genome Res* **14**, 1188-90 (2004).
166. Schneider, T.D. & Stephens, R.M. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* **18**, 6097-100 (1990).
167. Podust, L.M., Ioanoviciu, A. & Ortiz de Montellano, P.R. 2.3 Å X-ray structure of the heme-bound GAF domain of sensory histidine kinase DosT of *Mycobacterium tuberculosis*. *Biochemistry* **47**, 12523-31 (2008).
168. Romling, U. Cyclic Di-GMP (c-Di-GMP) goes into host cells--c-Di-GMP signaling in the obligate intracellular pathogen *Anaplasma phagocytophilum*. *J Bacteriol* **191**, 683-6 (2009).
169. Ryjenkov, D.A., Tarutina, M., Moskvina, O.V. & Gomelsky, M. Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: insights into biochemistry of the GGDEF protein domain. *J Bacteriol* **187**, 1792-8 (2005).
170. De, N., Navarro, M.V., Raghavan, R.V. & Sondermann, H. Determinants for the Activation and Autoinhibition of the Diguanylate Cyclase Response Regulator WspR. *J Mol Biol* (2009).
171. Mendez-Ortiz, M.M. & Membrillo-Hernandez, J. [Proteins with GGDEF and EAL domains: their role in bacterial metabolism]. *Rev Latinoam Microbiol* **47**, 130-9 (2005).

172. Tamayo, R., Schild, S., Pratt, J.T. & Camilli, A. Role of cyclic Di-GMP during el tor biotype *Vibrio cholerae* infection: characterization of the in vivo-induced cyclic Di-GMP phosphodiesterase CdpA. *Infect Immun* **76**, 1617-27 (2008).
173. Wassmann, P. et al. Structure of BeF₃⁻-modified response regulator PleD: implications for diguanylate cyclase activation, catalysis, and feedback inhibition. *Structure* **15**, 915-27 (2007).
174. Barends, T.R. et al. Structure and mechanism of a bacterial light-regulated cyclic nucleotide phosphodiesterase. *Nature* **459**, 1015-8 (2009).
175. Rao, F., Yang, Y., Qi, Y. & Liang, Z.X. Catalytic mechanism of cyclic di-GMP-specific phosphodiesterase: a study of the EAL domain-containing RocR from *Pseudomonas aeruginosa*. *J Bacteriol* **190**, 3622-31 (2008).
176. Galperin, M.Y. A census of membrane-bound and intracellular signal transduction proteins in bacteria: bacterial IQ, extroverts and introverts. *BMC Microbiol* **5**, 35 (2005).
177. Ryan, R.P., Fouhy, Y., Lucey, J.F. & Dow, J.M. Cyclic di-GMP signaling in bacteria: recent advances and new puzzles. *J Bacteriol* **188**, 8327-34 (2006).
178. Ryan, R.P. et al. Physical interactions between HD-GYP and GGDEF domain proteins mediate virulence-related signaling in *Xanthomonas campestris*. (2010).
179. Ryan, R.P. et al. HD-GYP domain proteins regulate biofilm formation and virulence in *Pseudomonas aeruginosa*. *Environ Microbiol* **11**, 1126-36 (2009).
180. Kelley, L.A. & Sternberg, M.J. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* **4**, 363-71 (2009).
181. Chiang, P. et al. Functional role of conserved residues in the characteristic secretion NTPase motifs of the *Pseudomonas aeruginosa* type IV pilus motor proteins PilB, PilT and PilU. *Microbiology* **154**, 114-26 (2008).
182. Turner, L.R., Lara, J.C., Nunn, D.N. & Lory, S. Mutations in the consensus ATP-binding sites of XcpR and PilB eliminate extracellular protein secretion and pilus biogenesis in *Pseudomonas aeruginosa*. *J Bacteriol* **175**, 4962-9 (1993).
183. Nunn, D., Bergman, S. & Lory, S. Products of three accessory genes, pilB, pilC, and pilD, are required for biogenesis of *Pseudomonas aeruginosa* pili. *J Bacteriol* **172**, 2911-9 (1990).
184. Jakovljevic, V., Leonardy, S., Hoppert, M. & Sogaard-Andersen, L. PilB and PilT are ATPases acting antagonistically in type IV pilus function in *Myxococcus xanthus*. *J Bacteriol* **190**, 2411-21 (2008).
185. Robien, M.A., Krumm, B.E., Sandkvist, M. & Hol, W.G. Crystal structure of the extracellular protein secretion NTPase EpsE of *Vibrio cholerae*. *J Mol Biol* **333**, 657-74 (2003).
186. Planet, P.J., Kachlany, S.C., DeSalle, R. & Figurski, D.H. Phylogeny of genes for secretion NTPases: identification of the widespread tadA subfamily and development of a diagnostic key for gene classification. *Proc Natl Acad Sci U S A* **98**, 2503-8 (2001).
187. Crowther, L.J., Yamagata, A., Craig, L., Tainer, J.A. & Donnenberg, M.S. The ATPase activity of BfpD is greatly enhanced by zinc and allosteric interactions with other Bfp proteins. *J Biol Chem* **280**, 24839-48 (2005).
188. Merz, A.J. & So, M. Interactions of pathogenic neisseriae with epithelial cell membranes. *Annu Rev Cell Dev Biol* **16**, 423-57 (2000).
189. NCBI. National Center for Biotechnology Information U.S. National Library of Medicine - <http://www.ncbi.nlm.nih.gov/>.
190. Wong, K.R., Hughes, C. & Koronakis, V. A gene, yaeQ, that suppresses reduced operon expression caused by mutations in the transcription elongation gene rfaH

