Alexandre Moutran

**Modelagem Molecular das proteínas
captadoras de Molibdato (ModA) e
oligopeptídeos (OppA) de *Xanthomonas
axonopodis* pv. *citri*.**

Tese apresentada ao Instituto de Ciências
Biomédicas/Departamento de Microbiologia
da Universidade de São Paulo, para
concorrer ao Título de Doutor, pelo curso de
Pós-Graduação em Microbiologia.

SÃO PAULO

2009

Alexandre Moutran

**Modelagem Molecular das proteínas captadoras de
Molibdato (ModA) e oligopeptídeos (OppA) de
Xanthomonas axonopodis pv. *citri*.**

Tese apresentada ao Instituto de Ciências Biomédicas da
Universidade de São Paulo, para obtenção do Título de
Doutor em Ciências.

Área de concentração: Microbiologia

Orientadora: Prof.^a Dra. Rita de Cássia Café Ferreira

Co-orientadora: Dra. Andrea Balan

São Paulo
2009

DADOS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
Serviço de Biblioteca e Informação Biomédica do
Instituto de Ciências Biomédicas da Universidade de São Paulo

© reprodução total

Moutran, Alexandre.

Modelagem molecular das proteínas captadoras de Molibdato (ModA) e Oligopeptídeo (OppA) de *Xanthomonas axonopodis* pv. *citri* / Alexandre Moutran. -- São Paulo, 2009.

Orientador: Rita de Cassia Café Ferreira.

Tese (Doutorado) – Universidade de São Paulo. Instituto de Ciências Biomédicas. Departamento de Microbiologia. Área de concentração: Microbiologia. Linha de pesquisa: Bioinformática.

Versão do título para o inglês: Molecular modeling of molybdate (ModA) and oligopeptide (OppA) uptake proteins in *Xanthomonas axonopodis* pv. *citri*.

Descritores: 1. Modelagem Molecular 2. Oligopeptídeos 3. *Xanthomonas* 4. Bioinformática 5. Proteína de transporte I. Ferreira, Rita de Cassia Café II. Universidade de São Paulo. Instituto de Ciências Biomédicas. Programa de Pós-Graduação em Microbiologia. III. Título.

ICB/SBIB24/2009

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS

Candidato(a): Alexandre Moutran.

Título da Tese: Modelagem molecular das proteínas captadoras de Molibdato (ModA) e Oligopeptídeo (OppA) de *Xanthomonas axonopodis* pv. *citri*.

Orientador(a): Rita de Cassia Café Ferreira.

A Comissão Julgadora dos trabalhos de Defesa da Tese de Doutorado, em sessão pública realizada a ...../...../....., considerou

( ) Aprovado(a)

( ) Reprovado(a)

Examinador(a): Assinatura: .....
Nome: .....
Instituição: .....

Examinador(a): Assinatura: .....
Nome: .....
Instituição: .....

Examinador(a): Assinatura: .....
Nome: .....
Instituição: .....

Examinador(a): Assinatura: .....
Nome: .....
Instituição: .....

Presidente: Assinatura: .....
Nome: .....
Instituição: .....

Dedico

*Aos meus pais, irmãos e minha noiva
Andrea por todo o apoio, compreensão, paciência
e amor.*

*Dedico também aos meus avós que
faleceram em 2008 e que durante mais de 90
anos deixaram um exemplo de vida a ser seguido.
Um casal que após 70 anos de casamento, ainda
trocavam juras de amor e recebiam a todos com
um carinho celestial. Aos meus amados avós,
dedico esta Tese e que Deus os tenha.
Sinto muita saudades.....*

AGRADECIMENTOS

A Professora Doutora Rita de Cássia Café Ferreira, que durante muitos anos acreditou no meu trabalho e na minha capacidade, com paciência e entusiasmo. Apesar das diferenças de sotaques “carioquês” e “caipires”, tive a possibilidade de obter um grande crescimento pessoal e profissional. Foi um grande prazer poder trabalhar durante praticamente uma década juntos e desenvolver uma amizade duradoura.

Ao Professor Doutor Luis Carlos de Souza Ferreira, uma pessoa de grande sabedoria e iluminação, um professor e pesquisador exemplar que muito colaborou com o meu trabalho durante todos estes anos.

A Doutora Andrea Balan, co-orientadora e amiga que “me seguiu” até a Itália, aonde participamos de um congresso e realizei parte do meu trabalho. Agradeço toda a sua colaboração e espero encontrar novamente aquele bar em Florença aonde tomamos uma das melhores cervejas da minha vida e que misteriosamente sumiu no labirinto de ruas daquela cidade.

Ao Professor Doutor Goran Neshich e Doutora Paula Falcão por permitirem que eu realizasse parte do meu trabalho no seu laboratório e por todo o ensinamento oferecido.

A Professora Doutora Anna Tramontano que me aceitou em seu laboratório, sem saber qualquer referência minha anterior. Agradeço a todos os alunos e pesquisadores do departamento de bioinformática da Università di Roma La Sapienza pela colaboração.

A todos meus colegas do ICB II, não somente do meu laboratório, mas de todos os departamentos, que durante anos colaboraram com informações, dados, experimentos e também pedindo pizza, Mc Donalds entre outras colaborações ditas “vitais” ao bom andamento da pesquisa.

Agradeço a minha noiva, talvez já possa dizer esposa nesta data, que logo que me conheceu, percebeu que não seria fácil conviver comigo. Ela sempre me incentivou nos momentos de dificuldade. Eu a considero minha “psicóloga particular”, porque comecei a perceber o mundo de uma forma diferente e finalmente compreendi que “nós cientistas somos loucos e babamos muito”, afinal, que ser humano normal pode ficar olhando para bactérias e chamando-as carinhosamente de filhas, ou conversando com o seu monitor pensando que tem alguém realmente na sua frente (auto-descrição).

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro.

Enfim a todos que contribuíram para a realização deste trabalho.

Obrigada por tudo!

“Vida”

Já escrevi sobre caminhos da vida,
e quantos caminhos trilhei desde então,
quando o caminho do mestrado
completava.

Eu sequer imaginava,
que pegaria um avião,
rumo a uma nova vida.

Agora escrevo somente sobre esta vida,
para lembrar o que me faltava,
pois eu sempre achava,
que algo me faltava.

Sempre buscava,
e nada encontrava,
mas foi quando te perdi,
que percebi,
que nesta vida,
feita exclusivamente de uma ida,
somente o tempo me faltará,
para vivê-la.

Adeus meus avós,
A DEUS os meus avós.
obrigado por ensinar,
como caminhar,
nesta minha vida.

(por Alexandre Moutran)

RESUMO

MOUTRAN, A. **Modelagem Molecular das proteínas captadoras de Molibdato (ModA) e oligopeptídeos (OppA) de *Xanthomonas axonopodis* pv. *citri***. 122 f. Tese (Doutorado em Microbiologia) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2009.

Sistemas de transporte tipo ABC são responsáveis pelo transporte de uma grande variedade de substâncias dentre elas os oligopeptídeos e molibdato.

Neste trabalho estudamos dois sistemas de transportadores do tipo ABC (*mod*, envolvido na captação de molibdato e o *opp* na captação de oligopeptídeos) presentes na bactéria *Xanthomonas axonopodis* pv. *citri* (*Xac*). Em particular analisamos a organização genética dos óperons *mod* e as proteínas ModA e OppA, componentes solúveis localizados no periplasma e responsáveis pela ligação aos substratos. Por meio de técnicas de modelagem molecular, definimos modelos estruturais para as proteínas ModA e OppA. Para a proteína ModA caracterizamos cinco resíduos que compõem a cavidade ligadora e são responsáveis pelas interações com o íon molibdato, assim como a sua similaridade estrutural e sequencial com ortólogos de 3 grupos distintos de bactérias. Para a OppA, descrevemos o seu comportamento na ancoragem de diferentes oligopeptídeos. Avaliamos parâmetros como a extensão da cadeia e estabelecemos uma ordem crescente de afinidade entre os oligopeptídeos com diferente composição residual e a proteína OppA.

Palavras-chave: OppA. ModA. Modelagem molecular. *Xanthomonas axonopodis* pv. *citri*. Transportadores ABC.

ABSTRACT

MOUTRAN, A. **Molecular modeling of molybdate (ModA) and oligopeptide (OppA) uptake proteins in *Xanthomonas axonopodis* pv. *citri***. 122 f. Tese (Doutorado em Microbiologia) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2009.

ABC transport systems are responsible for the uptake of a large variety of substrates, including oligopeptides and molybdate. In this work we studied two ABC transporter systems (mod and opp responsible for molybdate and oligopeptide uptake, respectively) present in plant pathogen *Xanthomonas axonopodis* pv. *citri* (Xac). We investigated the genetic organization of mod operon and, particularly, structural features of periplasmic components, ModA and OppA proteins, of the uptake systems. Using molecular modeling techniques, we defined the structural models of both ModA and OppA proteins. Based on the ModA structural model, amino acid residues involved in molybdate interaction were identified. In addition, both the structural and sequence similarities of Xac ModA and other bacterial orthologs with experimentally defined structural organizations were described. Based on the OppA structural model, we applied molecular docking tools to determine the binding specificity for different oligopeptides regarding number and amino acid composition. Collectively, our results represent an important contribution to the study of ABC transporters in an economically relevant phytopathogen bacterial species.

Key word: OppA. ModA. Molecular modeling. *Xanthomonas axonopodis* pv. *citri*, ABC Transporters.

LISTA DE FIGURAS

Figura 1	Organização dos transportadores ABC em bactérias gram-negativas.....	19
Figura 2	Transporte e regulação de molibdato em <i>Eco</i>	23
Figura 3	Estrutura tridimensional da proteína ModA de <i>Eco</i>	24
Figura 4	Organização genética do operon <i>opp</i> de <i>Stm</i> e <i>Xac</i>	26
Figura 5	Estrutura tridimensional da proteína OppA de <i>Stm</i>	28
Figura 6	Estrutura da Appa de <i>Bsu</i>	30
Figura 7	Organização genética do óperon <i>mod</i> nas linhagens de <i>Xac</i> , <i>Xcc</i> , <i>Xcv</i> e <i>Xoo</i>	46
Figura 8	Alinhamento sequencial de ModA das linhagens de <i>Xac</i> , <i>Xoo</i> , <i>Xcv</i> e <i>Xcc</i>	47
Figura 9	Organização genética de óperons <i>mod</i> encontrados em diferentes espécies de bactérias	50
Figura 10	Alinhamentos das sequencias de aminoácidos de 15 ortólogos de ModA encontrados em bactérias.....	51
Figura 11	Análise filogenética de diferentes ortólogos da proteína ModA de <i>bactérias</i>	52
Figura 12	Gráfico de Ramachandran do modelo gerado para ModA de <i>Xac</i>	54
Figura 13	Modelo estrutural da proteína ModA de <i>Xac</i>	55
Figura 14	Alinhamento estrutural entre as proteínas ModA de <i>Eco</i> e do modelo de <i>Xac</i>	56
Figura 15	Alinhamento estrutural do modelo da ModA de <i>Xac</i> definido no presente estudo (vermelho) com a estrutura cristalográfica.....	58
Figura 16	Resíduos da proteína ModA de <i>Xac</i> que interagem com o molibdato.....	60
Figura 17	Sítios de ligação da proteína ModA de <i>Xac</i>	62
Figura 18	Comparação dos sítios de ligação entre o modelo da proteína ModA de <i>Xac</i> e o íon molibdato comparado com sua estrutura cristalina.....	63
Figura 19	Sobreposição dos carbonos alfa das estruturas tridimensionais ..	64
Figura 20	Modelo estrutural da OppA de <i>Xac</i>	67
Figura 21	Ancoragem do tripeptídeo KWK na cavidade ligadora da OppA de <i>Xac</i>	69
Figura 22	Análise de ancoragem com diferentes oligopeptídeos de polilisinas.....	70
Figura 23	Análise de ancoragem com diferentes oligopeptídeos e a proteína OppA de <i>Xac</i>	71
Figura 24	Análise de ancoragem com diferentes oligopeptídeos à proteína OppA de <i>Xac</i>	72

LISTA DE TABELAS

Tabela 1	Similaridade entre os ortólogos das proteínas ModA, ModB e ModC.....	48
Tabela 2	Interações moleculares entre as proteínas ModA com a estrutura definida e o íon molibdato.....	62
Tabela 3	Alinhamento dos carbonos alfa entre os ortólogos da ModA.....	64

LISTA DE ABREVIATURAS E SIGLAS

a.a. - aminoácido
ABC - ATP-binding cassette
ADP - adenosina difosfato
Afu - *Archaeoglobus fulgidus*
ATP - adenosina trifosfato
Avi - *Azotobacter vinelandii*
Bbu - *Borrelia burgdorferie*
Bsu - *Bacillus subtilis*
DIM - domínio integral de membrana
DLN - domínios de ligação de nucleotídeos
Eco - *Escherichia coli*
Gsu – *Geobacter sulfurreducens*
 K_d - Coeficiente de dissociação
Kda - Kilo Dalton
Lla - *Lactococcus lactis*
LS - ligador de substrato
Mlo – *Mezorhizobium loti*
Mo - símbolo do elemento químico Molibdênio
Mo-co - cofator de molibdênio
Mtu - *Mycobacterium tuberculosis*
nM - nano mols
Pae - *Pseudomonas aeruginosa*
pb - pares de base
PDB - Protein Data Bank
pv. - patovares
Rca - *Rhodobacter capsulatus*
RMN - Ressonância magnética nuclear
Rso - *Ralstonia solanacearum*
SBP - sulfate binding protein
Sco - *Streptomyces coelicolor*
Sfl – *shigella flexneri*

Sme - Sinorhizobium meliloti

Stm - Salmonella typhimurium

Sub - Streptococcus uberis

Tma - Thermotoga maritima

W - símbolo do elemento químico tungstênio

Xac - Xanthomonas axonopodis pv. citri

Xag - Xanthomonas axonopodis pv. glycines

Xav - Xanthomonas axonopodis pv. vesicatoria

Xcc - Xanthomonas campestris pv. campestris

Xcv - Xanthomonas campestris vesicatoria

Xoo - Xanthomonas oryzae pv. oryzae

Ype – Yersinia pestis

LISTA DE SÍMBOLOS

Aminoácidos	Código de três letras	Código de uma letra
Alanina	Ala	A
Arginina	Arg	R
Asparagina	Asn	N
Ácido aspártico	Asp	D
Asparagina ou Ácido aspártico	Asx	B
Cisteína	Cys	C
Ácido glutâmico	Glu	E
Glutamina	Gln	Q
Glutamina ou Ácido glutâmico	Glx	Z
Glicina	Gly	G
Histidina	His	H
Isoleucina	Ile	I
Leucina	Leu	L
Lisina	Lys	K
Metionina	Met	M
Fenilalanina	Phe	F
Prolina	Pro	P
Serina	Ser	S
Treonina	Thr	T
Triptofano	Try	W
Tirosina	Tyr	Y
Valina	Val	V

SUMÁRIO

1	INTRODUÇÃO.....	16
1.1	Captação de Metais de Transição: íons Molibdênio/Tungstênio.....	20
1.2	Captação de Oligopeptídeos.....	25
1.3	Doenças Fitopatológicas e seu impacto econômico.....	31
1.4	O Projeto Genoma de <i>Xac</i> e <i>Xcc</i>.....	32
1.5	Biologia Computacional.....	34
2	OBJETIVOS.....	38
3	MATERIAIS E MÉTODOS.....	39
3.1	Análise Computacional.....	39
3.2	Processo de Modelagem.....	40
3.2.1	Seleção das Estruturas Moldes.....	40
3.2.2	Alinhamentos.....	41
3.2.3	Processo de Modelagem Estrutural da OppA e ModA de <i>Xac</i>	41
3.3	Dinâmica Molecular do melhor modelo da estrutura da ModA de <i>Xac</i>.....	42
3.4	Ancoragem do íon molibdênio da ModA de <i>Xac</i>.....	43
3.5	Ancoragem de oligopeptídeos à OppA de <i>Xac</i>.....	43
4	RESULTADOS.....	45
4.1	Identificação dos óperons <i>mod</i> e <i>opp</i> em <i>Xanthomonas</i>.....	45
4.2	Organização dos óperons <i>mod</i> e respectivos cistron em diferentes espécies de <i>Xanthomonas</i>.....	45
4.3	Análise da organização gênica do óperon <i>mod</i> em diferentes espécies bacterianas	48
4.3.1	Análise de similaridade da proteína ModA expressa por diferentes espécies de bactérias.....	51
4.4	Modelo tri-dimensional da proteína ModA de <i>Xac</i>.....	53
4.4.1	A cavidade ligadora de substrato da ModA de <i>Xac</i>	59
4.5	Modelagem molecular da proteína ligadora de oligopeptídeos (OppA) de <i>Xac</i>.....	65
4.5.1	Afinidade por oligopeptídeos da OppA de <i>Xac</i>	68
5	DISCUSSÃO.....	73
5.1	Modelagem da OppA.....	79
6	CONCLUSÃO.....	84
	REFERÊNCIAS BIBLIOGRÁFICAS	85
	ANEXOS A - Structural model and ligand interactions of the <i>Xanthomonas axonopodis</i> pv. <i>citri</i> oligopeptide-binding protein.....	97
	ANEXOS B - Crystallographic structure and substrate-binding interactions of the molybdate-binding protein of the phytopathogen <i>Xanthomonas axonopodis</i> pv. <i>citri</i>.....	106
	ANEXOS C - The molybdate-binding protein (ModA) of the plant pathogen <i>Xanthomonas axonopodis</i> pv. <i>citri</i>.....	115
	ANEXOS D - Purification and in vitro characterization of the maltose-binding protein of the plant pathogen <i>Xanthomonas citri</i>.....	118
	ANEXOS E – Estágios Realizados.....	126

1 INTRODUÇÃO

O envoltório celular bacteriano pode ser definido como um conjunto de membranas e estruturas macromoleculares que delimita a fronteira entre o meio intracelular e o ambiente externo (POXTRON, 1993). O envoltório celular de bactérias gram-positivas é formado pela membrana citoplasmática, uma densa camada de peptidoglicana rica em ácidos teicóicos e, em alguns casos, cápsulas (BEVERIDGE e GRAHAM, 1991). Por outro lado, o envoltório celular de bactérias gram-negativas mostra-se estruturalmente mais complexo e formado pela membrana citoplasmática; periplasma, mureína ou peptidoglicana, membrana externa e apêndices como flagelos e fímbrias (LUGTENBERG e VAN ALPHEN, 1983; SLEYTR *et al.*, 1993). Dentre as inúmeras funções da membrana plasmática, destaca-se a capacidade de controlar o tráfego de substâncias, atuando como uma barreira altamente seletiva que impede a passagem livre de moléculas e íons e possibilita assim a concentração de metabólitos específicos dentro da célula. Além disso, a excreção de substâncias pela célula também é feita através da membrana.