- in *Escherichia coli* and *Salmonella typhimurium*. *Mol Gen Genet* **257**, 693-6 (1998).
191. Vicari, D. & Artsimovitch, I. Virulence regulators RfaH and YaeQ do not operate in the same pathway. *Mol Genet Genomics* **272**, 489-96 (2004).
  192. Schulze-Gahmen, U., Pelaschier, J., Yokota, H., Kim, R. & Kim, S.H. Crystal structure of a hypothetical protein, TM841 of *Thermotoga maritima*, reveals its function as a fatty acid-binding protein. *Proteins* **50**, 526-30 (2003).
  193. Christendat, D. et al. The crystal structure of hypothetical protein MTH1491 from *Methanobacterium thermoautotrophicum*. *Protein Sci* **11**, 1409-14 (2002).
  194. Brakoulias, A. & Jackson, R.M. Towards a structural classification of phosphate binding sites in protein-nucleotide complexes: an automated all-against-all structural comparison using geometric matching. *Proteins* **56**, 250-60 (2004).
  195. Guzzo, C.R., Nagem, R.A., Barbosa, J.A. & Farah, C.S. Structure of *Xanthomonas axonopodis* pv. citri YaeQ reveals a new compact protein fold built around a variation of the PD-(D/E)XK nuclease motif. *Proteins* **69**, 644-51 (2007).
  196. Ban, C. & Yang, W. Structural basis for MutH activation in *E. coli* mismatch repair and relationship of MutH to restriction endonucleases. *EMBO J* **17**, 1526-34 (1998).
  197. Nishino, T., Komori, K., Tsuchiya, D., Ishino, Y. & Morikawa, K. Crystal structure of the archaeal holliday junction resolvase Hjc and implications for DNA recognition. *Structure* **9**, 197-204 (2001).
  198. Singleton, M.R., Dillingham, M.S., Gaudier, M., Kowalczykowski, S.C. & Wigley, D.B. Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature* **432**, 187-93 (2004).
  199. Kosinski, J., Feder, M. & Bujnicki, J.M. The PD-(D/E)XK superfamily revisited: identification of new members among proteins involved in DNA metabolism and functional predictions for domains of (hitherto) unknown function. *BMC Bioinformatics* **6**, 172 (2005).
  200. Tamulaitiene, G. et al. The crystal structure of the rare-cutting restriction enzyme SdaI reveals unexpected domain architecture. *Structure* **14**, 1389-400 (2006).
  201. Nishino, T., Komori, K., Ishino, Y. & Morikawa, K. X-ray and biochemical anatomy of an archaeal XPF/Rad1/Mus81 family nuclease: similarity between its endonuclease domain and restriction enzymes. *Structure* **11**, 445-57 (2003).
  202. Newman, M., Strzelecka, T., Dorner, L.F., Schildkraut, I. & Aggarwal, A.K. Structure of restriction endonuclease bamHI phased at 1.95 Å resolution by MAD analysis. *Structure* **2**, 439-52 (1994).
  203. Newman, M., Strzelecka, T., Dorner, L.F., Schildkraut, I. & Aggarwal, A.K. Structure of restriction endonuclease BamHI and its relationship to EcoRI. *Nature* **368**, 660-4 (1994).
  204. Viadiu, H. & Aggarwal, A.K. The role of metals in catalysis by the restriction endonuclease BamHI. *Nat Struct Biol* **5**, 910-6 (1998).
  205. Wah, D.A., Hirsch, J.A., Dorner, L.F., Schildkraut, I. & Aggarwal, A.K. Structure of the multimodular endonuclease FokI bound to DNA. *Nature* **388**, 97-100 (1997).
  206. Alegria, M.C. et al. Conditions affecting production of functional muscle recombinant alpha-tropomyosin in *Saccharomyces cerevisiae*. *Protein Expr Purif* **30**, 105-11 (2003).
  207. Rapoport, G. & Dedonder, R. [Specificity of the levansucrase activity of *Bacillus subtilis*: clarification]. *Bull Soc Chim Biol (Paris)* **48**, 1311-22 (1966).

208. Steinmetz, M., Le Coq, D., Djemia, H.B. & Gay, P. [Genetic analysis of *sacB*, the structural gene of a secreted enzyme, levansucrase of *Bacillus subtilis* Marburg]. *Mol Gen Genet* **191**, 138-44 (1983).
209. Ried, J.L. & Collmer, A. An *nptI-sacB-sacR* cartridge for constructing directed, unmarked mutations in gram-negative bacteria by marker exchange- eviction mutagenesis. *Gene* **57**, 239-46 (1987).
210. Pelicic, V., Reyrat, J.M. & Gicquel, B. Expression of the *Bacillus subtilis sacB* gene confers sucrose sensitivity on mycobacteria. *J Bacteriol* **178**, 1197-9 (1996).
211. Campbell, E.A. et al. Structural mechanism for rifampicin inhibition of bacterial rna polymerase. *Cell* **104**, 901-12 (2001).
212. Newman, M. et al. Structure of an XPF endonuclease with and without DNA suggests a model for substrate recognition. *EMBO J* **24**, 895-905 (2005).
213. Iyer, R.R., Pluciennik, A., Burdett, V. & Modrich, P.L. DNA mismatch repair: functions and mechanisms. *Chem Rev* **106**, 302-23 (2006).
214. Junop, M.S., Yang, W., Funchain, P., Clendenin, W. & Miller, J.H. In vitro and in vivo studies of MutS, MutL and MutH mutants: correlation of mismatch repair and DNA recombination. *DNA Repair (Amst)* **2**, 387-405 (2003).
215. Lee, J.Y. et al. MutH complexed with hemi- and unmethylated DNAs: coupling base recognition and DNA cleavage. *Mol Cell* **20**, 155-66 (2005).
216. Marti, T.M., Kunz, C. & Fleck, O. DNA mismatch repair and mutation avoidance pathways. *J Cell Physiol* **191**, 28-41 (2002).
217. Pukkila, P.J., Peterson, J., Herman, G., Modrich, P. & Meselson, M. Effects of high levels of DNA adenine methylation on methyl-directed mismatch repair in *Escherichia coli*. *Genetics* **104**, 571-82 (1983).
218. Geier, G.E. & Modrich, P. Recognition sequence of the *dam* methylase of *Escherichia coli* K12 and mode of cleavage of Dpn I endonuclease. *J Biol Chem* **254**, 1408-13 (1979).
219. Marinus, M.G. Location of DNA methylation genes on the *Escherichia coli* K-12 genetic map. *Mol Gen Genet* **127**, 47-55 (1973).
220. May, M.S. & Hattaman, S. Deoxyribonucleic acid-cytosine methylation by host- and plasmid-controlled enzymes. *J Bacteriol* **122**, 129-38 (1975).
221. Zell, R. & Fritz, H.J. DNA mismatch-repair in *Escherichia coli* counteracting the hydrolytic deamination of 5-methyl-cytosine residues. *EMBO J* **6**, 1809-15 (1987).
222. Hendrich, B., Hardeland, U., Ng, H.H., Jiricny, J. & Bird, A. The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature* **401**, 301-4 (1999).
223. Bunting, K.A. et al. Crystal structure of the *Escherichia coli* dcm very-short-patch DNA repair endonuclease bound to its reaction product-site in a DNA superhelix. *Nucleic Acids Res* **31**, 1633-9 (2003).
224. Drotschmann, K., Aronshtam, A., Fritz, H.J. & Marinus, M.G. The *Escherichia coli* MutL protein stimulates binding of Vsr and MutS to heteroduplex DNA. *Nucleic Acids Res* **26**, 948-53 (1998).
225. Macintyre, G., Doiron, K.M. & Cupples, C.G. The Vsr endonuclease of *Escherichia coli*: an efficient DNA repair enzyme and a potent mutagen. *J Bacteriol* **179**, 6048-52 (1997).
226. Kulakauskas, S., Barsomian, J.M., Lubys, A., Roberts, R.J. & Wilson, G.G. Organization and sequence of the *HpaII* restriction-modification system and adjacent genes. *Gene* **142**, 9-15 (1994).

227. Tsutakawa, S.E., Jingami, H. & Morikawa, K. Recognition of a TG mismatch: the crystal structure of very short patch repair endonuclease in complex with a DNA duplex. *Cell* **99**, 615-23 (1999).
228. Tsutakawa, S.E. et al. Crystallographic and functional studies of very short patch repair endonuclease. *Mol Cell* **3**, 621-8 (1999).
229. Dalhus, B., Laerdahl, J.K., Backe, P.H. & Bjoras, M. DNA base repair-- recognition and initiation of catalysis. *FEMS Microbiol Rev* **33**, 1044-78 (2009).
230. Kiley, P.J. & Beinert, H. The role of Fe-S proteins in sensing and regulation in bacteria. *Curr Opin Microbiol* **6**, 181-5 (2003).
231. Zheng, L. & Dean, D.R. Catalytic formation of a nitrogenase iron-sulfur cluster. *J Biol Chem* **269**, 18723-6 (1994).
232. Johnson, D.C., Dean, D.R., Smith, A.D. & Johnson, M.K. Structure, function, and formation of biological iron-sulfur clusters. *Annu Rev Biochem* **74**, 247-81 (2005).
233. Johnson, D.C., Dos Santos, P.C. & Dean, D.R. NifU and NifS are required for the maturation of nitrogenase and cannot replace the function of isc-gene products in *Azotobacter vinelandii*. *Biochem Soc Trans* **33**, 90-3 (2005).
234. Lill, R. Function and biogenesis of iron-sulphur proteins. *Nature* **460**, 831-8 (2009).
235. Booker, S.J., Cicchillo, R.M. & Grove, T.L. Self-sacrifice in radical S-adenosylmethionine proteins. *Curr Opin Chem Biol* **11**, 543-52 (2007).
236. Imlay, J.A. Cellular defenses against superoxide and hydrogen peroxide. *Annu Rev Biochem* **77**, 755-76 (2008).
237. Rudolf, J., Makrantonis, V., Ingledew, W.J., Stark, M.J. & White, M.F. The DNA repair helicases XPD and FancJ have essential iron-sulfur domains. *Mol Cell* **23**, 801-8 (2006).
238. Fontecave, M. & Ollagnier-de-Choudens, S. Iron-sulfur cluster biosynthesis in bacteria: Mechanisms of cluster assembly and transfer. *Arch Biochem Biophys* **474**, 226-37 (2008).
239. Meyer, J. Iron-sulfur protein folds, iron-sulfur chemistry, and evolution. *J Biol Inorg Chem* **13**, 157-70 (2008).
240. Tokumoto, U., Kitamura, S., Fukuyama, K. & Takahashi, Y. Interchangeability and distinct properties of bacterial Fe-S cluster assembly systems: functional replacement of the isc and suf operons in *Escherichia coli* with the nifSU-like operon from *Helicobacter pylori*. *J Biochem* **136**, 199-209 (2004).
241. Outten, F.W., Djaman, O. & Storz, G. A suf operon requirement for Fe-S cluster assembly during iron starvation in *Escherichia coli*. *Mol Microbiol* **52**, 861-72 (2004).
242. Huet, G., Daffe, M. & Saves, I. Identification of the *Mycobacterium tuberculosis* SUF machinery as the exclusive mycobacterial system of [Fe-S] cluster assembly: evidence for its implication in the pathogen's survival. *J Bacteriol* **187**, 6137-46 (2005).
243. Goldsmith-Fischman, S. et al. The SufE sulfur-acceptor protein contains a conserved core structure that mediates interdomain interactions in a variety of redox protein complexes. *J Mol Biol* **344**, 549-65 (2004).
244. Ollagnier-de-Choudens, S. et al. Mechanistic studies of the SufS-SufE cysteine desulfurase: evidence for sulfur transfer from SufS to SufE. *FEBS Lett* **555**, 263-7 (2003).
245. Ramelot, T.A. et al. Solution structure of *Vibrio cholerae* protein VC0424: a variation of the ferredoxin-like fold. *Protein Sci* **12**, 1556-61 (2003).