Moléculas como aminoácidos, açúcares, íons e sais minerais não conseguem passar livremente pela membrana e, por isso, devem ser especificamente transportadas. O transporte de substâncias através da membrana do meio externo para o meio interno e vice-versa ocorre com o auxílio de proteínas de transporte de membrana (SAURIN *et al.*, 1994). A característica mais importante do transporte mediado por carregadores protéicos é a sua natureza altamente específica. No entanto, embora a maioria dos carregadores protéicos mostrem grande afinidade por apenas um único tipo de molécula, alguns são capazes de interagir com toda uma classe de moléculas (DAVIDSON e MALONEY, 2007).

As proteínas transportadoras podem ser divididas em proteínas de canais e carreadoras. Proteínas do tipo canal funcionam como poros seletivos que são abertos em resposta a estímulos químicos ou eletrofisiológicos. Proteínas do tipo carreadoras dependem de energia para transportar o substrato contra um gradiente de concentração (DAHL *et al.*, 2004).

Os sistemas de transportes dependentes de energia são classificados como primários e secundários em função do tipo de energia empregada. Os sistemas primários de transporte utilizam energia gerada pela clivagem de ATP, enquanto que

os sistemas secundários fazem uso de outras fontes de energia celulares, como a força protomotora ou aquela gerada por sistemas de fosfotransferases. Entre os transportadores primários, destacam-se aqueles que ligam o ATP, também chamados de transportadores ABC, expressão derivada do inglês *ATP-binding cassette*, com ocorrência amplamente difundida entre os seres vivos, em particular, entre as bactérias (AMES *et al.*, 1992). Os sistemas de transporte do tipo ABC são constituídos por complexos formados por várias subunidades, denominados de permeases, envolvidos no transporte de substâncias, como diversos íons, aminoácidos, peptídeos, antibióticos, polissacarídeos e proteínas, através de membranas biológicas (BRAIBANT *et al.*, 2000). Transportadores ABC são classificados como importadores ou exportadores dependendo da direção de deslocamento da substância em relação à célula. Sistemas ABC de transporte dedicados à importação ou captação de substâncias são encontrados exclusivamente em bactérias (BRAIBANT *et al.*, 2000).

Os sistemas do tipo ABC são constituídos por três componentes funcionais e estruturais. Dois componentes apresentam-se associados à membrana citoplasmática enquanto o terceiro e último mostra-se solúvel nas bactérias gram-negativas, ou associado à membrana citoplasmática nas bactérias gram-positivas. Um dos componentes associados as membranas é responsável pela formação do poro por onde a substância é transportada. Esse componente é, em geral, formado por duas proteínas hidrofóbicas que atravessam a membrana citoplasmática, e por isto, denominado domínio integral de membrana (DIM) (Figura 1). Em bactérias, os DIM são formados por proteínas que apresentam de quatro a oito α -hélices (usualmente seis) que se organizam como homodímeros que delimitam o poro na membrana citoplasmática através do qual será feito o transporte do substrato (MOUREZ *et al.*, 1997).

O segundo componente dos sistemas ABC de transporte de bactérias é representado por dois domínios geradores de energia, proteínas associadas à face interna da membrana citoplasmática que ligam ATP e chamadas de domínios de ligação de nucleotídeos (DLN) (Figura 1). Os DLNs representam os elementos mais característicos e conservados dos sistemas ABC de transporte (YOUNG e HOLLAND, 1999). Os DLNs de transportadores ABC bacterianos acoplam a hidrólise do ATP com o processo de transporte e possuem várias seqüências conservadas que compartilham de 30 a 40% de identidade (LINTON e HIGGINS, 1998). O

elevado grau de conservação dessas seqüências permite a identificação presuntiva de transportadores ABC em genomas bacterianos. Todas as DLNs conhecidas estão envolvidas com o transporte de substratos através de membranas biológicas, com exceção da proteína UvrA de *Escherichia coli* (*Eco*), o fator 3 de alongamento da tradução e a proteína reguladora GCN20, ambas encontradas em leveduras (DOOLITTLE *et al.*, 1986; CHAKRABURTLY, 1999).

Os quatros domínios de transportadores ABC, dois DIMs e dois DLNs, podem ser encontrados livres ou fusionados de diversas maneiras (HIGGINS, 1992). Os sistemas de transporte ABC de bactérias envolvidos na captação de nutrientes possuem, em geral, um terceiro e último componente envolvido com a ligação ao substrato, também chamado de ligador de substrato (LS), responsável pela especificidade e afinidade do sistema (Figura 1). Em bactérias gram-negativas, esse componente é solúvel e encontra-se livre no periplasma enquanto que em bactérias gram-positivas os LS são encontrados na forma de lipoproteínas ancoradas por meio de uma acil-gliceril-cisteína N-terminal à face externa da membrana citoplasmática (DASSA, 2000).

Importadores**Exportadores**

Figura 1: Organização dos transportadores ABC em bactérias gram-negativas. Sistemas importadores ou exportadores são definidos em função da direção do tráfego de seus substratos (indicados por setas). O protótipo procariótico é composto de dois domínios integrais de membrana (DIM) e dois domínios de ligação a nucleotídeos (DLN). O componente de ligação do substrato (LS) está presente nos sistemas de importação de nutrientes (figura adaptada de BRAIBANT *et al.*, 2000 e HOLLENSTEIN *et al.*, 2007).

Os diversos componentes dos transportadores ABC bacterianos são codificados por múltiplos genes, que codificam polipeptídeos independentes, enquanto que em eucariontes um único gene codifica para uma proteína com múltiplos domínios. Os genes responsáveis pela codificação dos diferentes componentes que constituem os sistemas ABC de bactérias estão em geral organizados na forma de operons (YOUNG e HOLLAND, 1999).

A importância dos sistemas de transporte ABC pode ser avaliada pelo número de unidades encontradas em diferentes espécies de bactérias que tiveram seus genomas determinados. Em *Eco*, os genes responsáveis pela síntese de componentes de sistemas de transporte ABC representam 5% do total do genoma,

enquanto que em *Mycobacterium tuberculosis* (*Mtu*) esse total atinge 2,5% do genoma (BRAIBANT *et al.*, 2000; HIGGINS, 2001). De um modo geral, podem ser encontrados de 25 a cerca de 100 diferentes sistemas ABC, que ocupam de 2 a 7% dos respectivos genomas, em diferentes espécies de bactérias (TOMII e KANEHISA, 1998; YOUNG e HOLLAND, 1999).

A diversidade de funções desses sistemas de transporte demonstra a complexidade dos seus componentes e a importância de estudá-los e compreendê-los estruturalmente e fisiologicamente. Entre os sistemas de transporte responsáveis pela captação de substratos necessários para a sobrevivência da célula encontram-se os mecanismos capazes de transportar metais, açúcares, aminoácidos e enzimas, entre outros. A captação desses substratos na célula é mediada por receptores periplasmáticos que constituem uma classe de proteínas ligadoras que diferem em tamanho, sequência e especificidade pelo ligante. No entanto, quase todos apresentam uma estrutura com conformação α/β com características similares, consistindo de dois domínios globulares entre os quais o componente ligante fica ancorado (LAWSON *et al.*, 1998).

1.1 Captação de Metais de Transição: íons Molibdênio/Tungstênio

Todas as formas de vida requerem elementos orgânicos ou minerais para desempenhar os seus processos vitais. Estes elementos precisam ser mantidos dentro de estreitos limites para que a integridade funcional e estrutural das células possa ser preservada. Ao contrário de outros nutrientes, os minerais geralmente são metais e não podem ser sintetizados pelos organismos vivos, atuando em funções importantes como: (1) componentes estruturais de órgãos e tecidos, (2) constituintes de fluidos corporais como eletrólitos e (3) catalisadores em sistemas hormonais e enzimáticos (CANTARROW e SCHEPARTZ, 1969; McDowell, 1992). Os metais mostram afinidade por alguns resíduos de aminoácidos como histidina, cisteína, serina e tirosina (HILLE, 2002).

Dentre os metais requeridos por seres vivos, o molibdênio é utilizado por diversos sistemas biológicos. O molibdênio é o único metal de transição requerido pela maioria dos organismos vivos, incluindo bactérias, plantas e animais. Esse

metal é incorporado dentro dos sítios ativos de enzimas envolvidas nos processos de óxido-redução e que participam ativamente dos ciclos de nitrogênio, enxofre e carbono (KISKER *et al.*, 1997). O uso generalizado de molibdênio em sistemas biológicos talvez reflita semelhanças na composição dos nossos oceanos primitivos com os atuais. Algumas espécies utilizam tungstênio como alternativa ao molibdênio, o qual também apresenta versatilidade química e biodisponibilidade (HILLE, 2002).

As bactérias capturam molibdênio eficientemente do ambiente para a síntese dos cofatores de pelo menos 20 enzimas com atividade de óxido-redução (MOUNCEY *et al.*, 1995). Tanto o molibdênio como o tungstênio apresentam uma versatilidade química útil em sistemas biológicos e mostram capacidade de óxido-redução ativa sob condições fisiológicas (oscilando entre os estados de oxidação VI e IV), sendo que o estado de valência V é também acessível. Esses íons podem atuar como transdutores entre sistemas de óxido-redução de um ou dois elétrons e podem catalisar reações como a de hidroxilação dos centros de carbono sob condições mais moderadas requeridas por determinadas enzimas (HILLE, 2002).

O papel fisiológico das molibdoenzimas é fundamental e incluem a catálise dos passos chave nos metabolismos do carbono, nitrogênio e enxofre (HILLE, 2002). Com exceção das nitrogenases que contêm um grupamento de ferro-molibdênio-enxofre, o molibdênio e o tungstênio são incorporados a proteínas como cofator de molibdênio (Mo-co), que contem o átomo mononuclear Mo (ou W) complexado por um cofator orgânico denominado molibdopterina. Embora a função do ligante molibdoterina não tenha sido estabelecida de forma conclusiva, interações deste ligante com o metal são sensíveis às mudanças no estado de oxidação, indicando que a molibdoterina pode estar diretamente envolvida no mecanismo enzimático (KISKER *et al.*, 1997).

As molibdoenzimas estão divididas em três famílias que se diferenciam pelas estruturas do sítio-ativo e no tipo de reação catalisada. As enzimas da primeira família catalisam tipicamente a hidroxilação dos centros de carbono; as enzimas da segunda família catalisam a transferência de um átomo de oxigênio para um par de elétrons no substrato; e a terceira família, diversifica em estrutura e função, é constituída por membros com dois equivalentes do cofator pterina ligados ao metal (HILLE, 2002).

Em *Eco*, o principal cofator em molibdo proteínas é o dinucleotídeo molibdoterina guanina (MAUPIN-FURLOW *et al.*, 1995), requerido para a formação e

a função de múltiplas enzimas incluindo a nitrato redutase, a formato desidrogenase, a dimetil-sulfóxido redutase, a trimetilamina-N-óxido redutase, e a biotina-sulfóxido redutase (RAJAGOPALAN e JOHNSON, 1992). Essas enzimas, exceto pela biotina-sulfóxido redutase, são sintetizadas durante o crescimento anaeróbico de *Eco* e outras bactérias entéricas. Elas participam na respiração anaeróbica e colaboram na geração de energia celular (MAUPIN-FURLON *et al.*, 1995; RECH *et al.*, 1995; JOHANN e HINTON, 1987). Em *Eco*, o molibdato é captado para dentro da célula, reduzido, e incorporado ao dinucleotídeo molibdoterina guanina, o qual é requerido para o correto enovelamento e a função das enzimas (MAUPIN-FURLON *et al.*, 1995; RECH *et al.*, 1995; JOHANN e HINTON, 1987).

A produção das molibdoenzimas em *Eco* requer a captação de molibdato (GRUNDEN *et al.*, 1999). Em *Eco* o *locus moa* é requerido para a biossíntese de molibdopterina, enquanto que o óperon *modABCD* codifica o sistema tipo ABC de captação de molibdato com alta afinidade (RECH *et al.*, 1995). Linhagens defectivas neste *locus* são incapazes de sintetizar as molibdoenzimas, mas a suplementação do meio de cultura com altos níveis de molibdato restaura a atividade dessas enzimas, o que sugere a existência de outras vias alternativas de captação do molibdato (KUPER *et al.*, 2003).

O óperon *mod* de transporte de molibdato (*modABCD*) foi seqüenciado e caracterizado em *Eco* (MAUPIN-FURLON *et al.*, 1995; RECH *et al.*, 1995; JOHANN e HINTON, 1987). O óperon codifica um sistema de transporte tipo ABC semelhante ao encontrado nos transportadores de maltose, histidina e leucina-isoleucina (GILSON *et al.*, 1982; HIGGINS *et al.*, 1992; AMES, 1986; SHUMAN, 1987). A análise das proteínas codificadas pelo óperon *modABCD* sugere que a ModA atuaria como proteína periplasmática ligadora específica de molibdato; ModB como a proteína integral de membrana, formadora do canal ou poro na membrana; e ModC, como proteína energizante ligadora de ATP (RECH *et al.*, 1996).

Por analogia com outras proteínas periplasmáticas ligadoras e mais recentemente pela resolução de diversas estruturas ligadas ao substrato (HOLLENSTEIN *et al.*, 2007; HU *et al.*, 1997; LAWSON *et al.*, 1998; SANTACRUZ *et al.*, 2006), sabe-se que a proteína ModA liga molibdato e tungstênio e os transferem para a proteína ModB na superfície da membrana citoplasmática. ModB em conjunto com a proteína ModC, transporta-os através da membrana citoplasmática até o citoplasma celular, e ModC hidrolisa o ATP para o fornecimento de energia ao

processo de transporte (RECH *et al.*, 1995). A dependência por ATP é deduzida pela similaridade na sequência do gene *modC* com outros genes bacterianos transportadores ABC, incluindo *hisP* e *malk* (Ames, 1986) e pela estrutura resolvida do operon *mod* em *Archaeoglobus fulgidus* (*Afu*) (HOLLENSTEIN *et al.*, 2007).

Em *Eco*, além do óperon *mod*, encontra-se em orientação oposta outro óperon que codifica para as proteínas ModE e ModF. ModE regula a expressão do óperon *modABC* pela ligação na região operadora do mesmo, mas a função da proteína ModF ainda não foi estabelecida, uma vez que a deleção do gene *modF* não apresentou efeito detectável no metabolismo de molibdato em *Eco* (SELF *et al.*, 2001). É possível que a proteína ModF esteja relacionada com a exportação de molibdato (SELF *et al.*, 2001)(Figura 2).

Figura 2: Transporte e regulação de molibdato em *Eco*. O transportador de molibdato de alta afinidade está representado pelas proteínas ModABC. Um transportador ABC para sulfatos (CysUWA e SBP, ou proteína ligadora de sulfato), que também transporta outros anions, contribui para o transporte de molibdênio. ModE-Mo serve como repressor do operon *modABC* e aumenta a expressão dos óperons *hyc* e *narXL*. Figura adaptada de Self *et al.*, 2001.

Em *Eco*, o primeiro gene do óperon *mod*, *modA*, se estende por 774 bp e codifica para uma proteína de 257 aminoácidos (SELF *et al.*, 2001) (Figura 2). A sequência de aminoácidos do extremo N-terminal da ModA é similar às sequências sinal encontradas em proteínas ligadoras periplasmáticas solúveis. Assim, a proteína

ModA é exportada para o periplasma e a pré-proteína clivada entre os aminoácidos alanina e aspartato, deixando uma proteína madura de 233 aminoácidos (MAUPIN-FURLOW *et al.*, 1995). A proteína ModA é extremamente específica e seletiva e pode discriminar entre ligantes com tamanhos similares, formas e propriedades de ligação com o hidrogênio (LAWSON *et al.*, 1998). ModA liga molibdato ou tungstato em quantidades equimolares com um K_d de 20 ± 8 nM (RECH *et al.*, 1996; IMPERIAL *et al.*, 1998; SELF *et al.*, 2001).

A estrutura cristalina desta proteína foi resolvida por Hu e colaboradores (1997), e revela a presença de dois domínios globulares, bem separados e conectados por uma região central (Figura 3). Os dois domínios juntos criam uma profunda fenda onde o ligante se ancora. Na ausência deste, a fenda se abre e a abertura fica acessível ao solvente (SELF *et al.*, 2001).

Figura 3: Estrutura tridimensional da proteína ModA de *Eco* definida experimentalmente por cristalografia. fonte(HU *et al.*, 1997).

Tanto molibdato como tungstato se ligam na proteína ModA como um complexo tetraédrico mantido por sete pontes de hidrogênio formadas entre o oxigênio do ânion ligado e os grupos de aminoácidos dos dois domínios (HU *et al.*, 1997). As estruturas cristalinas de ModA-MoO₄ e ModA-WO₄ de *Eco* como também

de ModA-MoO₄ de *Azotobacter vinelandii* (Avi), mostram que a estrutura terciária da forma ligada de ModA é muito similar a outras proteínas ligadoras periplasmáticas, incluindo a proteína ligadora de sulfato, SBP (SELF *et al.*, 2001).

1.2 Captação de Oligopeptídeos

Peptídeos representam um componente importante na nutrição da maioria dos seres vivos, em particular das bactérias. Em bactérias, peptídeos também estão envolvidos com a sinalização intra e intercelular, desenvolvimento de estados fisiológicos específicos, como a esporulação, e a síntese da parede celular (PAYNE e GILBARG, 1968; HILLES e HIGGINS, 1986; FUQUA *et al.*, 1994). A capacidade de captar peptídeos torna-se, portanto, uma propriedade essencial para a sobrevivência das bactérias em diferentes ambientes, bióticos e abióticos (PAYNE e GILBARG, 1968; HILLES e HIGGINS, 1986). Em bactérias, a captação de peptídeos é feita por meio de três sistemas de transporte principais que foram detalhadamente estudados em *Salmonella typhimurium* (*Stm*) e *Eco* (GUYER *et al.*, 1985; HILLES e HIGGINS, 1986). Nessas bactérias três sistemas de transporte do tipo ABC se encarregam da captação de peptídeos, a saber: (i) um sistema específico de captação de dipeptídeos, denominado dipeptide permease (Dpp); (ii) um sistema de captação de tripeptídeos, denominado tripeptide permease (Tpp); (iii) e um sistema de captação de oligopeptídeos, ou oligopeptídeo permease (Opp).

O sistema Dpp é específico para dipeptídeos e, geralmente, está envolvido com o comportamento quimiotático de bactérias a peptídeos (MANSON *et al.* 1986). O sistema Tpp possui grande afinidade e especificidade para tripeptídeos hidrofóbicos e é induzido apenas durante o crescimento em anaerobiose (GIBSON *et al.*, 1984). O terceiro sistema de captação de peptídeos em bactérias, o Opp, reconhece peptídeos formados por três ou mais aminoácidos. O sistema Opp de bactérias se caracteriza pela grande afinidade, mas baixa especificidade em relação ao tamanho e à sequência do peptídeo transportado (HILLES e HIGGINS, 1986; GOODELL e HIGGINS, 1987).

O sistema Opp representa o principal sistema de transporte de peptídeos nas enterobactérias e, além do papel na nutrição, está envolvido com a reciclagem de

peptídeos da parede celular (GOODELL e HIGGINS, 1987). Em *Stm*, o sistema Opp é codificado por 5 genes *opp*, organizados na forma de um operon e localizados na região correspondente aos 37 minutos do mapa genético (HIGGINS *et al.*, 1983; HILLES e HIGGINS, 1986; YOUNG e HOLLAND, 1999) (Figura 4). O componente DIM do sistema Opp de *Stm* é responsável pela definição do poro na membrana citoplasmática e corresponde aos produtos dos genes *oppB* e *oppC* (HILLES e HIGGINS, 1986). O segundo componente do sistema Opp de *Stm*, os DLNs são codificados pelos genes *oppD* e *oppF* (YOUNG e HOLLAND, 1999). Por fim, o componente LS, codificado pelo primeiro cístron do operon *opp* (*oppA*), é denominado OppA, proteína encarregada de ligar o peptídeo e levá-lo ao poro formado pelos outros elementos DIM do sistema de captação. Em *Xanthomonas axonopodis* pv *citri* (*Xac*), o sistema Opp é codificado por 4 genes em um único transcrito. Os componentes DLN encontram-se fusionados e denominados *oppDF*, os componentes DIM compostos por dois genes *oppB* e *oppC* e um último gene, *oppA*, define o componente LS (MOUTRAN *et al.*, 2004).

Figura 4: Organização genética do operon *opp* de *Stm* e *Xac*. A organização sequencial dos cístrons constituintes, assim como os tamanhos em número de aminoácidos dos peptídeos codificados, estão indicados na figura.

As proteínas ligadoras de oligopeptídeos (OppA) expressas por diferentes espécies de bactérias possuem tamanho que variam de 55 e 70 KDa e possuem grande afinidade mas baixa especificidade a substratos peptídicos (GUYER *et al.*, 1985; HILLES e HIGGINS, 1986). A proteína OppA geralmente representa uma das mais abundantes no periplasma de enterobactérias (HILLES e HIGGINS, 1986).

Diferenças na especificidade entre as OppAs de *Eco*, *Stm* e *Lactococcus lactis* (*Lla*) foram descritas. A OppA de *Eco* liga-se a di e tripeptídeos e, com menor afinidade, pentapeptídeos e tetrapeptídeos. A OppA de *Lla* tem uma maior afinidade por penta e hexapeptídeos e é incapaz de captar di e tripeptídeos, sendo esses captados pelo sistema de transporte que utiliza a energia gerada pela força proton-motora (TYNKKYNEN *et al.*, 1993; JUILLARD *et al.*, 1998). Em *Lla*, a translocação de peptídeos depende da carga e peso molecular, sendo que peptídeos hidrofóbicos possuem uma maior preferência em relação aos peptídeos hidrofílicos (JUILLARD *et al.*, 1998). Estudos de captação com a OppA de *Lla* e nonapeptídeos indicaram que os seis primeiros resíduos do peptídeo ficam imersos na cavidade formada pelos domínios, enquanto que a porção restante interage com a superfície externa da proteína ligadora (DETMERS *et al.*, 2001).

Em *Streptococcus uberis* (*Sub*), foram testados diversos ligantes peptídicos quanto à captação mediada pela OppA1 assim como o parólogo OppA2. Em todos os casos a perda da OppA2 não alterou o fenótipo, indicando que a OppA1 pode utilizar peptídeos variando entre 3 a 8 aminoácidos. Cepas parentais mostram-se incapazes de transportar di-, nona- e decapeptídeos sugerindo que *Sub* não utiliza tais peptídeos (TAYLOR *et al.*, 2003).

Em *Borrelia burgdorferie* (*Bbu*), assim como em outras espécies do gênero, foram encontradas 5 cópias do gene responsável pela síntese do componente LS que diferem na habilidade de reconhecer peptídeos (WANG *et al.*, 2005). As OppA-1, OppA-2 e OppA-3 reconhecem diferentes heptapeptídeos enquanto que nenhuma afinidade foi encontrada para as OppA-4 e OppA5. A OppA-1 mostrou sua máxima atividade em uma faixa de pH variando de 5.5 até 7.5 e uma maior amplitude de afinidade por substratos (maior afinidade por FGTRWFN, SGTKFFN e FTTNQQD), enquanto as OppA-2 (maior afinidade por HSMPTVL) e OppA-3 (maior afinidade por AVPQSAL e LTPPTSL) mostraram-se ativas em pHs abaixo de 5.5 e uma maior afinidade por aminoácidos pequenos apolares. As OppA-1 e OppA-2 possuem maior preferência por tripeptídeos do que dipeptídeos. Diferenças nas afinidades entre os diversos parólogos em *Bbu*, sugerem que possuam distintas funções no organismo (WANG *et al.*, 2005).

A compreensão das propriedades funcionais das proteínas OppA de bactérias gram-negativas foi facilitada com a definição da estrutura da proteína expressa por *Stm* (TAME *et al.*, 1994). Ao todo foram identificados três domínios estruturais na

proteína OppA de *Stm*: o domínio I que compreende os aminoácidos 1 a 44, 169 a 270 e 487 a 517 da proteína madura, o domínio II que compreende os aminoácidos 45 a 168, e o domínio III correspondente aos aminoácidos 271 a 486 da proteína madura (Figura 5). De forma semelhante a outros componentes LS de sistemas ABC de transporte, a OppA de *Stm* possui dois grandes domínios estruturais que circundariam o substrato a ser direcionado ao complexo transportador associado à membrana. Esses dois domínios, I e III, são ligados por dois segmentos que permitiriam um movimento de abertura e fechamento da proteína ao redor do substrato. Por outro lado, o domínio II não possui função conhecida e não possui análogos estruturais em outros componentes LS conhecidos (TAME *et al.*, 1994).

Figura 5: Estrutura tridimensional da proteína OppA de *Stm*. Estrutura determinada experimentalmente por cristalografia (fonte TAME *et al.*, 1994).

O ligante peptídico fica completamente embebido em cavidades da proteína OppA de *Stm* e sua porção N-terminal forma ligações salinas com as cadeias laterais de aminoácidos presentes no domínio I (aminoácidos 32 a 34) e III (aminoácidos 413 a 419). As cadeias laterais dos peptídeos ficam acomodadas em bolsas hidratadas onde ocorrem poucos contatos diretos com a proteína. A presença de moléculas de água nessas bolsas satisfazem os requerimentos de pontes de hidrogênio para a ligação entre o ligante e a proteína OppA e ocultaria cargas

eventualmente presentes no ligante, o que explica a baixa especificidade do sistema já que peptídeos de composição diversa podem se acomodar na proteína (TAME *et al.*, 1994).

Em *Stm*, alguns aminoácidos da proteína OppA são importantes na acomodação do ligante. Entre eles os aminoácidos Arg⁴¹³ e His³⁷¹, formadores de pontes salinas com o grupo carboxílico de tri e tetrapeptídeos, e o aminoácido Lis³⁰⁷ para pentapeptídeos. Os resíduos Asn⁴¹⁹ e Arg⁴¹³ formam ligações salinas com as porções N-terminal e C-terminal do ligante, respectivamente. Duas cadeias laterais são importantes para a conformação das bolsas que envolvem o ligante. A primeira cadeia lateral é composta pelos aminoácidos Val³⁴ e Cis^{271 e 417} enquanto que a segunda é composta pelos aminoácidos Trp^{397 e 416} e Leu⁴⁰¹ (TAME *et al.*, 1994).

Em *Bacillus subtilis* (*Bsu*), foram identificados dois lóbulos estruturais na proteína AppA: o lóbulo 1 compreende os aminoácidos 17 a 264 e 491 a 520 (domínio I) e o lóbulo 2 compreende os aminoácidos 265 a 490 (domínio III). O lóbulo 1 ainda é dividido em duas subestruturas. A primeira correspondente aos aminoácidos 17 a 53, 184 a 264 e 491 a 520 e possui uma organização estrutural do tipo α/β na qual sete fitas centrais são rodeadas por α -hélices. A segunda subestrutura consiste os resíduos 54 a 183 e tem como base uma estrutura com quatro fitas em folha β -pregueada e dois “ β -hairpin” (Figura 6). O lóbulo 2 consiste em um domínio simples do tipo α/β com uma estrutura em folha β -pregueada ao centro contendo seis fitas, flanqueada por α -hélices e alças. De forma semelhante a outros componentes LS de sistemas ABC de transporte, a AppA de *Bsu* possui uma estrutura que circundaria o substrato a ser direcionado ao complexo transportador associado à membrana. Essa estrutura é ligada por dois segmentos que permitiriam um movimento de abertura e fechamento da proteína ao redor do substrato (LEVDIKOV *et al.*, 2004).

Figura 6: Estrutura da AppA de *Bsu* representada na forma de diagrama de fitas. Domínios, domínios I, II e III estão coloridas em azul escuro, azul claro e verde, respectivamente. As extremidades N- e C- terminais estão marcadas. O peptídeo está colorido em vermelho (fonte LEVDIKOV *et al.*, 2004).

Em *Bsu* e *Stm*, o ligante peptídico fica completamente embebido em cavidades da proteína AppA e sua porção C- e N-terminais formam ligações de hidrogênio com as cadeias principais da proteína assim como com as cadeias laterais ou ocasionalmente com moléculas de água. A presença de moléculas de água nessas bolsas satisfazem os requerimentos de pontes de hidrogênio para a ligação entre o ligante e a proteína OppA e ocultaria cargas eventualmente presentes no ligante, o que explica a baixa especificidade do sistema já que peptídeos de composição diversa podem se acomodar na proteína (TAME *et al.*, 1994).

Alguns aminoácidos da proteína AppA são importantes na acomodação do ligante. Entre eles o aminoácido Arg³⁷³ fornece carga positiva, formando um par de íons para balancear o grupo carboxílico da cadeia principal do peptídeo, que também forma uma interação dipolo-carregada com o aminoácido Tyr⁴⁸⁷. O resíduo Arg³⁷³ exerce uma outra função importante ao interagir com os últimos quatro resíduos C-terminais do peptídeo transportado (LEVDIKOV *et al.*, 2004).

1.3 Doenças Fitopatológicas e seu impacto econômico

Desde o começo da civilização, as doenças fitopatológicas têm merecido especial atenção devido à sua importância econômica, pois causam grandes prejuízos à produção agrícola. As plantas são hospedeiras de um grande número de fitopatógenos que incluem fungos, bactérias, vírus e parasitas. O aumento da população humana e o decréscimo da área agrícola cultivável tornam cada vez mais crítico o suprimento alimentício mundial. O entendimento dos mecanismos de defesa da planta em conjunto com o estudo do processo de infecção, podem levar a uma melhor compreensão das inter-relações planta-patógeno e ao desenvolvimento de mecanismos de proteção mais efetivos contra os mesmos.

A citricultura é de fundamental importância na atividade agrícola brasileira. O Brasil é o principal produtor e o maior exportador de suco de laranja (*Citrus spp.*) (ESTANISLAU *et al.*, 2001), respondendo por cerca de 80% do comércio mundial, gerando divisas superiores a um bilhão de dólares anuais e movimentando indiretamente cerca de dois bilhões de dólares (NEVES *et al.*, 2001). Apesar do grande potencial produtivo, a cultura de *Citrus* é alvo constante de inúmeras pragas e doenças que causam queda da produtividade e qualidade dos frutos.

O gênero *Xanthomonas* engloba um grupo de bactérias gram-negativas estritamente aeróbicas, móveis, e pertencentes à subdivisão gama das proteobactérias (TWING e SWINGS, 1993). A maioria das cepas produz um pigmento amarelo, a xantomonadina, e as colônias apresentam aspecto mucóide devido à presença de um polímero hidrossolúvel comumente chamado de goma xantana que está envolvida na virulência da bactéria, mas apresenta uma ampla gama de aplicações na indústria de alimentos (COPLIN e COOK, 1990; POPLAWSKY *et al.*, 2000). Bactérias do gênero *Xanthomonas* encontram-se principalmente no solo, mas também podem comportar-se como patógenos oportunistas para animais e plantas. Seu metabolismo é sempre respiratório, ou aeróbio (a maioria utiliza O₂ como aceptor de elétrons) e podem utilizar substratos variados como fonte de carbono, o que demonstra grande versatilidade metabólica. O gênero compreende inúmeras espécies identificadas por características bioquímicas, fisiológicas, morfológicas e fitopatogênicas. No entanto, utiliza-se com frequência um esquema de nomenclatura baseado em patovares (pv), que definem

propriedades de fitopatogenicidade. A espécie mais conhecida e estudada é *Xanthomonas campestris* com aproximadamente 141 patovares (CHAN e GOODWIN, 1998). Através de uma classificação baseada em grupos de homologia de DNA, Vauterin e colaboradores (1996) propuseram uma nova divisão composta por 20 grupos. O grupo de *X. campestris*, sendo o maior e mais diversificado, foi dividido em 16 sub-grupos, onde o grupo 9, que inclui *X. axonopodis*, é o mais heterogêneo e contém a maioria dos patovares de *X. campestris*.

A *X. campestris* pv. *campestris* (*Xcc*) causa a putrefação negra, doença que afeta crucíferas, como as do gênero *Brassica*, e *Arabidopsis* (CHAN e GOODWIN, 1998). A pústula bacteriana é uma das principais doenças que afetam o pimentão em todo o mundo e pode estar associada a *Xanthomonas axonopodis* pv. *vesicatoria* (*Xav*), *Xanthomonas campestris vesicatoria* (*Xcv*) e *X. Gardneri*. Já *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) é responsável por prejuízos em cultivos de arroz no mundo todo.

A bactéria *Xanthomonas axonopodis* pv. *citri* (*Xac*) ataca laranjeiras, e outras plantas cítricas, e causa o cancro cítrico, doença responsável por perdas anuais estimadas em dezenas de milhões de dólares aos citricultores paulistas.

1.4 O Projeto Genoma de *Xac* e *Xcc*

Na última década a Biologia, em especial a área de Biologia Molecular, passou por uma grande mudança com o advento de projetos de sequenciamento de genomas de diversos organismos como bactérias, arqueobactérias e eucariotos. O interesse na determinação da sequência completa do genoma de um organismo baseia-se na possibilidade de identificação de genes que conferem características únicas aos mesmos e que tenham importância econômica, científica ou social (FERREIRA, 2000). Pela primeira vez, tornou-se possível a comparação do conjunto completo de proteínas presentes na célula de um determinado organismo e a identificação de proteínas responsáveis por muitas funções celulares. As perspectivas geradas com a elucidação do conteúdo genético completo de bactérias são grandes e busca-se, a partir destas sequências, o aproveitamento ou o melhoramento dos produtos derivados do metabolismo microbiano.

A análise computacional de genomas procarióticos completos mostrou, até o momento, que as proteínas microbianas são, em geral, bem conservadas (~ 70% delas contêm regiões conservadas) (KOONIN *et al.*, 1998). Esse fato permitiu delinear famílias de ortólogos em uma ampla extensão filogenética, e em muitos casos, propiciou predição de funções com considerável precisão.