246. Balasubramanian, R., Shen, G., Bryant, D.A. & Golbeck, J.H. Regulatory roles for IscA and SufA in iron homeostasis and redox stress responses in the cyanobacterium *Synechococcus* sp. strain PCC 7002. *J Bacteriol* **188**, 3182-91 (2006).
247. Sendra, M., Ollagnier de Choudens, S., Lascoux, D., Sanakis, Y. & Fontecave, M. The SUF iron-sulfur cluster biosynthetic machinery: sulfur transfer from the SUFS-SUFE complex to SUFA. *FEBS Lett* **581**, 1362-8 (2007).
248. Layer, G. et al. SufE transfers sulfur from SufS to SufB for iron-sulfur cluster assembly. *J Biol Chem* **282**, 13342-50 (2007).
249. Unciuleac, M.C. et al. In vitro activation of apo-aconitase using a [4Fe-4S] cluster-loaded form of the IscU [Fe-S] cluster scaffolding protein. *Biochemistry* **46**, 6812-21 (2007).
250. Chandramouli, K. et al. Formation and properties of [4Fe-4S] clusters on the IscU scaffold protein. *Biochemistry* **46**, 6804-11 (2007).
251. Layer, G., Ollagnier-de Choudens, S., Sanakis, Y. & Fontecave, M. Iron-sulfur cluster biosynthesis: characterization of *Escherichia coli* CYaY as an iron donor for the assembly of [2Fe-2S] clusters in the scaffold IscU. *J Biol Chem* **281**, 16256-63 (2006).
252. Ayala-Castro, C., Saini, A. & Outten, F.W. Fe-S cluster assembly pathways in bacteria. *Microbiol Mol Biol Rev* **72**, 110-25, table of contents (2008).
253. Mettert, E.L., Outten, F.W., Wanta, B. & Kiley, P.J. The impact of O(2) on the Fe-S cluster biogenesis requirements of *Escherichia coli* FNR. *J Mol Biol* **384**, 798-811 (2008).
254. Butland, G. et al. eSGA: *E. coli* synthetic genetic array analysis. *Nat Methods* **5**, 789-95 (2008).
255. Vinella, D., Brochier-Armanet, C., Loiseau, L., Talla, E. & Barras, F. Iron-sulfur (Fe/S) protein biogenesis: phylogenomic and genetic studies of A-type carriers. *PLoS Genet* **5**, e1000497 (2009).
256. Killmann, H., Herrmann, C., Torun, A., Jung, G. & Braun, V. TonB of *Escherichia coli* activates FhuA through interaction with the beta-barrel. *Microbiology* **148**, 3497-509 (2002).
257. Beck, B.J. & Downs, D.M. The apbE gene encodes a lipoprotein involved in thiamine synthesis in *Salmonella typhimurium*. *J Bacteriol* **180**, 885-91 (1998).
258. Skovran, E. & Downs, D.M. Metabolic defects caused by mutations in the isc gene cluster in *Salmonella enterica* serovar typhimurium: implications for thiamine synthesis. *J Bacteriol* **182**, 3896-903 (2000).
259. Liu, G. et al. High-quality homology models derived from NMR and X-ray structures of *E. coli* proteins YgdK and Suf E suggest that all members of the YgdK/Suf E protein family are enhancers of cysteine desulfurases. *Protein Sci* **14**, 1597-608 (2005).
260. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760-3 (1994).
261. DeLano, W.L. The PyMOL molecular graphics system. <http://www.pymol.org>. (2002).
262. Santos, J.M., Freire, P., Vicente, M. & Arraiano, C.M. The stationary-phase morphogene bolA from *Escherichia coli* is induced by stress during early stages of growth. *Mol Microbiol* **32**, 789-98 (1999).
263. Cecilian, F. et al. Cloning, overexpression, and purification of *Escherichia coli* quinolinate synthetase. *Protein Expr Purif* **18**, 64-70 (2000).

264. Hagervall, T.G., Edmonds, C.G., McCloskey, J.A. & Bjork, G.R. Transfer RNA(5-methylaminomethyl-2-thiouridine)-methyltransferase from *Escherichia coli* K-12 has two enzymatic activities. *J Biol Chem* **262**, 8488-95 (1987).
265. Nocek, B., Maltseva, N., Stam, J., Anderson, W., Joachimiak, A. Crystal structure of cysteine desulfuration protein SufE from *Salmonella typhimurium* LT2. *To be Published*.
266. ACD/ChemSketch. ACD/ChemSketch Freeware version 12.01, www.acdlabs.com. Advanced Chemistry Development, Inc (2009).
267. Holm, L., Kaariainen, S., Rosenstrom, P. & Schenkel, A. Searching protein structure databases with DaliLite v.3. *Bioinformatics* **24**, 2780-1 (2008).
268. Fujiwara, K. et al. Crystal structure of lipoate-protein ligase A from *Escherichia coli*. Determination of the lipoic acid-binding site. *J Biol Chem* **280**, 33645-51 (2005).
269. Jeyakanthan, J., Kanaujia, S.P., Sekar, K., Agari, Y., Ebihara, A., Shinkai, A., Kuramitsu, S., Yokoyama, S. Crystal structure of Iron-sulfur cluster biosynthesis protein IscU (TTHA1736) from *thermus thermophilus* HB8. *To be Published*
270. Kornhaber, G.J., Swapna, G.V.T., Ramelot, T.A., Cort, J.R., Kennedy, M.A., Montelione, G.T. Solution structure of Iron-Sulfur cluster assembly protein IscU from *Bacillus subtilis*, with Zinc bound at the active site. Northeast Structural Genomics Consortium Target SR17. *To be Published*
271. Shimomura, Y., Wada, K., Fukuyama, K. & Takahashi, Y. The asymmetric trimeric architecture of [2Fe-2S] IscU: implications for its scaffolding during iron-sulfur cluster biosynthesis. *J Mol Biol* **383**, 133-43 (2008).
272. Kornhaber, G.J., Swapna, G.V.T., Ramelot, T.A., Cort, J.R., Aramini, J.M., Kennedy, M.A., Montelione, G.T. Solution NMR Structure of Zn-Ligated Fe-S Cluster Assembly Scaffold Protein SufU From *Bacillus subtilis*. *To be published*
273. Nakanishi, T., Tomizawa, T., Koshiba, S., Inoue, M., Kigawa, T., Yokoyama, S. Solution structure of Iron-sulfur cluster protein U (IscU). *To be Published*
274. Holm, L. Unification of protein families. *Curr Opin Struct Biol* **8**, 372-9 (1998).
275. Deckert, G. et al. The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature* **392**, 353-8 (1998).
276. Mihara, H., Kurihara, T., Yoshimura, T. & Esaki, N. Kinetic and mutational studies of three NifS homologs from *Escherichia coli*: mechanistic difference between L-cysteine desulfurase and L-selenocysteine lyase reactions. *J Biochem* **127**, 559-67 (2000).
277. Mihara, H. et al. A nifS-like gene, *csdB*, encodes an *Escherichia coli* counterpart of mammalian selenocysteine lyase. Gene cloning, purification, characterization and preliminary x-ray crystallographic studies. *J Biol Chem* **274**, 14768-72 (1999).
278. Mihara, H. & Esaki, N. Bacterial cysteine desulfurases: their function and mechanisms. *Appl Microbiol Biotechnol* **60**, 12-23 (2002).
279. Loiseau, L., Ollagnier-de-Choudens, S., Nachin, L., Fontecave, M. & Barras, F. Biogenesis of Fe-S cluster by the bacterial Suf system: SufS and SufE form a new type of cysteine desulfurase. *J Biol Chem* **278**, 38352-9 (2003).
280. Shen, G. et al. SufR coordinates two [4Fe-4S]²⁺, 1⁺ clusters and functions as a transcriptional repressor of the *sufBCDS* operon and an autoregulator of *sufR* in cyanobacteria. *J Biol Chem* **282**, 31909-19 (2007).
281. Shen, G. & Golbeck, J.H. In *Photosystem I: The Plastocyanin*