Os genomas das espécies *Xac* e *Xcc* foram elucidados como parte do programa ONSA da FAPESP (da SILVA *et al.*, 2002) em parceria com o Fundo de Defesa da Citricultura de São Paulo (FUNDECITRUS). Entretanto, aproximadamente 20-40% dos genes identificados em cada um dos genomas bacterianos seqüenciados até o momento têm função desconhecida (da SILVA *et al.*, 2002). Sendo assim, a informação gerada no projeto ainda é insuficiente para a compreensão das bases moleculares da patogenicidade destas bactérias. Atualmente (10/01/2009) estão disponíveis sete genomas completos para o gênero *Xanthomonas*: *Xac* 306, *Xcc* 8004, *Xcc* ATCC33913, *Xcv* 85-10, *Xoo* KACC10331, *Xoo* PXO99A e *Xoo* MAFF311018, além de dois genomas incompletos, sendo de *Xoo* BLS256 e *Xanthomonas axonopodis* pv. *glycines* (*Xag*).

O genoma de *Xac* é constituído por um DNA cromossômico circular com 5.175.554 pb e dois plasmídeos com 33.699 e 64.920 pb. Por outro lado o genoma da cepa de *Xcc* ATCC33913 é constituído por apenas um DNA cromossômico circular com 5.076.187 pb que codifica para 4.182 genes. Embora filogeneticamente relacionadas, a análise dos genomas das duas espécies de *Xanthomonas* revelou que muitos genes são exclusivos para cada espécie. O genoma de *Xoo* é constituído por um DNA cromossômico circular com 4.941.439 pb, que codifica para 4.364 genes. O genoma de *Xcv* é constituído por um DNA cromossômico circular com 5.178.466 pb, e quatro plasmídeos com 182.572, 38.116, 19.146 e 1.852 pb que definem 4.855 genes. A disponibilidade dos genomas assim como a noção de que muitos genes são exclusivos de cada espécie abre perspectivas para estudos que visem compreender as diferenças genéticas e fisiológicas que respondam pelas características de virulência e especificidade de hospedeiro para as espécies do gênero *Xanthomonas* (da SILVA *et al.*, 2002).

A fitopatogenicidade de *Xac* é um fenômeno complexo e pouco estudado. Resultados obtidos com outras espécies do gênero revelam que aproximadamente 50 a 100 genes estão direta ou indiretamente relacionados com a capacidade da bactéria causar doenças em plantas. Esses genes codificam, em sua maioria,

proteínas que se localizam no envoltório celular, ou são secretadas para o meio extracelular. O envoltório representa, portanto, um dos componentes celulares mais importantes envolvidos na patogenicidade e na evasão de sistemas de defesa do hospedeiro (da SILVA *et al.*, 2002).

Nos últimos 20 anos, foram feitos avanços significativos na descrição de alguns componentes do envoltório celular de bactérias do gênero *Xanthomonas*, porém, pouco ou nada se sabe sobre as características bioquímicas e fisiológicas desses elementos nessa espécie de bactéria. O sequenciamento do genoma da *Xac* permitiu a identificação de genes responsáveis pela síntese de proteínas do envoltório que poderiam estar relacionadas com diferentes processos celulares, inclusive a patogenicidade.

1.5 Biologia Computacional

A biologia teórica e computacional existe há décadas “à margem” das ciências biológicas. Mas de alguns poucos anos para cá, o fluxo de novos dados biológicos produzidos pelos esforços genômicos e, pela necessidade, a aplicação de computadores à análise dos dados genômicos, começou a afetar todos os aspectos das ciências biológicas. As pesquisas que costumavam ser iniciadas em laboratório começam agora no computador, quando cientistas pesquisam os bancos de dados em busca de informações que possam sugerir novas hipóteses. Os bancos de dados de eletrônicos permitem um acesso mais fácil aos dados armazenados em relação aos registros não eletrônicos. Além da sua utilidade para armazenar, analisar e visualizar os dados, os computadores são dispositivos úteis para entender qualquer sistema que possa ser descrito de forma matemática, dando origem às disciplinas da biologia computacional e, mais recentemente, à bioinformática. A bioinformática é a aplicação da tecnologia de informação ao gerenciamento de dados biológicos, portanto, uma ciência biológica (GIBAS, 2001).

Além de estudar as relações entre as sequências, uma das aplicações mais interessantes dos alinhamentos múltiplos é a descoberta de novas seqüências relacionadas. Essa análise baseada no perfil ou no padrão utiliza o conhecimento derivado de alinhamentos para construir e pesquisar padrões de sequências. Destas

análises, podemos extrair a predição da estrutura secundária, considerada o primeiro passo na predição da estrutura espacial de uma proteína (GIBAS, 2001).

Estudos de estruturas de proteínas mostram-se cada vez mais relevantes. Detalhes atômicos de enzimas consideradas chaves permitem uma melhor compreensão do seu funcionamento e propiciam o desenvolvimento de compostos que possam interagir com essas enzimas. A cristalografia é a técnica mais utilizada para a determinação de estruturas tridimensionais de proteínas (BABU *et al.*, 1995), seguida pela Ressonância Magnética (RMN). Porém, a técnica esbarra na dificuldade de obtenção de cristais, passo considerado crucial na determinação de estruturas. Nesse contexto a técnica de modelagem molecular por homologia representa uma ferramenta importante no sentido de predizer a estrutura tridimensional de proteínas.

Dentre os métodos teóricos de obtenção de estruturas listam-se os métodos físicos e empíricos. Os métodos físicos baseiam-se nas interações entre átomos, utilizando as teorias de dinâmicas molecular e minimização de energias entre outras, enquanto que nos métodos empíricos se enquadram a técnica de modelagem por homologia estrutural ou modelagem comparativa (SILVA, 2007; MAKINO, 2000). Nos casos em que a elucidação estrutural não é possível, ou ainda não foi definida, modelos virtuais da proteína podem ser elaborados por comparação da similaridade com as sequências de proteínas homólogas com estruturas resolvidas e depositadas no “Protein Data Bank” (PDB). A qualidade de modelos estruturais de proteínas gerados por modelagem comparativa e a sua aplicabilidade dependem, predominantemente, do grau de similaridade seqüencial entre a proteína com estrutura resolvida (proteína-molde) e a proteína a qual se deseja modelar (proteína-alvo). O princípio do método baseia-se no fato de que haverá pouca alteração na estrutura tridimensional de uma proteína quando sua sequência primária é levemente alterada (SALI e BLUNDELL, 1993). O método de modelagem molecular por homologia estrutural consiste em quatro passos sequenciais. O primeiro passo é identificar as proteínas com estruturas resolvidas homólogas a sequência alvo. O segundo consiste em alinhar a sequência alvo com a(s) sequência(s) da(s) proteína(s) com estrutura(s) resolvida(s) experimentalmente. O terceiro passo envolve a construção de modelos tridimensionais para a sequência alvo, fornecendo para o programa o alinhamento e arquivo de coordenadas da proteína com estrutura

resolvida. O quarto e último passo consiste na avaliação dos modelos gerados utilizando-se uma variedade de critérios.

A maior diferença entre os métodos de modelagem por homologia envolve o cálculo do modelo tridimensional a partir de um dado alinhamento. O método de modelagem mais amplamente usado é denominado de corpo rígido. Esse método permite a construção do modelo a partir de corpos rígidos que incluem regiões centrais, alças (loops) e cadeias laterais, todos obtidos de estruturas relacionadas. Outra família de métodos de modelagem utiliza ajustes de segmentos em que posições aproximadas para átomos conservados são atribuídas a partir do cálculo das coordenadas de outros átomos no molde. Isso é alcançado pelo uso de um banco de dados de pequenos fragmentos de estrutura de proteínas, regras de energia ou geométrica, ou combinações desses critérios. O terceiro grupo de métodos envolve modelagem por satisfação de restrições espaciais, que usa técnicas de otimização para satisfazer as restrições espaciais obtidas do alinhamento da sequência alvo com os moldes homólogos de estrutura conhecida (SÁNCHEZ, 2000; LASKOWSKI, 2003).

Uma abordagem automatizada para a modelagem molecular por homologia, baseada em restrições espaciais foi implementado pelo programa Modeller, que calcula modelos de proteínas impondo restrições espaciais obtidas das estruturas homólogas (SALI e BLUNDELL, 1993). Este programa foi avaliado com sucesso nos diversos encontros de “Critical Assessment of Techniques for Protein Structure Prediction” (CASP) que, desde 1994, ocorre com o objetivo de avaliar métodos de predição de estrutura de proteínas. A modelagem por homologia realizada pelo programa Modeller inicia-se com o alinhamento da sequência alvo com a sequência da proteína que possui estrutura determinada. O arquivo de saída, obtido sem nenhuma intervenção do usuário, contém os modelos tridimensionais com todos os átomos das cadeias principal e lateral. A etapa final consiste na avaliação dos modelos gerados. Desta forma, o modelo por homologia auxilia na investigação de fatores relevantes para a interação entre inibidor/enzima, bem como para o entendimento dos mecanismos de reação e ação molecular de enzimas.

No nosso trabalho, procuramos caracterizar dois componentes ligantes do sistema de transporte ABC (OppA e ModA) de *Xac* por meio de ferramentas de bioinformática. Os resultados obtidos representam contribuições significativas ao

conhecimento da biologia molecular e estrutural de componentes do envoltório celular desses importantes fitopatógenos.

2 OBJETIVOS

A presente tese teve como principal objetivo fornecer informações sobre organização genética, similaridade sequencial e estrutura espacial das proteínas ModA e OppA de *Xac*. A partir de modelos estruturais gerados com programas específicos elucidamos características relacionadas à interação com os respectivos ligantes e geramos dados que auxiliam na definição experimental das respectivas estruturas moleculares por meio de técnicas cristalográficas. Para isso algumas etapas do trabalho foram cumpridas:

- a) análises comparativas das organizações genéticas dos óperons *mod* e *opp* em diferentes espécies bacterianas;
- b) alinhamento sequencial e estrutural de ortólogos de ModA ou OppA;
- c) modelagem por satisfação das restrições espaciais e validação das estruturas obtidas das duas proteínas alvos;
- d) análise das interações moleculares entre as proteínas OppA e ModA com os seus respectivos ligantes;
- e) determinação da afinidade da proteína OppA de *Xac* a seus respectivos ligantes.

3 MATERIAIS E MÉTODOS

3.1 Análise Computacional

Encontram-se disponíveis no banco de dados do National Center of Bioinformatics (NCBI) (<http://www.ncbi.nlm.nih.gov/>) as sequências correspondente aos aminoácidos e nucleotídeos de genes da *oppA* (gi|21241626|ref|NP_641208.1|) e *modA* (gi|21244083|ref|NP_643665.1|) de *Xac*. Determinamos a sequência do peptídeo sinal com os programas SignalP (BENDTSEN *et al.*, 2004) e DAS (CSERZO *et al.*, 1997). Todos os parâmetros para a proteína OppA e ModA de *Xac* foram calculados com a aplicação dos programas disponíveis no servidor do ExPASy Molecular Biology (<http://www.expasy.org/>). Utilizando a ferramenta de predição CDART (Conserved Domain Architecture Retrieval Tool) disponível no NCBI, analisamos os prováveis domínios conservados. Foi utilizado o código de 3 letras para os nomes das linhagens utilizadas nos alinhamentos segundo a classificação do banco de dados do Kegg (<http://www.genome.ad.jp>).

3.2 Processo de Modelagem

O Conceito de modelagem por homologia baseia-se na conservação do enovelamento tridimensional entre membros de uma mesma família de proteínas (BLUNDELL *et al.*, 1987; GREER, 1990, SANCHEZ e SALI, 1997).

O programa MODELLER (SALI e BLUNDELL, 1993) é frequentemente utilizado em modelagem estrutural tridimensional de proteínas por homologia. Neste programa o usuário fornece o alinhamento da sequência a ser modelada contra estruturas relacionadas e um modelo atômico será automaticamente calculado.

O MODELLER modela a estrutura 3D da proteína por satisfação de restrições espaciais. A princípio, as restrições podem vir de diferentes tipos de fontes como: estruturas homólogas, regras de empacotamento da estrutura secundária, mutagênese sítio dirigida, potencial de campo de força médio de átomo-átomo e resíduo-resíduo, etc. O MODELLER consegue obter automaticamente as restrições espaciais apenas de estruturas conhecidas homólogas sendo as outras restrições fornecidas pelo usuário.

3.2.1 Seleção das Estruturas Moldes

Para selecionar as melhores estruturas moldes utilizamos o programa BLASTp contra o Banco de Dados de Proteínas (PDB, <http://www.rcsb.org/pdb/>). As estruturas com definição por difração de raio-X utilizadas para a predição da estrutura tridimensional da proteína OppA de *Xac* foram a OppA de *Stm* (código PDB: 1jev) (TAME *et al.*, 1996) e a AppA de *Bsu* (código PDB: 1xoc) (LEVDIKOV *et al.*, 2005). Para a predição da estrutura tridimensional da proteína ModA de *Xac* foi utilizada a estrutura da ModA de *Eco* (código PDB:1amf) (HU *et al.*, 1997).

3.2.2 Alinhamentos

Os alinhamentos seqüenciais dos respectivos ortólogos das OppAs e ModAs, e os alinhamentos utilizados para a modelagem da OppA e ModA de *Xac*, foram obtidos com o programa ClustalX (THOMPSON *et al.*, 1997). Edições dos alinhamentos foram realizadas no programa Genedoc (NICHOLAS *et al.*, 1997). Para a modelagem, quando necessário, introduzimos “GAPs” nas regiões de alças respeitando o posicionamento das estruturas secundárias previstas pelo programa JPRED – A Secondary Structure Prediction Server” (CUFF *et al.*, 1998). Para os alinhamentos estruturais e super-imposição de cadeias principais dos ortólogos respectivos de OppA e ModA, utilizamos o método CE (combinatorial extension) (SHINDYALOV e BOURNE, 1998).

3.2.3 Processo de Modelagem Estrutural da OppA e ModA de *Xac*

Utilizando as coordenadas das proteínas 1Xoc e 1Jev como moldes para a modelagem da OppA e 1amf_A como molde para a modelagem da ModA, empregamos a versão 6v2 do programa Modeller (SALI e BLUNDELL, 1993). Foram gerados 100 modelos das estruturas tridimensionais de 2 proteínas estudada de *Xac*. Diversas metodologias existem para checagem e avaliação da qualidade de uma estrutura, seja ela determinada experimentalmente ou por modelagem por homologia. A princípio os modelos foram selecionados pelos menores valores obtidos nas funções objetivas do modeller (“MODELLER OBJECTIVE FUNCTION”) (SALI e BLUNDELL, 1993).

Posteriormente utilizamos o programa PROCHECK (LASKOWSKI *et al.*, 1993) que analisa a estereoquímica de cada resíduo individualmente quanto a estereoquímica global da estrutura tendo como parâmetros informações derivadas de estruturas resolvidas a alta resolução (MORRIS *et al.*, 1992) que constituem sua base de dados. As informações estereoquímicas checadas pelo programa são: ligações covalentes,

quiralidade, interações não-covalente, ligações de hidrogênio da cadeia principal e pontes de dissulfeto. Dentre as várias análises do PROCHECK estão: análise da estrutura secundária, acessibilidade estimada de cada resíduo, gráfico de Ramachandran etc. O diagrama de Ramachandran (RAMACHANDRAN e SASISEKHARAN, 1968), indica a distribuição dos ângulos Φ e Ψ dos resíduos pertencentes a uma estrutura. Seus indicadores de qualidade baseiam-se na análise de estruturas com resolução de até 2 Å. Nos seus critérios uma estrutura de boa qualidade deve ter 90% ou mais dos seus resíduos nas regiões mais favoráveis. Dados sobre as interações entre resíduos, visualizações e comparações com outras proteínas ligadoras de substratos, foram feitas com o pacote de ferramentas Diamond Sting Suite Package “Java Contact Table” (DMS 3.0) (NESHICH *et al.*, 2005), juntamente com o programa Pymol (DELANO, 2002).

3.3 Dinâmica Molecular do melhor modelo da estrutura da ModA de *Xac*

A proteína ModA modelada a partir das coordenadas estruturais do ortólogo de *E. coli* foi submetida a simulações de dinâmica molecular através do programa Gromacs 3.2 (VAN DER SPOEL *et al.*, 2005) para que houvesse um refinamento do nosso melhor modelo. Este método consiste em resolver as equações de Newton acopladas para todas as partículas do sistema. O primeiro passo para iniciar uma minimização foi inserir o arquivo modelado em “pdb”. O complexo foi solvatado considerando-o em uma caixa octaédrica preenchida por água e os aminoácidos ionizados, totalizando 21.650 moléculas que foram protonadas a um pH 7,0. Para neutralizar o sistema dois íons de Na^+ foram adicionados. O sistema foi submetido a 5000 passos para minimização das energias e remoção de contatos inapropriados de van der Waals. Para um relaxamento estrutural adicional, realizamos uma dinâmica molecular durante 500 ps em uma temperatura de 312 Kelvin e uma pressão de 1 atm.

3.4 Ancoragem do íon molibdênio da ModA de *Xac*

A determinação do sítio de ancoragem do molibdato na proteína ModA de *Xac* foi feita pelo cruzamento dos dados estruturais entre 1amf, 1atg e o nosso modelo com o programa Sting (NESHICH *et al.*, 2005). Extraíndo os dados estruturais físico-químicos dos resíduos responsáveis pela ancoragem do molibdato nas estruturas 1amf e 1atg. Em seguida, observamos quais as características que se mantiveram constantes ou próximas em ambas as estruturas. A energia relativa de conservação calculada pelo Sting foi calibrada para selecionar os resíduos que continham energia maior ou igual a 15 kcal juntamente com um potencial eletrostático de superfície maior ou igual a 19 kcal. Aplicamos esses parâmetros à estrutura do nosso modelo e utilizamos o programa Pymol para selecionar os resíduos que estivessem a uma distância física de não mais que 5 Å entre os resíduos da ModA de *Xac* e o elemento químico molibdênio, determinando os resíduos que interagiam com o ligante. Outra estratégia utilizada na ancoragem do molibdato à ModA de *Xac* baseia-se no método manual no qual as coordenadas estruturais do molibdato obtidas no PDB de 1amf foram transferidas para o arquivo do modelo da ModA. O arquivo foi editado para o melhor posicionamento do íon.

3.5 Ancoragem de oligopeptídeos à OppA de *Xac*

Submetemos o modelo da proteína OppA de *Xac* a análises de ancoragem com diferentes oligopeptídeos. As seqüências dos ligantes, foram as mesmas descritas em ensaios funcionais para os parólogos da OppA de *Bbu* com heptapeptídeos (WANG *et al.*, 2005); e OppA de *Lla* (RDMPIQA), (RDMPIQAF) e (SLSQSSLSQSKVLPVPQ) (PICON *et al.*, 2000). Também foram utilizados os dados descritos nas estruturas cristalográficas de AppA de *Bsu* (LEVDIKOV *et al.*, 2005) e tripeptídeos, tetrapetídeos

utilizados na definição da estrutura da OppA de *Stm* (TAME *et al.*, 1996). Para este estudo, utilizamos o programa de ancoragem via provedor GRAMM-X (TOVCHIGRECHKO e VAKSER, 2005). A partir dos dois arquivos no formato “.pdb”, buscou-se de forma exaustiva os 100 prováveis melhores perfis de ancoragens. A seleção dos melhores resultados (denominados eventos positivos) foram baseadas na comparação com dados estruturais existentes para 1jev e 1xoc. Esta seleção foi realizada com o programa STING (NESHICH *et al.*, 2005) e Pymol (DELANO, 2002), que permitem a visualização de interações e o posicionamento dos ligantes dentro da estrutura prevista para a OppA. Foi considerado um evento positivo, o ligante que se encontrava imerso na fenda existente entre os domínios da OppA de *Xac*, com as devidas precauções para que não houvesse a sobreposição (“space crash”) de cadeias laterais e/ou principais dos resíduos.

4 RESULTADOS

4.1 Identificação dos óperons *mod* e *opp* em *Xanthomonas*

Até o presente momento, estão disponíveis no banco de dados do NCBI sete genomas completos para espécies e/ou patovares do gênero *Xanthomonas* (atualizado em 10/01/2009): *Xac* 306, *Xcc* 8004, *Xcc* ATCC33913, *Xcv* 85-10, *Xoo* KACC10331, *Xoo* PXO99A e *Xoo* MAFF311018 e dois genomas incompletos de *Xoo* BLS256 e *Xag*. Buscas por ortólogos dos óperons *opp* e *mod* nos genomas e plasmídeos dessas linhagens revelaram a ausência do óperon *opp* em todos os genomas exceto naquele da linhagem 306 de *Xac* (SILVA *et al.*, 2002; MOUTRAN *et al.*, 2004). Por outro lado, o óperon *mod* foi identificado em todos os genomas completos de *Xanthomonas*. Nenhuma evidência dos óperons *mod* e *opp* foi observada para a linhagem de *Xor* cujo sequenciamento genômico ainda está incompleto.

4.2 Organização dos óperons *mod* e respectivos cistron em diferentes espécies de *Xanthomonas*

O óperon *mod* de *Xac* 306 é constituído por três genes (*modA*, *modB* e *modC*) que apresentam similaridades com os ortólogos previamente identificados em outras espécies de bactérias. A mesma organização do óperon é encontrada em todos os genomas disponíveis de *Xanthomonas* com exceção de *Xoo* em que o sentido de transcrição do óperon encontra-se invertido em relação às outras espécies de *Xanthomonas* (Figura 7). Não foi encontrado no gênero de *Xanthomonas* o gene *modE*, responsável pelo controle da transcrição do óperon *mod* em *Eco*.

O cistron *modA* é constituído por 774 pb em todas as linhagens de *Xanthomonas* e codifica uma única proteína de 258 aminoácidos incluindo um peptídeo sinal com 24 aminoácidos (Figuras 7, 8 e 9A), exceto em *Xoo* na qual o

gene *modA* foi anotado com 807 pb o que corresponderia a uma proteína de 269 aminoácidos (LEE *et al.*, 2005). Entretanto, alinhamentos sequenciais baseados na similaridade entre os genes *modA* de *Xanthomonas* indicam claramente que se trata de um erro de anotação no qual o códon de iniciação foi posicionado 11 resíduos à montante da metionina inicial encontrada nas proteínas codificadas por todas as outras espécies de *Xanthomonas* com genomas determinados (Figura 8).

O cóstron *modB* (690 pb) é responsável pela síntese de uma proteína de 230 aminoácidos em *Xac*, *Xcc* e *Xoo* e 228 aminoácidos (684 pb) em *Xcv* (Figuras 7 e 9A). O terceiro cóstron, *modC*, é formado por 612 pb, codificando uma proteína de 204 aminoácidos em *Xac*, *Xcc*, *Xcv* e 217 aminoácidos (651 pb) em *Xoo* (Figuras 7 e 9A). Tanto à montante quanto à jusante não foram anotados genes relacionados a transposons ou bacteriófagos assim como genes de regulação do próprio sistema de transporte de molibdato.

Figura 7: Organização genética do operon *mod* nas linhagens de *Xac*, *Xcc*, *Xcv* e *Xoo*. Ao lado de cada operon, estão indicadas as posições dos mesmos dentro dos seus respectivos genomas, enquanto que no interior das setas estão indicados os tamanhos de cada cóstron em número de pares de bases. Em vermelho representamos o gene *modA*, amarelo *modB* e azul *modC*.

Os alinhamentos sequenciais dos aminoácidos das proteínas ModA, ModB e ModC das diferentes espécies de *Xanthomonas* revelaram uma elevada conservação entre os ortólogos. As identidades entre as proteínas variaram entre 82% a 94% para ModA, 96% a 98% para ModB e de 88% a 95% para ModC

(Tabela 1, Figura 8). Os alinhamentos mostram ainda uma maior similaridade entre os ortólogos encontrados em *Xac* e *Xcv*, quando consideramos as proteínas ModA e ModC, seguido pelos ortólogos de *Xoo*. As sequências com menor grau de similaridade foram aquelas provenientes de *Xcc* que apresentaram valores como 82% de similaridade para ModA em relação ao ortólogo de *Xoo* (Tabela 1, Figura 8).

Figura 8: Alinhamento sequencial de ModA das linhagens de *Xac*, *Xoo*, *Xcv* e *Xcc*. Em azul estão indicados os aminoácidos provenientes de um erro de anotação no genoma de *Xoo*. Em vermelho está demarcada a região do peptídeo sinal. Em preto e cinza, os alinhamentos de aminoácidos com identidades de 100% e 75%, respectivamente, entre todas as seqüências de *Xanthomonas* analisadas. Os aminoácidos responsáveis pela interação com o molibdato estão representados em amarelo, quando estes formam ligações de hidrogênio, ou em verde, quando apresentam átomos de carbono com outros tipos de interação com o ligante. A numeração e os asteriscos acima das seqüências indicam a posição de cada resíduo na proteína ModA.

Tabela 1: Similaridade entre os ortólogos das proteínas ModA, ModB e ModC presentes nas diversas espécies de *Xanthomonas* com genomas concluídos.

Ortólogos ModA	Similaridade	Ortólogos ModB	Similaridade	Ortólogos ModC	Similaridade
<i>Xoo</i> vs <i>Xcv</i>	91%	<i>Xoo</i> vs <i>Xcv</i>	97%	<i>Xoo</i> vs <i>Xcv</i>	95%
<i>Xoo</i> vs <i>Xac</i>	93%	<i>Xoo</i> vs <i>Xac</i>	98%	<i>Xoo</i> vs <i>Xac</i>	93%
<i>Xoo</i> vs <i>Xcc</i>	82%	<i>Xoo</i> vs <i>Xcc</i>	96%	<i>Xoo</i> vs <i>Xcc</i>	88%
<i>Xcv</i> vs <i>Xac</i>	94%	<i>Xcv</i> vs <i>Xac</i>	98%	<i>Xcv</i> vs <i>Xac</i>	95%
<i>Xcv</i> vs <i>Xcc</i>	85%	<i>Xcv</i> vs <i>Xcc</i>	97%	<i>Xcv</i> vs <i>Xcc</i>	90%
<i>Xcc</i> vs <i>Xac</i>	85%	<i>Xcc</i> vs <i>Xac</i>	98%	<i>Xcc</i> vs <i>Xac</i>	89%

4.3 Análise da organização gênica do óperon *mod* em diferentes espécies bacterianas

A comparação do óperon *mod* de *Xac* com outras espécies de bactérias revelou algumas diferenças interessantes. Em *Mesorhizobium loti* (*Mlo*), *Shigella flexneri* (*Sfl*), *Rhodobacter capsulatus* (*Rca*), *Geobacter sulfurreducens* (*Gsu*) (Figuras 9A, B e C), o óperon *mod* é transcrito na direção oposta. Em *Ralstonia solanacearum* (*Rso*), o gene *modB* é transcrito em direção oposta aos demais cístrons do óperon (Figura 9C). Genes adicionais estão presentes no óperon *mod* de algumas espécies bacterianas quando comparadas aos genes encontrados em *Xanthomonas*. No simbionte *Mlo*, identificamos uma cópia do gene *modE*, responsável pela codificação do repressor do óperon *mod* assim como o gene responsável pela síntese da proteína molibdopterina localizados à montante e à jusante do óperon *modABC*, respectivamente (Figura 9A).

Para o grupo das enterobactérias, além dos genes responsáveis pela síntese de ModE e molibdopterina, está presente uma segunda cópia do gene *modF*, responsável pela síntese do componente energizante do sistema ABC de transporte. Nesse grupo de bactérias, apenas os genes de *Sfl* são transcritos em sentido oposto em relação às demais espécies de enterobactérias (Figura 9B). Em *Gsu*, o óperon *mod* possui os genes *modABC* separados do gene *modE* pelo cistron *modD* com

função desconhecida (Figura 9C). O gene *modD* também foi encontrado em *Rca* (Figura 9C).

Figura 9: Organização genética de óperons mod encontrados em diferentes espécies de bactérias. A) Grupo das bactérias que interagem com plantas: Xac, Xcc, Xoo, Xcv, Mlo e *Sinorhizobium meliloti*, (Sme); B) Grupo das enterobactérias: Eco, Stm, Sfl e Ype; C) Grupo das bactérias de solo: Avi, Gsu, *Streptomyces coelicolor* (Sco), Rso e Rca.

4.3.1 Análise de similaridade da proteína ModA expressa por diferentes espécies de bactérias

Alinhamentos das sequencias de aminoácidos dos ortólogos da ModA encontrados em 14 proteobactérias e 1 actinobactéria (Sco) revelaram alto grau de conservação, incluindo seis regiões onde se encontram os resíduos que interagem com o molibdato em ortólogos com estrutura tri-dimensional resolvida (HU *et al.*, 1997; LAWSON *et al.*, 1998, HOLLENSTEIN *et al.*, 2007; BALAN *et al.*, 2008) (Figura 10).

		*		20		*		40		*		60		*		80											
Xac :	VTVFAAA	SLKES	MDEA	ATAYE	KAT-GTP	VRVSYAA	SSALAR-QIE	QGAPADVFFS	ADLEWMDYL	QQHGLV	LP	QQRH	NLLG	NTLV	LV	VAP :	86										
Xoo :	VTVFAAA	SLKEAM	DEAAAA	YEKAT-GTP	VRVSYAA	SSALAR-QIE	QGAPADVFFS	ADLEWMDDL	QQRGLV	LP	QQRH	NLLG	NTLV	LI	AP :	86											
Xcv :	VTVFAAA	SLKESMD	QAAAAYE	KAT-GTP	VRVSYAA	SSALAR-QIE	GGAPADVFFS	ADLEWMDYL	QDCKGL	VP	QQRH	NLLG	NTLV	LI	AP :	86											
Xcc :	LTVFAAA	SLKES	MDEA	AAAA	YEQAT-GTP	VRVSYAA	SSALAR-QIE	QGAPADVFFS	ADLEWMDYL	QQRGVL	PAQR	RNLLG	NTLV	LI	AP :	86											
Sme :	VTVFAAA	SLKNA	LD	AINGE	WLKQT-GKE	AI	VSYAA	SSALAK-QIE	QGAPADVFFS	ADLA	WMDYL	ADK	KLK	ADTR	SNLLG	NR	IV	LV	VSG :	86							
Sco :	VTVFAAA	SLKNA	LD	AINGE	WLKQT-GKE	AI	VSYAA	SSALAK-QIE	QGAPADVFFS	ADLA	WMDYL	ADK	KLK	ADTR	SNLLG	NR	IV	LV	VSG :	86							
Mlo :	VVFFAAA	SLKDA	LD	AVN	KACE	ADV-GEA	ATVSYAA	SSALAK-QIE	GGAPADVFFS	ADLD	WMKY	LS	DKK	LT	KPD	TEV	KLLG	NE	IV	LV	VAP :	86					
Eco :	ITVFAAA	SLTNAM	QDIAT	QFKKEK-GVDV-VSS	FASSSTL	ARQIE	AGAPADLFFS	ADQK	WMDY	AVDK	KAID	TAT	RQT	LLG	NSL	VV	VAP :	86									
Sfl :	ITVFAAA	SLTNAM	QDIAT	QYKKEK-GVDV-VSS	FASSSTL	ARQIE	AGAPADLFFS	ADQK	WMDY	AVDK	KAID	TAT	RQT	LLG	NSL	VV	VAP :	86									
Stm :	ITVFAAA	SLTNAM	QDIAE	YKKEK-NVDV-VSS	FASSSTL	ARQIE	AGAPADLFFS	ADQK	WMDY	AAADK	KA	VD	T	TRE	T	LLG	NSL	VV	VAP :	86							
Ype :	ITVFAAA	SLTNAL	QDIA	VYKQEK-QVDV-VASY	SSSTL	ARQIE	QGAPADLFFS	ADQK	WMDY	AI	DKQ	IV	AN	TR	Y	T	LLG	NE	IV	LV	LI	AP :	86				
Gsu :	LSLSVAA	SLKEA	INEL	TA	EY	GKKKPGIT	FLK	NYG	SSALAK-QIE	QGAPADLFFS	ANRE	WVDYL	QNR	KL	TD	EA	SR	CA	FAYN	SL	V	FAG :	86				
Rso :	LVSFAAA	SLTN	AFK	TLA	QSYEA	QHPDP	TRV	VLN	FG	SDVLMQ-QIV	RGA	PADVFA	SSAD	QEAM	KAV	SEK	VIQ	SGSR	RD	FVA	NQ	LV	IVP :	87			
Avi :	LKVVTAT	NFLGT	LE	QLA	GQFAKQT-GH	AV	VISSGS	SGPVYA-QIV	NGAPYNVFFS	AD	EK	SP	EK	LD	NQ	GFAL	PGSR	FTY	Y	IG	KL	V	WSA :	86			
Rca :	VIAAVAA	NFTE	PAKE	IGAL	FTTEKT-GHT	V	TYA	FGPS	GQFYT-QIT	QGA	PF	EV	FLS	AD	ASR	PK	KA	VEE	YAV	PGSD	FTY	Y	IG	KL	V	WSA :	86