- (Cytochrome c6)/Ferredoxin (Flavodoxin) Oxidoreductase in Photosynthesis. (Golbeck, J., ed) Kluwer Academic Publishers, Amsterdam, 529–548 (2005).
282. Whitehead, N.A., Barnard, A.M., Slater, H., Simpson, N.J. & Salmond, G.P. Quorum-sensing in Gram-negative bacteria. *FEMS Microbiol Rev* **25**, 365-404 (2001).
  283. Waters, C.M. & Bassler, B.L. Quorum sensing: cell-to-cell communication in bacteria. *Annu Rev Cell Dev Biol* **21**, 319-46 (2005).
  284. Camilli, A. & Bassler, B.L. Bacterial small-molecule signaling pathways. *Science* **311**, 1113-6 (2006).
  285. Williams, P., Winzer, K., Chan, W.C. & Camara, M. Look who's talking: communication and quorum sensing in the bacterial world. *Philos Trans R Soc Lond B Biol Sci* **362**, 1119-34 (2007).
  286. Wilkinson, A., Danino, V., Wisniewski-Dye, F., Lithgow, J.K. & Downie, J.A. N-acyl-homoserine lactone inhibition of rhizobial growth is mediated by two quorum-sensing genes that regulate plasmid transfer. *J Bacteriol* **184**, 4510-9 (2002).
  287. Engebrecht, J., Nealson, K. & Silverman, M. Bacterial bioluminescence: isolation and genetic analysis of functions from *Vibrio fischeri*. *Cell* **32**, 773-81 (1983).
  288. Engebrecht, J. & Silverman, M. Identification of genes and gene products necessary for bacterial bioluminescence. *Proc Natl Acad Sci U S A* **81**, 4154-8 (1984).
  289. Engebrecht, J. & Silverman, M. Nucleotide sequence of the regulatory locus controlling expression of bacterial genes for bioluminescence. *Nucleic Acids Res* **15**, 10455-67 (1987).
  290. Bassler, B.L., Wright, M., Showalter, R.E. & Silverman, M.R. Intercellular signalling in *Vibrio harveyi*: sequence and function of genes regulating expression of luminescence. *Mol Microbiol* **9**, 773-86 (1993).
  291. Bassler, B.L., Wright, M. & Silverman, M.R. Multiple signalling systems controlling expression of luminescence in *Vibrio harveyi*: sequence and function of genes encoding a second sensory pathway. *Mol Microbiol* **13**, 273-86 (1994).
  292. Bassler, B.L., Wright, M. & Silverman, M.R. Sequence and function of LuxO, a negative regulator of luminescence in *Vibrio harveyi*. *Mol Microbiol* **12**, 403-12 (1994).
  293. Freeman, J.A. & Bassler, B.L. Sequence and function of LuxU: a two-component phosphorelay protein that regulates quorum sensing in *Vibrio harveyi*. *J Bacteriol* **181**, 899-906 (1999).
  294. Freeman, J.A., Lilley, B.N. & Bassler, B.L. A genetic analysis of the functions of LuxN: a two-component hybrid sensor kinase that regulates quorum sensing in *Vibrio harveyi*. *Mol Microbiol* **35**, 139-49 (2000).
  295. Freeman, J.A. & Bassler, B.L. A genetic analysis of the function of LuxO, a two-component response regulator involved in quorum sensing in *Vibrio harveyi*. *Mol Microbiol* **31**, 665-77 (1999).
  296. Lilley, B.N. & Bassler, B.L. Regulation of quorum sensing in *Vibrio harveyi* by LuxO and sigma-54. *Mol Microbiol* **36**, 940-54 (2000).
  297. Taga, M.E. & Bassler, B.L. Chemical communication among bacteria. *Proc Natl Acad Sci U S A* **100 Suppl 2**, 14549-54 (2003).
  298. Slater, H., Alvarez-Morales, A., Barber, C.E., Daniels, M.J. & Dow, J.M. A two-component system involving an HD-GYP domain protein links cell-cell signalling to pathogenicity gene expression in *Xanthomonas campestris*. *Mol Microbiol* **38**, 986-1003 (2000).

299. He, Y.W. et al. Dual signaling functions of the hybrid sensor kinase RpfC of *Xanthomonas campestris* involve either phosphorelay or receiver domain-protein interaction. *J Biol Chem* **281**, 33414-21 (2006).
300. Zhang, Z. et al. Dual regulation role of GH3.5 in salicylic acid and auxin signaling during *Arabidopsis*-*Pseudomonas syringae* interaction. *Plant Physiol* **145**, 450-64 (2007).
301. Barber, C.E. et al. A novel regulatory system required for pathogenicity of *Xanthomonas campestris* is mediated by a small diffusible signal molecule. *Mol Microbiol* **24**, 555-66 (1997).
302. He, Y.W. et al. Genome scale analysis of diffusible signal factor regulon in *Xanthomonas campestris* pv. *campestris*: identification of novel cell-cell communication-dependent genes and functions. *Mol Microbiol* **59**, 610-22 (2006).
303. Dow, M. Diversification of the function of cell-to-cell signaling in regulation of virulence within plant pathogenic xanthomonads. *Sci Signal* **1**, pe23 (2008).
304. Siciliano, F. et al. Analysis of the molecular basis of *Xanthomonas axonopodis* pv *citri* pathogenesis in *Citrus limon*. *electron. j. biotechnol.* **9**, 4 (2006).
305. Newman, K.L., Almeida, R.P., Purcell, A.H. & Lindow, S.E. Cell-cell signaling controls *Xylella fastidiosa* interactions with both insects and plants. *Proc Natl Acad Sci U S A* **101**, 1737-42 (2004).
306. Chatterjee, S., Newman, K.L. & Lindow, S.E. Cell-to-cell signaling in *Xylella fastidiosa* suppresses movement and xylem vessel colonization in grape. *Mol Plant Microbe Interact* **21**, 1309-15 (2008).
307. Chatterjee, S., Almeida, R.P. & Lindow, S. Living in two worlds: the plant and insect lifestyles of *Xylella fastidiosa*. *Annu Rev Phytopathol* **46**, 243-71 (2008).
308. Chatterjee, S., Wistrom, C. & Lindow, S.E. A cell-cell signaling sensor is required for virulence and insect transmission of *Xylella fastidiosa*. *Proc Natl Acad Sci U S A* **105**, 2670-5 (2008).
309. Kato, M., Mizuno, T., Shimizu, T. & Hakoshima, T. Refined structure of the histidine-containing phosphotransfer (HPT) domain of the anaerobic sensor kinase ArcB from *Escherichia coli* at 1.57 Å resolution. *Acta Crystallogr D Biol Crystallogr* **55**, 1842-9 (1999).
310. Ikegami, T. et al. Solution structure and dynamic character of the histidine-containing phosphotransfer domain of anaerobic sensor kinase ArcB from *Escherichia coli*. *Biochemistry* **40**, 375-86 (2001).
311. Raghava, G.P.S. APSSP: Advanced Protein Secondary Structure Prediction Server (<http://imtech.res.in/raghava/apssp/>).