		*		100		*		120		*		140		*		160		*																																																													
Xac :	ASSKLR	-VD	PR	AP	GAI	AKAL	G	ENG	R	LAV	G	QTAS	VFAG	KY	AAAA	LR	KL	G	QW	D	SV	SN--	RLA	ES	SV	RA	AL	M	L	V	S	R	G	E	A	P	L	G	I	V	Y	G	S	D	A	R	:	171																															
Xoo :	ASSKLR	-VD	PR	AP	GAI	AKAL	G	ES	G	R	LAV	G	QTAS	VFAG	KY	AAAA	LR	KL	G	QW	D	SV	SN--	RLA	ES	SV	RA	AL	M	L	V	S	R	G	E	A	P	L	G	I	V	Y	D	S	D	A	R	:	171																														
Xcv :	ANSKLH	-V	D	L	R	P	GAI	AKAL	G	QGR	L	AV	G	QTAS	VFAG	KY	AAAA	LR	KL	G	QW	D	SV	GN--	RLA	ES	SV	RA	AL	M	L	V	S	R	G	E	A	P	L	G	I	V	Y	G	S	D	A	R	:	171																													
Xcc :	ATSKVR	-V	D	L	R	P	GAI	AKAL	G	S	D	G	R	LAV	G	QTAS	VFAG	KY	AAAA	LR	AL	S	QW	E	G	V	S	A--	HLA	ES	SV	RA	AL	M	L	V	S	R	G	E	A	P	L	G	I	V	Y	G	S	D	A	R	:	171																									
Sme :	KSDAPA	-V	E	I	E	P	G	L	D	L	A	G	L	L	G-DG	R	L	A	M	G	A	V	D	S	VFAG	KY	G	K	A	A	E	K	L	G	L	W	P	N	V	A	G--	KV	A	E	S	V	R	A	A	L	L	V	S	R	G	E	A	P	Y	G	I	V	Y	R	T	D	A	:	170										
Sco :	KSDAPA	-V	E	I	E	P	G	L	D	L	A	G	L	L	G-DG	R	L	A	M	G	A	V	D	S	VFAG	KY	G	K	A	A	E	K	L	G	L	W	P	N	V	A	G--	KV	A	E	S	V	R	A	A	L	L	V	S	R	G	E	A	P	Y	G	I	V	Y	R	T	D	A	:	170										
Mlo :	QDSTVE	-T	K	I	E	G	F	D	L	A	K	L	V	G-DG	K	L	A	M	G	D	F	K	A	V	VFAG	KY	G	K	A	A	E	S	L	G	V	W	S	S	V	E	G--	KV	A	Q	A	E	N	V	R	A	A	L	K	L	V	S	T	G	E	A	A	L	G	I	V	Y	A	T	D	A	:	170							
Eco :	KASVQ	K	D	F	T	I	D	S	K	T	N	W	T	S	L	L	N-GG	R	L	A	V	G	D	P	E	H	V	VFAG	I	Y	A	K	E	A	L	Q	K	L	G	A	W	D	T	L	S	P--	KL	A	P	A	E	D	V	R	G	A	L	A	L	V	E	R	N	E	A	P	L	G	I	V	Y	G	S	D	A	V	:	171	
Sfl :	KASEQ	K	D	F	T	I	D	S	K	T	N	W	T	S	L	L	N-GG	R	L	A	V	G	D	P	E	H	V	VFAG	I	Y	A	K	E	A	L	Q	K	L	G	A	W	D	T	L	S	P--	KL	A	P	A	E	D	V	R	G	A	L	A	L	V	E	R	N	E	A	P	L	G	I	V	Y	G	S	D	A	V	:	171	
Stm :	KASEQ	K	P	F	T	I	D	N	K	T	D	W	I	H	L	L	N-GG	R	L	A	V	G	D	P	Q	H	V	VFAG	I	Y	A	K	E	A	L	Q	K	L	G	A	W	T	L	S	P--	KL	A	P	A	E	D	V	R	G	A	L	A	L	V	E	R	N	E	A	P	L	G	I	V	Y	G	S	D	A	V	:	171		
Ype :	QDSQ	I	D	K	V	E	I	D	K	K	T	D	W	K	L	L	E-GG	R	L	A	V	G	D	P	D	H	V	VFAG	I	Y	A	K	E	S	L	E	N	L	G	A	W	S	T	L	A	P--	EM	A	R	A	N	N	V	R	S	A	M	L	V	E	R	A	E	A	P	L	G	I	V	Y	G	S	D	A	V	:	171		
Gsu :	-----	T	T	Q	P	V	A	G	M	K	D	L	A	K	L	R	K	I	A	I	G	S	P	G	S	VFAG	E	Y	A	M	E	S	F	K	A	A	G	I	D	K	A	L	D	G--	KL	V	M	A	K	D	V	R	E	A	M	K	Y	A	E	L	G	E	V	D	G	A	F	V	Y	R	T	D	A	:	164				
Rso :	SASAVP	---	V	H	A	L	G	D	L	A	R	P	E	V--	Q	R	V	A	I	G	N	P	A	S	VFAG	R	Y	A	K	R	A	L	E	A	A	G	L	W	Q	P	V	S	A--	K	A	V	L	A	Q	N	V	R	Q	A	L	D	Y	V	A	R	G	E	V	E	A	G	L	V	F	A	T	D	A	:	168				
Avi :	KPGLVD	---	---	N	Q	G	K	V	L	A	G	N	G	R	H	I	A	I	S	N	P	Q	I	A	P	Y	G	L	A	G	T	Q	V	L	T	H	L	G	L	L	D	K	L	A	Q	E	R	I	V	E	A	N	S	V	G	Q	A	H	S	Q	T	A	S	G	A	A	D	L	G	F	V	A	L	A	Q	I	I	:	168
Rca :	DPA	R	V	D	---	---	D	K	G	A	V	L	K	T	D	-L	E	H	V	A	I	A	D	P	K	S	A	Y	G	A	A	I	E	V	M	Q	K	L	G	V	Y	E	T	L	E	P--	KL	V	T	G	K	S	I	S	Q	A	Y	E	F	A	A	T	G	N	A	P	V	G	F	V	A	S	Q	V	I	:	165		

Figura 10: Alinhamentos das sequencias de aminoácidos de 15 ortólogos de ModA encontrados em bactérias. As regiões coloridas representam as seis posições conservadas que incluem todos os resíduos que interagem com o molibdato. Os aminoácidos responsáveis pela interação com o molibdato estão representados em amarelo, quando esses formam ligações de hidrogênio, ou em verde, quando apresentam átomos de carbono interagindo com o ligante. A numeração e os asteriscos acima das sequencias indicam a posição de cada resíduo na proteína ModA.

Utilizando o alinhamento da figura 10 geramos uma árvore consenso entre 100 árvores não-enraizadas obtidas pelo programa Phylip. Essa árvore revelou a presença de três grandes ramos (Figura 11). O primeiro ramo (grupo 1) inclui organismos pertencentes às espécies que interagem com plantas e uma actinobactéria (*Xac*, *Xcc*, *Xoo*, *Xcv*, *Mlo*, *Sme* e *Sco*). O segundo ramo (grupo 2) inclui os ortólogos de ModA pertencentes às enterobactérias (*Eco*, *Stm*, *Sfl* e *Ype*). Finalmente o terceiro ramo (Grupo 3) engloba as espécies mais distantes filogeneticamente em relação à proteína ModA de *Xac* e inclui *Avi*, *Gsu*, *Rso* e *Rca*. A análise de cada grupo revelou identidade de sequência de aminoácidos maior entre as enterobactérias e as bactérias que interagem com plantas. Os valores de identidade variaram entre 41% (*Sme* com *Ype*) a 51% (*Xac* com *Eco*). Por outro lado, o grupo 3 apresentou maior diversidade com valores de identidades menores entres os ortólogos de ModA (máximo de 15%).

Figura 11: Análise filogenética de diferentes ortólogos da proteína ModA de bactérias. Árvore não-enraizada gerada pelo programa PHYLIP pelo método de “Neighbor Joining” (NJ) e valor de “bootstrap” de 100. O grupo 1 está representado pelas bactérias que interagem com plantas: *Xac*, *Xcc*, *Xoo*, *Xcv*, *Mlo*, *Sco* e *Sme*. O grupo 2 inclui as enterobactérias: *Eco*, *Stm*, *Sfl* e *Ype*. O grupo 3 compreende as bactérias de solo: *Avi*, *Gsu*, *Rso* e *Rca*.

4.4 Modelo tri-dimensional da proteína ModA de *Xac*

Com as coordenadas estruturais disponíveis para a proteína ModA de *Eco* e resolução de 1.75 Å definimos um modelo estrutural para a proteína ModA de *Xac*. Para alcançar o melhor modelo estrutural empregamos três estratégias: 1) alinhamento sequencial e correção visual entre ModA de *Xac* e de *Eco* pelo programa ClustalX (THOMPSON *et. al.*, 1997); 2) a modelagem propriamente dita, realizada por meio do programa Modeller utilizando a configuração básica para gerar 100 modelos; 3) e, por fim, dinâmica molecular do melhor modelo feita com o programa Gromacs. O melhor modelo foi submetido à dinâmica molecular com o objetivo de se atingir os menores níveis de energia e estabilidade da proteína. O gráfico de Ramachandran revelou que 95.59% (195) dos resíduos encontram-se nas áreas mais favoráveis, enquanto que nenhum resíduo foi localizado em regiões desfavoráveis (Figura 12). A qualidade estereoquímica do modelo também foi confirmada pelo programa PROCHECK (LASKOWSKI *et al.*, 1993)

Figura 12: Gráfico de Ramachandran do modelo gerado para ModA de *Xac*. Os dados estatísticos mostram 195 resíduos (95,59%) nas regiões mais favoráveis, 9 resíduos (4,41%) nas regiões permitidas adicionais e nenhum resíduo presente tanto nas regiões generosamente permitidas como nas regiões não permitidas de um total de 204 resíduos (glicinas e prolinas ausentes).

O modelo da ModA de *Xac* apresenta uma estrutura tri-dimensional em forma elipsoidal com dimensões aproximadas de 33 Å x 48 Å x 62 Å com dois domínios similares (caracterizado por sanduíche de alfa/beta), designados de domínios N e C separados por uma região ligadora de ânions (Figura 13). O domínio N é formado pelos resíduos 1 a 82 e 196 a 234 englobando as regiões definidas por uma estrutura folha beta pregueada com 6 fitas ($\beta 1$, $\beta 2$, $\beta 3$, $\beta 4$, $\beta 10$, $\beta 11$) circundadas por 5 hélices alfa ($\alpha 1$, $\alpha 2$, $\alpha 3$, $\alpha 8$, $\alpha 9$). O domínio C, delimitado pelos resíduos 83 a 195, é formado por 5 fitas β -pregueadas ($\beta 5$, $\beta 6$, $\beta 7$, $\beta 8$, $\beta 9$) circundadas por 4 α -hélices ($\alpha 4$, $\alpha 5$, $\alpha 6$, $\alpha 7$) (Figura 13).

Figura 13: Modelo estrutural da proteína ModA de *Xac* apresentado na forma de “cartoon” com o íon molibdato indicado no sítio de ligação. A extremidade N-terminal (N) está localizada acima e à direita e a extremidade C-terminal (C) está localizada abaixo do ânion de molibdato.

As estruturas das proteínas ModA de *Eco* e *Xac* foram alinhadas pelos seus respectivos carbonos alfa. Como observa-se na Figura 14A e na tabela 2, o alinhamento das duas estruturas da ModA revelou uma elevada similaridade, com r.m.s.d. de 1.5 Å. A única região que não apresentou similaridade estrutural está compreendida entre os resíduos Thr¹⁰⁷ à Asn¹¹¹ em *Eco* e em *Xac* pelos resíduos Ile¹⁰⁵ à Leu¹⁰⁹ (Figura 14B)

Figura 14: Alinhamento estrutural entre as proteínas ModA de *Eco* e do modelo de *Xac* representadas em “cartoon”. Em (A) a ModA de *Eco* está colorida em azul e a ModA de *Xac* em vermelho. Em (B) as duas estruturas estão coloridas de acordo com as estruturas secundárias, hélices em vermelho, fitas em amarelo e alças em verde. A seta indica a hélice ausente em *Xac*.

Realizamos comparações estruturais entre o modelo de ModA, definido no presente estudo, e a estrutura cristalográfica (código PDB: 2h5y) resolvida pelo nosso grupo (BALAN *et al.*, 2008). De um modo geral o modelo de ModA definido por modelagem molecular apresentou uma excelente correlação com a estrutura definida experimentalmente. Apenas duas regiões presentes no domínio C apresentaram pequenas divergências com a estrutura cristalográfica (Figura 15). A primeira divergência envolve a ausência de uma hélice delimitada pelos resíduos Gli¹⁰³ até Leu¹⁰⁹ (Figura 15A, B e C), semelhante ao encontrado com a ModA de *Eco*, na qual esta estrutura secundária está ausente (Figura 14B). A segunda divergência consiste em uma pequena hélice não modelada entre os resíduos Pro¹⁸⁷ e His¹⁹¹ (Figura 15A). As demais diferenças estruturais concentraram-se em regiões permitidas das alças. Para o restante do modelo da ModA definido por técnicas de modelagem molecular, incluindo os trechos que delimitam a cavidade ligadora do molibdato, houve uma correlação perfeita com a estrutura real da proteína.

Figura 15: Alinhamento estrutural do modelo da ModA de Xac definido no presente estudo (vermelho) com a estrutura cristalográfica definida experimentalmente (amarelo). (A) Vista frontal do alinhamento com a fenda que compõe a cavidade ligadora separando os dois domínios N (à direita) e C (à esquerda). (B) Vista frontal após rotação horizontal de 180°. (C) Vista superior do alinhamento após rotação vertical de 90°. O domínio N está indicado na figura pela letra N. As setas 1 e 2 indicam as duas regiões não alinhadas.

4.4.1 A cavidade ligadora de substrato da ModA de *Xac*

A cavidade ligadora de molibdato define a afinidade e permite a interação do íon com as proteínas ModA bacterianas. Analisamos a cavidade ligadora de molibdato no modelo descrito e determinamos os resíduos que interagem com o íon.

Primeiramente, as coordenadas estruturais do molibdato, obtidas a partir do arquivo pdb da estrutura da ModA de *Eco*, foram transferidas para o arquivo do modelo da ModA de *Xac*. A posição do molibdato foi centrada na cavidade ligadora. A partir da estrutura modelada com o molibdato, os tipos de interações e resíduos envolvidos foram determinados e comparados com os resíduos que interagem com o íon nas estruturas de ModA definidas experimentalmente. Duas abordagens foram empregadas para caracterizar os resíduos que interagem com o molibdato. A primeira empregou como parâmetro as interações do molibdato nos diferentes ortólogos com estrutura definida. As interações foram aplicadas à estrutura modelo e oito resíduos foram identificados: Ala¹⁰, Ser¹², Ser³⁹, Ala⁵⁸, Val¹²³, Ala¹²⁵, Val¹⁵² e Tyr¹⁷⁰. Na segunda abordagem adicionamos as coordenadas do íon no modelo e verificamos as interações polares. O resultado excluiu os resíduos Ala¹⁰, Ala⁵⁸ e Val¹²³ (Figura 16). Como resultado final, cinco resíduos foram identificados como responsáveis pela ligação ao molibdato na cavidade ligadora na ModA. As linhas pontilhadas na Figura 16 indicam as distâncias das ligações entre os átomos dos resíduos da ModA e os quatro oxigênios do molibdato (MoO₄).

Figura 16: Resíduos da proteína ModA de *Xac* que interagem com o molibdato definidos a partir do modelo estrutural gerado no presente estudo. As linhas pontilhadas em azul indicam as distâncias dos átomos dos resíduos da ModA até os oxigênios do íon. Todos os átomos foram coloridos segundo o padrão: oxigênio em vermelho; nitrogênio em azul; carbonos e hidrogênios em cinza. Os átomos e resíduos que interagem com o ligante, segundo nosso modelo, estão indicados.

Segundo o nosso modelo, o hidrogênio (NH) da Ser¹² está distante 2.64 Å do oxigênio⁴ do molibdênio. O hidrogênio (NH) da Ser³⁹ está distante 2 Å e o hidrogênio (OγH) encontra-se apenas 1.44 Å, ambos, do oxigênio² do molibdato. A Ala¹²⁵ possui seu hidrogênio (NH) e o carbono^β distante de 3.85 Å e 2.56 Å, respectivamente do oxigênio³ do molibdato. O resíduo Val¹⁵² faz uma interação distante 2.84 Å do seu hidrogênio (NH) e o oxigênio¹ do molibdato. Por fim, os hidrogênios ε2 e Δ2 da Tyr¹⁷⁰ estão distantes 4.84 Å e 4.9 Å, respectivamente, do oxigênio³ do molibdênio e não interagem com o íon (Tabela 2). Todos os resíduos estão situados na porção N-terminal das α-hélices: Ser¹² na α1; Ser³⁹ na α2; Ala¹²⁵ na α4; Val¹⁵² na α6 e Tyr¹⁷⁰ na α7 (Figura 13) permitindo a formação de um dipolo que acopla perfeitamente por todos os ângulos o íon na cavidade ligadora da proteína (Figura 16).

Quando comparamos os dados obtidos no presente estudo com aqueles gerados após a definição da estrutura da proteína ModA de *Xac* (Balan *et al.* 2008), concluímos que a maioria das interações entre proteína e ligante foram previstas corretamente com apenas três exceções: 1) não conseguimos identificar uma segunda interação da Ser¹²/O^YH\O⁴; 2) acrescentamos uma interação para Ala¹²⁵/Cβ\O³; 3) ausência das interações da Tyr¹⁷⁰ definidas na estrutura cristalográfica definida experimentalmente (Tabela 2; Figuras 16 e 17). A ausência das interações da Tyr¹⁷⁰ e o ligante seriam resultado da escolha de um rotâmero inadequado. Para Ala¹²⁵ o processo de modelagem fez com que o posicionamento espacial deste resíduo fosse alterado ficando mais próximo do ligante e o afastou da Ser¹². Esse deslocamento foi responsável pela perda de uma interação entre a Ser¹² e o acréscimo de uma interação em Ala¹²⁵ (Figura 18). As interações encontradas na ModA de *Xac* e o ligante foram idênticas àsquelas encontradas no ortólogo de *Eco*. Em *Avi* duas interações adicionais são observadas entre os resíduos Thr⁹ e Ser³⁷ com os oxigênios 1 e 4 do ligante, respectivamente e que não estão presentes no nosso modelo (Tabela 2).

Figura 17: Sítios de ligação da proteína ModA de *Xac* resolvida por cristalografia e o ligante (molibdato) com visualização da cavidade ligadora da proteína ModA de *Xac* e os resíduos que interagem com o molibdato. Faixas vermelhos indicam o eixo das α -hélices e as linhas pontilhadas em azul indicam as ligações de hidrogênio. A porção vermelha dos resíduos e do ligante indicam a presença de átomos de oxigênio e em azul átomos de hidrogênio.

Tabela 2: Interações moleculares entre as proteínas ModA com a estrutura definida e o íon molibdato.

<i>Xac</i> Modelo*	<i>Xac</i> 2h5y	<i>Eco</i> 1amf	<i>Avi</i> 1atg
Val ¹⁵² /NH\O ¹	Val ¹⁵² /NH\O ¹	Val ¹⁵² /NH\O ¹	Val ¹⁴⁷ /NH\O ¹
			Thr ⁹ /OG1CB\O ¹
Ser ³⁹ /NH\O ²	Ser ³⁹ /NH\O ²	Ser ³⁹ /NH\O ²	Asn ¹⁰ /NH\O ²
Ser ³⁹ /O ^y H\O ²	Ser ³⁹ /O ^y H\O ²	Ser ³⁹ /O ^y H\O ²	Asn ¹⁰ /ND2\O ²
Ala ¹²⁵ /NH\O ³	Ala ¹²⁵ /NH\O ³	Ala ¹²⁵ /NH\O ³	Pro ¹¹⁷ /NH\O ³
Ala ¹²⁵ /C β \O ³			
	Tyr ¹⁷⁰ /O η H\O ³	Tyr ¹⁷⁰ /O η H\O ³	Tyr ¹¹⁸ /O η H\O ³
Ser ¹² /NH\O ⁴	Ser ¹² /NH\O ⁴	Ser ¹² /NH\O ⁴	Ser ³⁷ /NH\O ⁴
	Ser ¹² /O ^y H\O ⁴	Ser ¹² /O ^y H\O ⁴	Ser ³⁷ /O ^y H\O ⁴

*Comparação dos resultados obtidos com o modelo da ModA de *Xac* modelo e os dados cristalográficos da ModA de *Xac* (2h5y) e seus ortólogos em *Eco* (1amf) e *Avi* (1atg).

Figura 18: Comparação dos sítios de ligação entre o modelo da proteína ModA de *Xac* e o íon molibdato comparado com sua estrutura cristalina. Em azul estão representados os resíduos do modelo da ModA de *Xac* de acordo com o modelo definido no presente estudo e em vermelho da estrutura definida experimentalmente. O molibdato está representado no formato de barra com os seus átomos de oxigênios em vermelho e o metal em cinza.

As estruturas cristalográficas das proteínas ModAs de *Eco*, *Azo* e *Afu* (HU *et al.*, 1997; LAWSON *et al.*, 1998, HOLLENSTEIN *et al.*, 2007; BALAN *et al.*, 2008) foram sobrepostas com o modelo estrutural da ModA de *Xac* e os carbonos α alinhados. Como é observado na Figura 19 e na tabela 3, o alinhamento dos quatro ortólogos da ModA revelou uma similaridade crescente entre a estrutura modelada da ModA e os ortólogos de *Afu*, *Avi* e *Eco* com identidades de 24%, 26% e 52% e um r.m.s.d. de 1.9 Å, 1.8 Å, 1.6 Å e 1.5 Å, respectivamente. A região em preto, indicada na figura 19, foi descrita como responsável pela interação com a ModB e por não apresentar sobreposição com os carbonos alfas da ModA de *Eco*. Essa região não se sobrepôs aos outros ortólogos analisados (LAWSON *et al.*, 1998).

Tabela 3: Alinhamento dos carbonos alfa entre os ortólogos da ModA.

Organismo 1	R.M.S.D (Å)	Organismo 2	Identidade (%)
<i>Xac Modelo</i>	1.5	<i>Eco</i>	52
<i>Xac Modelo</i>	1.6	<i>Avi</i>	26
<i>Xac Modelo</i>	1.8	<i>Xac 2h5y</i>	93
<i>Xac Modelo</i>	1.9	<i>Afu</i>	24
<i>Xac 2h5y</i>	1.1	<i>Eco</i>	52
<i>Xac 2h5y</i>	1.6	<i>Avi</i>	26
<i>Xac 2h5y</i>	1.8	<i>Afu</i>	22
<i>Eco</i>	1.9	<i>Afu</i>	21
<i>Avi</i>	1.8	<i>Eco</i>	24
<i>Avi</i>	2.1	<i>Afu</i>	19

Figura 19: Sobreposição dos carbonos alfa das estruturas tridimensionais das ModA de *Xac* cristal (salmão), *Xac* modelo (vermelho), *Eco* (azul), *Avi* (amarelo), *Afu* (verde). A região colorida em preto delimita a alça presente em *Avi* e que não apresenta alinhamento com as outras estruturas.

4.5 Modelagem molecular da proteína ligadora de oligopeptídeos (OppA) de *Xac*

O modelo estrutural da OppA de *Xac* utiliza as coordenadas estruturais das OppAs de *Bsu* (AppA) complexada com um nonapeptídeo na resolução de 1.55 Å (1xoc) (LEVDIKOV *et al.*, 2005), e *Stm* (OppA), complexada com um tripeptídeo na resolução de 1.75Å (1jev) (TAME *et al.*, 1996). A estrutura do modelo da OppA de *Xac* mostra-se do tipo α/β , consistindo em uma mistura de folhas β rodeadas por α -hélices. O modelo apresenta uma forma elipsoidal com dimensões aproximadas de 85 Å x 67 Å x 65 Å, composta por dois domínios similares, designados domínios I e III, separados por uma fenda na qual se localiza a região ligadora de oligopeptídeos (Figura 20A). O domínio I pode ser dividido em dois sub-domínios (I e II), como descrito para AppA de *Bsu* (LEVDIKOV *et al.*, 2005). O subdomínio I é formado por três sequências descontínuas, definidas pelos resíduos de 1 a 44, 154 a 260 e 491 a 521, que delimitam uma região do tipo folha β constituída por sete fitas rodeadas por quatro α -hélices. Os resíduos Asp⁴⁵ e Ser¹⁵³ estão envolvidos na conexão entre os sub-domínios I e II. Existe uma ligação dissulfídica entre Cys²⁵ e Cys¹⁶⁵, conectando os dois segmentos descontínuos do domínio I (1-44 e 154-260). O subdomínio II consiste nos resíduos 45 a 153 e está formado por uma folha β com quatro fitas anti-paralelas, com um lado exposto para a superfície da molécula e outro rodeado por sete α -hélices que delimitam a região central altamente hidrofóbica. O outro lóbulo da proteína é formado pelo domínio III, definido pelos resíduos 261 a 490, representados por sete fitas rodeadas por nove α -hélices (Figura 20).

A qualidade do modelo foi avaliada com o programa PROCHECK (LASKOWSKI *et al.*, 1993). A qualidade estereoquímica geral da estrutura do modelo da OppA de *Xac*, demonstrada pelo gráfico de Ramachandran foi boa com 88,37% (395) dos resíduos incluídos nas regiões mais favoráveis, 11,63% (52) dos resíduos em regiões permitidas e nenhum resíduo nas regiões não permitidas ou desfavoráveis. Os ortólogos da OppA de *Xac*, *Bsu* e *Stm* apresentam baixa similaridade sequencial com identidades de 18,5% e 18%, respectivamente. A reduzida similaridade sequencial também é observada na estrutura depositada

recentemente da bactéria gram-negativa *Thermotoga maritima* (*Tma*) (código PDB: 1vr5) que apresentou 12,5% de identidade com a OppA de *Xac* (r.m.s.d. de 3,8Å).

Apesar das diferenças sequenciais, o alinhamento estrutural dos C^α revelou que a OppA de *Xac* é conservada quando comparada aos ortólogos *Bsu* (r.m.s.d. 1,4 Å) e *Stm* (r.m.s.d. 1,8 Å) (Figura 20). As principais diferenças observadas concentram-se em duas alças extras na OppA de *Xac*, definidas pelas regiões entre os resíduos Ala¹⁹⁴ e His²⁰² e Gly⁴³⁵ e His⁴⁴⁰ (indicado em preto na Figura 20). O alinhamento estrutural das três estruturas das OppAs revelou também a presença de uma alça extra na proteína expressa por *Stm* delimitada pelos resíduos Asn¹²⁰ a Ala¹³¹ (indicada em cinza na Figura 20).

Figura 20: Modelo estrutural da OppA de *Xac*. A) Visão global da estrutura da OppA de *Xac* no formato “cartoon”. A proteína é formada por 3 domínios indicados por cores diferentes: o domínio I delimitado pelos aminoácidos 1 a 44, 154 a 260 e 491 a 521 (em vermelho), o domínio II delimitado pelos aminoácidos 45 a 153 (verde), e o domínio III formado pelos aminoácidos 261 a 490 (em amarelo). As extremidades N- (N-term) e C-terminal (C-term) da proteína estão indicadas. B) Sobreposição das estruturas da OppA de *Xac* (azul), OppA de *Stm* (vermelha) e AppA de *Bsu* (amarelo). Alças presentes apenas nos ortólogos de *Xac* e *Stm* estão indicadas em preto e cinza, respectivamente. A sobreposição dos átomos de carbonos alfa das três estruturas foi realizada pelo programa CE.

4.5.1 Afinidade por oligopeptídeos da OppA de *Xac*

Nossas análises de afinidade da proteína OppA de *Xac* por diferentes oligopeptídeos estão divididas em três partes em função da abordagem metodológica empregada. Na primeira análise utilizamos as informações dos ligantes das OppAs de *Eco* e *Bsu*, utilizadas como moldes para a definição do nosso modelo. Por meio dessa abordagem analisamos as interações da OppA de *Xac* com um oligopeptídeo de cadeia longa (nonapeptídeo, VDSKNTSSW, da AppA de *Bsu*) outro de cadeia curta (tripeptídeo, KWK, da OppA de *Stm*). Nossos resultados revelaram que a OppA de *Xac* possui um perfil de ancoragem idêntico ao observado em *Stm* para o tripeptídeo KWK, que permaneceu ancorado na fenda formada entre os domínios I e III (Figura 21 A e B). Os principais resíduos responsáveis pela cavidade ligadora da OppA de *Xac* foram: Met³⁴, Val³⁷, Arg⁴¹, Ala¹⁴⁵, Asn⁴¹⁴, His⁴⁴⁰, Thr⁴⁹², Pro⁵⁰⁷ e Leu⁵⁰⁹. A ancoragem de KWK com o nosso modelo, revelou interações polares entre Arg⁴¹/NH₂ (OppA) e Trp²/O (ligante), Arg⁴¹/O (OppA) e Lys¹/NZ (ligante) e Asn⁴¹⁴/ND2 (OppA) e Lys¹/O (ligante) (Figura 21).

Os alinhamentos estruturais das OppAs de *Xac* e *Stm* ligadas ao tripeptídeo KWK mostraram que os resíduos formadores de interações polares (Arg⁴¹) de ambas estão estruturalmente alinhados e interagem com o ligante. O resíduo Asn⁴¹⁴ da OppA de *Xac*, que corresponde ao resíduo Asp⁴¹⁹ na OppA de *Stm*, é conservado em diversas proteínas ligadoras de oligopeptídeos (DETMERS *et al.*, 2001).

Os resultados de ancoragem com nonapeptídeo VDSKNTSSW indicam que 17 resíduos da OppA de *Xac* interagem com o ligante, porém o oligopeptídeo não se acomoda perfeitamente no nosso modelo, restando parte da extremidade N-terminal do nonapeptídeo fora da fenda ligadora. Esses resultados indicam que, a partir do modelo gerado, a proteína ModA de *Xac* possui uma maior proximidade funcional com a proteína OppA encontrada em enterobactérias, especificamente com a proteína expressa por *Stm*.

Figura 21: Ancoragem do tripeptídeo KWK na cavidade ligadora da OppA de *Xac*. **A)** Representação da OppA de *Xac* e do oligopeptídeo ancorado na fenda formada pelos domínios I e III (representado em azul escuro). **B)** Principais resíduos formadores da cavidade ligadora da OppA de *Xac*. Os resíduos formadores de interações polares com KWK estão sublinhados e suas interações estão indicadas por linhas pontilhadas.

Uma segunda abordagem voltada para a caracterização da afinidade de OppA de *Xac* frente a diferentes cadeias de oligopeptídeos foi feita por meio do programa GRAMM-X. Oligopeptídeos previamente identificados em ensaios de

captação como capazes de se ligarem a diferentes OppAs bacterianas foram utilizados assim como outros criados para dar suporte as comparações. Nossos resultados com o tripeptídeo KKK mostraram que eventos de ancoragem positivos ocorrem mais frequentemente na fenda ligadora da OppA de *Xac* (38 eventos positivos em 100 tentativas de ancoragem) do que com oligopeptídeos com cadeias de peptídeos variando entre 5 ou mais resíduos (menor que 20 eventos positivos em 100 tentativas de ancoragem). Para validar a abordagem de determinação de afinidade da OppA de *Xac*, foram feitas simulações com polilisinas variando entre 3 a 7 resíduos de extensão. Os resultados demonstram que a proteína OppA de *Xac* possui maior afinidade a peptídeos de cadeia curta, como evidenciado pelo número de eventos positivos de ancoragem (Figura 22).

Figura 22: Análise de ancoragem com diferentes oligopeptídeos de polilisinas (K_n aonde n é o número de lisinas) e a proteína OppA de *Xac*.

Utilizamos ainda o tripeptídeo KKK para avaliar a influência do aminoácido central na ligação à proteína OppA de *Xac* (Figura 23). A substituição da lisina central do tripeptídeo por glicina, valina, prolina, treonina, ou serina aumentou em aproximadamente 10% o número de eventos de ancoragem positivos para a OppA de *Xac*. Esse resultado indica que tanto o tamanho do ligante como a sua composição, interferem na afinidade e na ligação com a proteína OppA de *Xac*.

Figura 23: Análise de ancoragem com diferentes oligopeptídeos e a proteína OppA de *Xac*. Relevância do segundo resíduo nas ancoragens com OppA para tripeptídeos KXK, onde X é substituído por um dos 20 aminoácidos existentes.

Finalmente, usando as coordenadas de peptídeos que foram descritas anteriormente por serem captadas pelos ortólogos de OppA de *Stm* (HILLES *et al.*, 1986; FENNO *et al.*, 2000; MONNET, 2003), *Bsu* (MONNET, 2003), *Bbu* (PICON *et al.*, 2000) e *Lla* (TAME *et al.*, 1995) demonstramos que o tetrapeptídeo VKPG, originalmente reconhecido pela OppA de *Stm* (Figura 24, colunas brancas), gerou o maior número de eventos positivos (43%) com a OppA de *Xac*, seguida por KKK (39%) e PIQAA (37%) (Figura 24). Os resultados confirmam as observações feitas com o tripeptídeo KKK, e indicam que a presença de aminoácidos apolares ou polares neutros, com cadeias laterais pequenas, aumenta a afinidade e o reconhecimento pela proteína OppA de *Xac* pelo oligopeptídeo.

Figura 21: Análise de ancoragem com diferentes oligopeptídeos à proteína OppA de *Xac*. Ensaio de ancoragem da OppA de *Xac* com oligopeptídeos reconhecidos por diferentes ortólogos bacterianos de OppA. Os peptídeos testados estão representados por colunas negras, oligopeptídeos reconhecidos pela OppA de *Lla*; colunas brancas, oligopeptídeos reconhecidos pela OppA de *Stm*; colunas com listras horizontais, oligopeptídeos reconhecidos pela OppA de *Bsu* e colunas listradas verticalmente, oligopeptídeos reconhecidos pela OppA de *Bbu*.

5 DISCUSSÃO

O advento da genômica criou um novo paradigma para a pesquisa biológica, em particular para a microbiologia. A disponibilidade de sequências gênicas completas permite que informações importantes sobre a regulação de genes e função de proteínas possam ser obtidas antes mesmo da clonagem gênica e purificação dos produtos codificados. A biologia estrutural surgiu como uma complementação importante aos estudos em genômica. A definição de modelos estruturais de proteínas, em conjunto com estudos funcionais, gera informações mais precisas sobre vias metabólicas e seus componentes.

Os sistemas de transporte ABC são sistemas primários de transporte presentes em archaea, eubactérias e eucariotos voltados para a captação ou a exportação de diferentes substâncias (HIGGINS, 1992; HIGGINS, 2001). O molibdênio é um elemento requerido para o crescimento da maioria dos organismos, incluindo plantas e bactérias. Utilizado como parte integral do complexo pterina para atuar como cofator de várias enzimas, este íon está comprometido em reações de oxido-redução e é essencial para o funcionamento dos ciclos do carbono e captação de nitrogênio nas formas de gás e de nitrato (KISKER e SCHINDELIN, 1997; WILLIAMS e da SILVA, 2002). Os peptídeos representam um componente importante na nutrição da maioria dos seres vivos, particularmente das bactérias. Em bactérias, peptídeos são descritos em processos de sinalização intra e intercelular, com o desenvolvimento de estados fisiológicos específico como a esporulação, e a síntese da parede celular (PAYNE e GILBARG, 1968; HILLES e HIGGINS, 1986; FUQUA *et al.*, 1994). No presente trabalho propusemos a análise dos óperons *mod* de *Xac*, assim como as proteínas ModA e OppA, por meio de ferramentas de bioinformática, voltadas para a análise de sequências gênicas e modelagem molecular, de forma a definir estruturas protéicas e suas interações com seus respectivos ligantes.

Em *Xac* os sistemas de transporte de oligopeptídeos e molibdato foram anotados e identificados, mas pouco se conhece sobre a função destes no metabolismo, crescimento ou mesmo, patogenicidade desta bactéria. Entre os sete genomas completos e dois incompleto de bactérias do gênero *Xanthomonas*,

apenas no genoma de *Xac* encontramos o óperon *opp*. O óperon *mod*, ao contrário, foi identificado em todas as espécies de *Xanthomonas* com genomas definidos de forma completa.

A organização genética mínima do óperon *mod* apresentada por todas as bactérias do gênero *Xanthomonas*, e demais microrganismos estudados, envolve os genes *modABC*. Somente em *Xoo* esse sistema é transcrito em sentido contrário comparado aos demais organismos fitopatógenos estudados. Nestes organismos, o transporte de molibdato é realizado por um sistema de alta afinidade composto pela proteína ligadora ModA, uma proteína integrada na membrana ModB e uma proteína energizante ModC, que compõem o policistron *modABC* com elevada identidade sequencial. Nas espécies de *Xanthomonas* estudadas a identidade dos cistrons *modA* variou entre 82% a 94% enquanto que *modB* a conservação variou entre 96% a 98% de identidade enquanto que em *modC* as identidades variaram entre 88% a 95%. A similaridade demonstrada pela linhagem de *Xac* e seus ortólogos fitopatógenos decresce na seguinte ordem: *Xcv*, *Xoo* e *Xcc*.