LISTA DE ANEXOS

- a. Expression, crystallization and preliminary crystallographic analysis of PilZ(XAC1133) from *Xanthomonas axonopodis* pv. *citri*. *Acta Crystallogr Sect F Struct Biol Cryst Commun.*; **65**:304-6. Guzzo CR, Farah CS.
Localizado na página 80
- b. PilZ Protein Structure and Interactions with PilB and the FimX EAL Domain: Implications for control of Type IV pilus biogenesis. Guzzo, C. R., Salinas, R. K., Andrade, M. O. & Farah, C. S. (to be publish).
Localizado na página 94
- c. Expression, purification, crystallization and preliminary X-ray analysis of YaeQ (XAC2396) from *Xanthomonas axonopodis* pv. *citri*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **61**, 493-5. Guzzo, C. R., Nagem, R. A., Galvao-Botton, L. M., Guimaraes, B. G., Medrano, F. J., Barbosa, J. A. & Farah, C. S. (2005).
Localizado na página 164
- d. Structure of *Xanthomonas axonopodis* pv. *citri* YaeQ reveals a new compact protein fold built around a variation of the PD-(D/E)XK nuclease motif. *Proteins* **69**, 644-51. Guzzo, C. R., Nagem, R. A., Barbosa, J. A. & Farah, C. S. (2007).
Localizado na página 178
- e. Expression, crystallization and preliminary crystallographic analysis of SufE (XAC2355) from *Xanthomonas axonopodis* pv. *citri*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **62**, 268-70. Guzzo, C. R., Silva, L. R., Galvao-Botton, L. M., Barbosa, J. A. & Farah, C. S. (2006).
Localizado na página 235

SÚMULA CURRICULAR

DADOS PESSOAIS

Nome: Cristiane Rodrigues Guzzo
 Data de Nascimento: 13/04/1982 Nacionalidade: Brasileira Naturalidade: São Paulo
 Estado Civil: Casada RG: 29.484.572-0 CPF: 311.890.928-59
 Email: crisguzzo@uol.com.br

Endereço de Trabalho
 Rua: Av. Prof. Lineu Prestes No. 748
 Bairro: Butantã CEP: 05508-900 Cidade/Estado: São Paulo/SP
 Telefone: 0xx(11) 3815.3257 Fax: 0xx(11)3815.5579

FORMAÇÃO

Graduação

Nome do Curso: Bacharelado em Química com atribuições biotecnológicas e industriais Período: 2001 - 2005
 Instituição: Universidade de São Paulo – Instituto de Química Cidade/Estado: São Paulo/SP

Pós-Graduação Strictu Sensu (Doutorado)

Nome do Curso: Bioquímica Ano de Ingresso: 2005
 Instituição: Universidade de São Paulo – Instituto de Química Cidade/Estado: São Paulo/SP
 Orientador: Dr. Shaker Chuck Farah
 Título da Tese/Dissertação: Estudos estruturais de proteínas do sistema Rpf e hipotéticas conservadas do fitopatógeno *Xanthomonas axonopodis* pv *citri*

Atividades extracurriculares (estágios e outras atividades que, comprovavelmente, não foram exigência de currículo de graduação)

1- Professora de orgânica no cursinho do CUCA
 Instituição: UNESP Cidade/Estado: Araquara/SP
 Período: Junho de 2001 a Fevereiro de 2002

2 – Iniciação científica
 Instituição: UNESP Cidade/Estado: Araquara/SP
 Período: Abril de 2001 a Fevereiro de 2002
 Orientador: Dr. Younes Messadeq

3- Iniciação científica
 Instituição: USP - IQ Cidade/Estado: São Paulo/SP
 Período: Agosto de 2002 a Dezembro de 2004
 Orientador: Dr. Shaker Chuck Farah
 Bolsa da FAPESP no período Janeiro de 2002 a Dezembro de 2004

Participação em Cursos Nacionais e Internacionais:

1- “Validação e Modelo de Estruturas de Proteínas”, Centro Nacional de Pesquisa Tecnológicas em Informática para a Agricultura – CNPTIA, em Campinas-SP, no dia 01 de julho de 2005.

2- “Workshop em Bioinformática Estrutural, no Laboratório Nacional de Luz Síncrotron”, realizado no campus do laboratório nos dias 28 e 29 de junho de 2007.

3-“Advanced Certificate in Protein Crystallography on the Web Course Index 2006 -2007” , School of Crystallography, Birkbeck College, University of London, England. November 2006 – November 2007.

4- “RapiData 2009, Collection and Structure Solving”. A practical course in macromolecular X-ray diffraction measurement. Biology department and National Synchrotron Light Source, Brookhaven National Laboratory, USA. April 20-24, 2009.

Cursos e Aulas Ministradas em Disciplinas de Pós-Graduação:

Aula:

1- “Cristalografia e difração de raios-X de proteínas” em 2007 e 2009 na disciplina: Estudos de proteínas de interesse diagnóstico pela aplicação de técnicas de biologia molecular e estrutural (FBC-5759), do Curso de pós-graduação em Farmácia, Área de Análises Clínicas. Faculdade de Ciências Farmacêuticas, Universidade de São Paulo, SP, Brazil.

Ministrou os cursos:

- 1- “1º curso de Verão em Bioquímica e Biologia Molecular” no Departamento de Bioquímica do Instituto de Química da USP, SP, Brazil. Período de 9-20 de Janeiro de 2006.
- 2- “1º curso de Inverno – Temas Avançados de Bioquímica e Biologia Molecular” no Departamento de Bioquímica do Instituto de Química da USP, SP, Brazil. Período de 10-21 de Julho de 2006.
- 3- “III curso de Inverno – Temas Avançados de Bioquímica e Biologia Molecular” no Departamento de Bioquímica do Instituto de Química da USP, SP, Brazil. Período de 21-01 de Agosto de 2008.

Publicações

- 1- Galvao-Botton LM, Katsuyama AM, Guzzo CR, Almeida FC, Farah CS, Valente AP, 2003. “High-throughput screening of structural proteomics targets using NMR.”. *FEBS Lett.* 25; 552(2-3):207-13.
- 2- C. R. Guzzo, R. A. P. Nagem, L. M. P. Galvão-Botton, B. G. Guimarães, F. J. Medrano, J. A. R. G. Barbosa and C. S. Farah, 2005. “Expression, purification, crystallization and preliminary X-ray analysis of YaeQ (XAC2396) from *Xanthomonas axonopodis* pv. *citri*”. *Acta Cryst.* **F61**, 493-495
- 3- Guzzo CR, Silva LR, Galvao-Botton LM, Barbosa JA, Farah CS. Expression, crystallization and preliminary crystallographic analysis of SufE (XAC2355) from *Xanthomonas axonopodis* pv. *citri*. *Acta Crystallograph Sect F Struct Biol Cryst Commun.* 2006 Mar 1;62(Pt 3):268-70. Epub 2006 Feb 24.
- 4- Andrade MO, Alegria MC, Guzzo CR, Docena C, Rosa MC, Ramos CH, Farah CS. The HD-GYP domain of RpfG mediates a direct linkage between the Rpf quorum-sensing pathway and a subset of diguanylate cyclase proteins in the phytopathogen *Xanthomonas axonopodis* pv. *citri*.
- 5- Cristiane R. Guzzo, Ronaldo A. P. Nagem, João A. R. G. Barbosa, and Chuck S. Farah. The structure of *Xanthomonas axonopodis* pv. *Citri* YaeQ reveals new compact protein fold built around a variation of the PD-(D/E)XK nuclease motif. *PROTEINS: Structure, Function, and Bioinform.* v. 69, p. 644-651, 2007
- 6- Guzzo, Cristiane R.; FARAH, C. S.. Expression, crystallization and preliminary crystallographic analysis of PilZ from pv. *Acta Crystallographica. Series F*, v. 65, p. 304-306, 2009
- 7- Guzzo CR, Salinas RK, Andrade MO, Farah CS. PILZ protein structure and interactions with PILB and the FIMX EAL domain: implications for control of type IV pilus biogenesis. *J Mol Biol.* 2009 Nov 6;393(4):848-66. Epub 2009 Jul 29.
- 8- Hauk, Pricila ; Guzzo, Cristiane R. ; Ho, Paulo L. ; Farah, Chuck S. . Crystallization and preliminary X-ray analysis of LipL32 from serovar Copenhageni. *Acta Crystallographica. Series F*, v. 65, p. 307-309, 2009.
- 9- Hauk, Pricila ; GUZZO, C. R. ; Henrique Roman Ramos ; Ho, Paulo L. ; Farah, Chuck S.. Structure and Calcium-Binding Activity of LipL32, the Major Surface Antigen of Pathogenic *Leptospira* sp.. *Journal of Molecular Biology*, v. 390, p. 722-736, 2009.

Textos em jornais de notícias/revistas

1- GUZZO, C. R. ; NAGEM, R. A. ; BARABOSA, J. A. ; FARAH, C. S.. Xanthomonas axonopodis pv citri YaeQ structure reveals new compact protein fold built around a variation of the PD-(D/E)XK nuclease motif. Activity Report/Brazilian Synchrotron Light Laboratory, Brasil, p. 55 - 59, 01 nov

Resumos publicados em anais de congressos

1- GUZZO, C. R. ; GALVÃO-BOTTON, L. ; F.J.M. Martin ; FARAH, C. S.. Purification and crystallization of three Xanthomonas axonopodis pv citri proteins for structural studies. In: XXXII Reunião Anual de SBBq, 2003, Caxambu - MG. Resumos da XXXII Reunião Anual de SBBq, 2003.

Palavras-chave: Xanthomonas axonopodis.

Grande área: Ciências Biológicas / Área: Bioquímica / Subárea: Química de Macromoléculas /

Especialidade: Proteínas.

Referências adicionais: Classificação do evento: Nacional; Brasil/ Inglês; Meio de divulgação: Impresso
Guzzo, C.R. Galvão-Botton, L. Martin, F.J.M. and Farah, C.S..

2- GUZZO, Cristiane R ; BOTTON, Leonor Galvao de ; MEDRANO, F. J. ; BARBOSA, Joao Alexandre Ribeiro Gonçalves ; GUIMARÃES, Beatriz Gomes ; FARAH, Chuck Shaker. Purification and crystallization of three Xanthomonas axonopodis pv citri proteins for structural studies. In: XIV Reunião Anual de Usuários do LNLS, 2004, Campinas, 2004.

3- GUZZO, Cristiane R ; BOTTON, Leonor Galvao de ; GUIMARÃES, Beatriz Gomes ; MEDRANO, F. J. ; BARBOSA, Joao Alexandre Ribeiro Gonçalves ; FARAH, Chuck Shaker. Purification and crystallization of three Xanthomonas axonopodis pv. citri proteins for structural studies. In: XXXIII Reunião Anual da SBBq, 2004, Caxambu, 2004.

4- GUZZO, C. R. ; NAGEM, R. A. P. ; GALVAO-BOTTON, L.M.P ; GUIMARAES, B. G. ; MEDRANO, F. J. ; BARBOSA, J. A. R. G. ; FARAH, C. S.. Crystalization and Structure Determination of YaeQ from Xanthomonas axonopodis pv citri.. In: XXXIV Reunião Anual de SBBq, 2005, Águas de Lindóia. Resumos da XXXIV Reunião Anual de SBBq, 2005.

Palavras-chave: Crystalization; Xanthomonas axonopodis.

Grande área: Ciências Biológicas / Área: Bioquímica / Subárea: Química de Macromoléculas /

Especialidade: Proteínas.

Referências adicionais: Classificação do evento: Nacional; Brasil/ Inglês; Meio de divulgação: Impresso

5- GUZZO, C. R. ; NAGEM, R. A. P. ; GALVAO-BOTTON, L.M.P ; GUIMARAES, B. G. ; MEDRANO, F. J. ; BARBOSA, J. A. R. G. ; FARAH, C. S.. New protein fold from Xanthomonas axonopodis pv. citri YaeQ. In: Keystone Symposium: Frontiers in Structural Biology, 2006, Colorado. Livro de Resumos, 2006.

Referências adicionais: Classificação do evento: Internacional; Estados Unidos/ Inglês; Meio de divulgação: Impresso

6- GUZZO, C. R. ; RIBEIRO, L. ; ZUCKER, P. J. ; GALVAO-BOTTON, L. P. G. ; BARBOSA, J. A. R. G. ; FARAH, C. S.. Purificação, Cristalização e Análise Estrutural da Proteína SufE e do domínio Hpt da RpfC de Xanthomonas axonopodis pv. citri e da GAPDH de Escherichia coli. In: XVI Reunião Anual de Utilitários do Laboratório Nacional de Luz Síncrotron, 2006, Campinas, SP. Livro de Resumos, 2006.

Referências adicionais: Classificação do evento: Nacional; Brasil/ Português; Meio de divulgação: Impresso

7- ANDRADE, M. O. ; GUZZO, C. R. ; ALEGRIA, M. C. ; FARAH, C. S.. Protein - Protein Interaction Network of the Quorum Sensing Response in Xanthomonas axonopodis pv. citri.. In: XIII International Congress on Molecular Plant-Microbe Interactions, 2007, Sorrento. Livro de Resumos, 2007.

Referências adicionais: Classificação do evento: Internacional; Itália/ Inglês; Meio de divulgação: Impresso

8- ANDRADE, M. O. ; ALEGRIA, M. C. ; GUZZO, C. R. ; FARAH, C. S.. The HD-GYP Domain of RpfG Mediates a Direct Linkage Between the Rpf Quorum-Sensing Pathway and a Subset of Diguanilate

Cyclase Proteins in the Phytopathogen *Xanthomonas axonopodis* pv. *citri*.. In: XXXVI Reunião Anual de SBBq e 10th IUBMB Conference, 2007, Salvador, BA. Livro de Resumos, 2007.

Referências adicionais: Classificação do evento: Internacional; Holanda/ Inglês; Meio de divulgação: Impresso

9- GUZZO, C. R. ; NAGEM, R. A. P. ; BARBOSA, J. A. R. G. ; FARAH, C. S.. The Structure of *Xanthomonas Axonopodis* pv. *citri* YaeQ Reveals a New Compact Protein Fold Built Around a Variation of the PD-(D/E)XK Nuclease Motif.. In: XXXVI Reunião Anual de SBBq e 10th IUBMB Conference, 2007, Salvador, BA. Livro de Resumos, 2007.

Referências adicionais: Classificação do evento: Internacional; Brasil/ Inglês; Meio de divulgação: Impresso

10- GUZZO, C. R. ; SILVA, L. R. ; GALVAO-BOTTON, L.M.P ; BARBOSA, J. A. R. G. ; FARAH, C. S.. Crystal Structure of SufE from *Xanthomonas axonopodis* pv. *citri*. In: XXXVI Reunião Anual de SBBq e 10th IUBMB Conference, 2007, Salvador, BA. Livro de Resumos, 2007.

Referências adicionais: Classificação do evento: Internacional; Brasil/ Inglês; Meio de divulgação: Impresso

11- Beton, Daniela ; Guzzo, Cristiane Rodrigues ; Ribeiro, Alberto de Freitas ; Farah, Chuck Shaker ; Terra, Walter Ribeiro ; during the XXXVIII Annual Meeting of The Brazilian Society for Biochemistry and Molecular Biology (SBBq) in the Convention Center of Hotel Monte Real, Águas de Lindóia, São Paulo, Brazil, on May 16 to 19, 2009.

12- Guzzo Cristiane Rodrigues; Salinas, Roberto Kopk ; Andrade, Maxuel de Oliveira, Farah, Chuck Shaker; presented the abstract entitled: "Structure determination of PilZ from *Xanthomonas axonopodis* pv. *citri* and their role in Type IV biogenesis," during the XXXVIII Annual Meeting of The Brazilian Society for Biochemistry and Molecular Biology (SBBq) in the Convention Center of Hotel Monte Real, Águas de Lindóia, São Paulo, Brazil, on May 16 to 19, 2009.

13- International Network of Protein Engineering Centers. PilZ protein structure and interactions with PilB and FimX EAL domain: implications for control of Type IV pilus Biogenesis. 2009. (Participações em eventos/Encontro).

Apresentação Oral

1- GUZZO, C. R. ; NAGEM, R. A. P. ; GALVAO-BOTTON, L.M.P ; GUIMARAES, B. G. ; MEDRANO, F. J. ; BARBOSA, J. A. R. G. ; FARAH, C. S.. Title: *The Crystal Structure of YaeQ from Xanthomonas axonopodis* pv. *citri*.. 2005. Cidade: Campinas; Inst. promotora/financiadora: 17ª - Reunião da Sociedade Brasileira de Cristalografia, LNLS. Apresentação oral da Cristiane Guzzo (aluna de DR).

2- GUZZO, C. R.; SALINAS, R. K.; ANDRADE, M. O.; FARAH, C. S. Title: "*Estrutura e Função de PilZ de Xanthomonas citri*" nos seminários gerais do departamento de bioquímica do programa de pós-graduação do Instituto de Química da Universidade de São Paulo, SP, Brazil em 2008. Apresentação oral da Cristiane Guzzo (aluna de DR).

3- GUZZO, C. R. ; NAGEM, R. A. P. ; GALVAO-BOTTON, L.M.P ; GUIMARAES, B. G. ; ; BARBOSA, J. A. R. G. ; ANDRADE, M. O.; Hauk, Pricila ; SALINAS, R. K.; Ho, Paulo L.; CICERO, D.O.; Contessa, G.C.; PERTINHEZ, T. ; M. Gallo; Katsuyama, A.M. ; PACI, M.; SPISNI, A.; Beton, Daniela; FARAH, C. S. Title: "*YaeQ and PilZ Proteins: Examples of Combining NMR and Crystallography in Structural Studies*" in the NMR in Structural Biology: from cloning to NMRPipe applications in 2009. Center for Structural Molecular Biology/Brazilian Synchrotron Light Laboratory, Campinas, Brazil. Apresentação oral da Cristiane Guzzo (aluna de DR).

4- GUZZO, C. R. ; NAGEM, R. A. P. ; GALVAO-BOTTON, L.M.P ; GUIMARAES, B. G. ; ; BARBOSA, J. A. R. G. ; ANDRADE, M. O.; Hauk, Pricila ; SALINAS, R. K.; Ho, Paulo L.; CICERO, D.O.; Contessa, G.C.; PERTINHEZ, T. ; M. Gallo; Katsuyama, A.M. ; PACI, M.; SPISNI, A.; Beton, Daniela; FARAH, C. S.. "*NMR Screening of Structural Biology Targets and the Protein Crystallography Laboratory at the Instituto de Química, Universidade de São Paulo*". In the IV CeBiME/LNLS Structural Molecular Biology

Workshop: techniques, tools and developments in protein crystallization in 2009. Center for Structural Molecular Biology/Brazilian Synchrotron Light Laboratory, Campinas, Brazil. Apresentação oral da Cristiane Guzzo (aluna de DR).