A regulação do óperon *mod* em algumas bactérias pode ser feita por proteínas expressas pelos genes *mop* ou pela proteína ModE (GRUNDEN e SHANMUGAM, 1997). A proteína ModE é capaz de reprimir a transcrição\ tradução dos genes *modAB in vitro* e esta repressão ocorre na presença de Molibdênio (GRUNDEN *et al.*, 1996). De acordo com estudos realizados por Zhang e Gladyshev em 2008, a regulação da captação de Mo promovida por ModE ocorre em menos de 30% dos organismos. A grande maioria dos microrganismos, destes, 70% das bactérias e 80% das arqueas, aparentam não necessitar da regulação promovida por ModE (ZHANG e GLADYSHEV, 2008). Em nossas análises não encontramos nenhum gene relacionado ao mecanismo de regulação nos organismos pertencentes ao gênero *Xanthomonas*, *Sme*, *Rca*, *Rso*, *Avi* e *Sco*.

Apesar de um elevado número de organismos que não apresentam o gene *modE*, o grupo das enterobactérias e os organismos *Mlo*, *Gsu* apresentaram a proteína reguladora ModE. A proteína ModE também controla a transcrição dos genes codificantes para a síntese de molidopterina e molidoenzimas (ZHANG e GLADYSHEV, 2008). Em bactérias capazes de reduzir dinitrogênio e nitrato em amônia é comum a presença das molibdoenzimas dinitrogenase e nitrato redutase, respectivamente (GRUNDEN e SHANMUGAM, 1997). Exceto para dinitrogenases,

que contém um cofator Fe-Mo, todas as demais molibdoenzimas caracterizadas apresentam uma única forma de molibdopterina (RAJAGOPALAN e JOHNSON, 1992). Assim, parece lógico a presença de molibdopterinas como componentes adicionais do óperon *mod* entre as enterobactérias estudadas e *Mlo*. Ainda entre as enterobactérias, a única diferença observada dentro do grupo, foi na organização genética do óperon *mod* em *Sfl*. Nesta bactéria, o óperon está com o sentido de transcrição invertido em relação às demais, exceto o gene da molibdopterina.

A proteína ModA, dentre todas supra citadas, é a que define a função do óperon Mod no processo de captação do substrato. Alinhamentos sequenciais dos ortólogos da ModA revelaram seis regiões conservadas onde encontramos todos os resíduos descritos que interagem com o molibdato (HU *et al.*, 1997; LAWSON *et al.*, 1998, HOLLENSTEIN *et al.*, 2007; BALAN *et al.*, 2008). A princípio pensamos que a cavidade ligadora fosse composta por esses resíduos, mas nossas análises mostraram que eles estão na superfície da proteína. Futuras análises de interações de ancoragem proteína-proteína, poderão demonstrar se esses seis resíduos realizam alguma interação com os outros componentes de permease do óperon *mod*.

Uma árvore não enraizada dos diversos ortólogos da ModA pelo método de “Neighbor Joining”, permitiu correlacioná-los em três grupos distintos, os quais descrevemos como 1) bactérias que interagem com plantas; 2) Enterobactérias e 3) de meio ambiente (solo e água). Eles são representados estruturalmente pelas ModAs de *Xac* (nosso modelo); *Eco* (1amf) e *Avi* (1atg), respectivamente. Nossos dados sugerem que exista uma maior proximidade entre o grupo de bactérias de plantas e as enterobactérias, fato interessante quando pensamos que todas as bactérias do grupo que interage com plantas podem ser encontradas no solo. As bactérias *Xac* e *Eco* são relativamente próximas e possuem identidade de 51%. Ainda entre o grupo 1 e 2, os mais distantes são *Sme* e *Ype* com 41% de identidade. O grupo três foi o que mostrou o menor valor de identidade (15%) com os demais grupos. No grupo dois, nossas análises incluíram dois simbioses (*Sme* e *Mlo*) que interagem com raízes de plantas leguminosas pertencentes ao gênero *Rhizobium* (GUO *et al.*, 2003). O agrupamento de fitopatógenos e simbioses de plantas foi outro fato inesperado. Apesar de encontrarmos o gênero de *Xanthomonas* em ambientes como o solo, onde encontramos os fixadores de nitrogênio, esperávamos que evolutivamente fossem mais distantes, já que desenvolveram interações

contrárias com seus hospedeiros. A proximidade desse grupo de bactérias com os fitopatógenos sugere uma provável correlação entre a função do óperon *mod* em simbiontes e o mecanismo de patogenicidade de *Xanthomonas*.

A necessidade de molibdênio para o crescimento de plantas foi demonstrado pela primeira vez por Arnon e Stout em 1939. Plantas crescendo em uma solução nutriente carente de fontes de molibdênio desenvolveram características fenotípicas que incluíram lesões estampadas nas folhas e morfologia alterada folicular, aonde, a lamela enrolou de forma helicoidal (ARNON e STOUT, 1939) e houve a necrose dos meristemas (HEWITT e BOLLE-JONES, 1952). O único elemento capaz de suprimir este fenótipo foi o molibdênio. Ao contrário disto, não existem relatos de que altas concentrações de molibdênio possam ser tóxicas às plantas (KAISER *et al.*, 2005). Em nosso laboratório estudos realizados em folhas de *Citrus* demonstraram que linhagens de *Xac* mutantes para a proteína ModA, possuem o seu fenótipo de colonização alterado em relação a linhagem selvagem. Estes dados ainda não estão completamente compreendidos, porém sugerem uma adaptação evolutiva de *Xac* no processo de patogênese de *Citrus*. A completa ausência de regulação do óperon em uma bactéria que não possui nitrogenases, nem tampouco é anaeróbica, nos permitiu refletir sobre a possibilidade de que este óperon possa competir com as células da planta pelo metal Mo, causando uma carência e deixando-a vulnerável a colonização bacteriana. Evidências desta provável função são observadas quando analisamos a influência do metal Mo na atividade da molibdoenzima XDH. Responsável pela mobilização e exportação do nitrogênio fixado para fora do nódulo, esta enzima sofre influência da deficiência de molibdênio e sua atividade sofre aumento na presença de fungos fitopatógenos de cereais e legumes. Ainda permanece desconhecida como estas e outras relações de defesas de plantas estão ligadas a presença de molibdênio. Existem pequenas evidências diretas que permitem concluir que elevar os níveis de molibdênio nas plantas resultaria em um decréscimo da doença (KAISER *et al.*, 2005), com exceção de alguns pequenos estudos que indicam que a fertilização com molibdênio pode aumentar a resistência de tomates ao “verticilo mucho” (GRAHAM e STANGOULIS, 2005). As molibdoenzimas estão também envolvidas na síntese de fito-hormônios ABA e Ácido indolilacético (AIA). O passo final na catálise do AIA e a oxidação para obtenção de ABA são dependentes de Moco. Baixos níveis de ABA resultam em plantas de

aspecto murcho favorecendo uma transpiração excessiva (KAISER *et al.*, 2005) o que permitiria aberturas e perda do controle nos estômatos ocasionando a perda de turgidez e permitindo a entrada de *Xac*. Assim, a presença de genes reguladores em bactérias simbiotes poderia ir além da regulação da expressão do óperon, mas também permitiria uma melhor interação simbiote aonde a planta não seria prejudicada.

A grande diversificação sequencial apresentada principalmente em relação ao grupo 3 de bactérias e as demais, assim como na composição do óperon, sugeriu inicialmente haver diferenças estruturais relevantes nas estruturas das proteínas ligadoras de molibdato (ModA) capazes de inferir diferenças funcionais em cada organismo. Predições das estruturas tridimensionais *in silico* auxiliam nas pesquisas *in vitro* sugerindo soluções mais rápidas e econômicas. Uma das claras aplicações dos dados do nosso projeto esteve na colaboração para a resolução da estrutura cristalográfica da ModA de *Xac* (BALAN *et al.*, 2005; BALAN *et al.*, 2008). Durante o processo de modelagem, diversas etapas foram atingidas, e outras forneceram dados relevantes que permitiram colaborar com os projetos de pesquisas do nosso grupo. Entre estes dados, a predição *in silico* do peptídeo sinal da ModA de *Xac*, os efeitos da sua remoção e a inserção de uma cauda de histidina na estrutura da ModA modelada.

Nosso modelo da *modA* de *Xac* é composto por 232 aminoácidos que compõem uma proteína globular com dimensões aproximadas de 33 Å x 48 Å x 62 Å. Como a maioria das outras proteínas ligadoras de substratos a ModA de *Xac* é composta de dois domínios similares com um total de 9 hélices e 11 fitas formando uma estrutura do tipo sanduíche de alfa/beta interligada pela região ligadora de ânions.

A recente estrutura resolvida de *Afu* (HOLLENSTEIN *et al.*, 2007), juntamente com as estruturas da ModA de *Eco* (HU *et al.*, 1997); *Avi* (LAWSON *et al.*, 1998) alinharam os seus carbonos alfa ao do nosso modelo com grande sobreposição espacial, resultando em uma identidade e r.m.s.d ("root mean square deviation") crescente na seguinte ordem: *Afu* (24%) com 1.8 Å; *Avi* (26%) com 1.9 Å e *Eco* (52%) com 1.5 Å, resultados bem similares ao obtido pela estrutura cristalina, que apresentou identidades de 22%, 26% e 52% e um r.m.s.d. de 1.8 Å, 1.6 Å e 1.1 Å para os carbonos alfa, respectivamente. Estes dados refletem o excelente

desempenho do processo de modelagem, que dentre todas as estruturas secundárias descritas no projeto estrutural da ModA de *Xac*, apenas foi incapaz de determinar duas hélices de pequenas proporções em relação à estrutura da ModA (2h5y) de *Xac*. Estas estruturas foram modeladas como alças que praticamente se enovelaram como uma pequena hélice. A semelhança entre os r.m.s.d. das estruturas disponíveis também indicam que, apesar de utilizarmos organismos distintos com baixa similaridade sequencial, suas proteínas adotam conformações cristalinas semelhantes principalmente nas regiões ligadoras. Na verdade, a grande maioria dos desvios concentrou-se em regiões de alças, enquanto que nos trechos relevantes, como a cavidade ligadora, demonstraram perfeita correlação com a estrutural real da proteína.

Reconhecimento molecular, por definição, implica na complementaridade entre a proteína e seu ligante. Os oxigênios do ânion molibdênio estão dispostos em estrutura tetraédrica e são ligados na interface entre os dois domínios da ModA. Estudos com a ModA de *Avi* demonstraram que esta região está completamente desidratada e a molécula de água mais próxima está a uma distância de 6.4 Å (LAWSON *et al.*, 1998). Os resíduos responsáveis pelas interações com o Mo (Ser¹², Ser³⁹, Ala¹²⁵, Val¹⁵² e Tyr¹⁷⁰) foram idênticos entre o nosso modelo e *Eco* e posteriormente a ModA de *Xac*. As únicas diferenças descritas na modelagem, que não foram confirmadas pela cristalografia, são três interações, sendo duas acrescentadas a mais nos resíduos Ala¹²⁵ e Tyr¹⁷⁰ e uma, presente em Ser¹² da ModA 2h5y, que não foi identificada pelo processo de ancoragem durante a modelagem. Apesar disto, acreditamos que os dados gerados foram eficientes o bastante para confiarmos neste processo. Ainda não é possível afirmar com segurança que os parâmetros poderão servir para todas as estruturas de ModA futuras, porém a confiabilidade na determinação da cavidade poderá ser otimizada após maiores testes com a nossa assinatura estrutural do sítio ativo em outras estruturas que serão depositadas.

Análises comparativas de genomas colaboram na elucidação de novas funções e mecanismos de patogênese em diversos genes. Para organismos aonde o óperon *mod* não apresenta genes relacionados à fixação de nitrogênio e tampouco são anaeróbicos, novas funções deverão surgir nos próximos anos. Em todos os três reinos vivos, Zhang e Gladyshev, observaram uma ampla utilização de Mo e

revelaram que mudanças evolutivas na utilização de Mo podem ter influenciado diversos fatores. A relação entre a utilização de Mo e Selênio (Se) em procariontes é um exemplo e sugere a possibilidade de que em procariontes existe a dependência de Mo na utilização de Se. Desta forma, em *Xac*, devem existir funções alternativas para o óperon *mod* que não esteja relacionada à nitrogenases, que estão ausentes. Acreditamos que neste gênero, pode ter ocorrido um processo evolutivo que permitiu a competição pelo metal Mo entre *Xanthomonas* e os seus hospedeiros. Esta competição ocasionaria uma carência do Metal nas plantas que começariam a expressar fenótipos deficientes para a produção de fito-hormônios ABA e Ácido indolilacético (AIA), que permitiriam uma proliferação da infecção do fitopatógeno através dos estômatos ou outro mecanismo de infecção que aproveitasse a ineficácia dos mecanismos de defesa da planta. Estes dados ainda não foram completamente avaliados e compreendidos, porém sinalizam a importância de continuarmos a estudar as possíveis funções desconhecidas deste óperon nestas bactérias, que além de não possuir um mecanismo de regulação, e ser uma bactéria aeróbica restrita, não possui genes relacionados à nitrogenases.

A resolução da estrutura da ModA foi o primeiro passo para uma maior compreensão da função deste óperon no gênero de *Xanthomonas*. A partir dos dados estruturas depositados por Hollestein e colaboradores em 2007 das estruturas das proteínas ModB e ModC, podemos também sugerir projetos futuros de modelagem destes componentes em *Xac* e obter uma maior compreensão da sua função e interação. Observações prévias descritas por Hollestein das regiões denominadas “Gate1” e “Gate2”, indicam uma forte possibilidade de que em *Xac*, a ModA realize interações semelhantes as observadas em *Afu* com ModB.

5.1 Modelagem da OppA

Durante o meu mestrado caracterizamos o óperon *opp* em *Xac* e o comparamos com diversos ortólogos. Sua semelhança com *Streptomyces roseofulvus* (*Sro*) e o seu agrupamento, entre bactérias gram-positivas nas árvores não enraizadas analisadas, sugeriu que este óperon poderia ser o resultado de um

evento de transferência horizontal (MOUTRAN *et al.*, 2004). Todavia nossos dados comprovaram que não há sinais da presença de fagos, transposons ou semelhança entre os códons usage de *Xac* e *Sro* que pudessem confirmar a possível transferência (MOUTRAN *et al.*, 2004). Entre as bactérias do gênero *Xanthomonas*, a presença exclusiva do óperon em *Xac* e de uma mutação pontual no gene *oppDF* (MOUTRAN *et al.*, 2004) poderia inviabilizar todo o sistema de transporte de oligopeptídeos neste organismo. Posteriormente, todas as investidas realizadas pelos membros do nosso laboratório na tentativa de elucidar esta questão foram frustradas tanto na obtenção cristalográfica da estrutura da OppA, quanto por ensaios de captação de peptídeos. O desenvolvimento da modelagem da OppA, assim como, ensaios *in silico* de captação de oligopeptídeos, surgiram como uma alternativa para esclarecer estas dúvidas.

A modelagem da OppA de *Xac*, que não dispõem de PDBs moldes com identidade superior a 19%, necessitou de técnicas avançadas. Valores de identidades inferiores a 25% são caracterizados como pertencentes a região “twin light zone”, com baixa cobertura seqüencial e quando não é impossível, é extremamente trabalhosa. Através da modelagem com múltiplos PDBs (1xoc e 1jev) conseguimos uma cobertura de praticamente toda a área da OppA de *Xac*. Nosso modelo superou as expectativas e a qualidade estereoquímica apresentada pela estrutura da OppA de *Xac*, revelou um modelo com 100% dos seus resíduos em regiões favoráveis, sendo distribuídos em 88,37% (395) dos seus resíduos nas regiões mais favoráveis, 11,63% (52) dos resíduos dentro das regiões permitidas, com nenhum resíduo nas regiões não permitidas ou desfavoráveis. O modelo da OppA de *Xac* possui uma estrutura globular com dimensões aproximadas de 85 Å x 67 Å x 65 Å, do tipo sanduiche de α/β , semelhante ao descrito para seus ortólogos. A cavidade ligadora da OppA localiza-se na fenda formada pelos dois domínios similares, denominados I e III. Observando o domínio I pudemos dividi-lo em dois sub-domínios (I e II), semelhante ao que é descrito para AppA de *Bsu* (LEVDIKOV *et al.*, 2005). O subdomínio I é formado por três sequências descontínuas.

Os ortólogos da OppA de *Xac* apresentam identidades sequenciais de 18,5% (*Bsu*) e 18% (*Stm*). A reduzida similaridade sequencial, também é observada na estrutura depositada recentemente da bactéria gram-negativa *Tma*, que apresentou 12,5% de identidade com a OppA de *Xac*.

Apesar das diferenças sequenciais, o alinhamento estrutural obtido entre os C^α revelou que a OppA de *Xac* é estruturalmente conservada quando comparada aos ortólogos *Bsu* (r.m.s.d. 1.4 Å), *Stm* (r.m.s.d. 1,8 Å) e contem maior divergência com a OppA de *Tma* (r.m.s.d. de 3,8 Å). As principais diferenças observadas entre *Xac*, *Bsu* e *Stm* concentram-se em duas alças extras presentes na OppA de *Xac*, que não se encontram alinhados estruturalmente ou sequencialmente. Esta diferença tanto pode ser real como pode ser um artifício natural do processo de modelagem, quando regiões descobertas por informações estruturais dos ortólogos moldes são colocadas em alças para minimizar os desvios e permitir continuidade do processo de predição estrutural no segmento seguinte. O alinhamento estrutural das três estruturas das OppAs, também permitiu a identificação de uma alça extra presente somente em *Stm*.

A capacidade de captar peptídeos é uma propriedade essencial para a sobrevivência das bactérias em diferentes ambientes (PAYNE e GILBARG, 1968; HILLES e HIGGINS, 1986). As proteínas ligadoras de oligopeptídeos (OppA), expressas por diferentes espécies de bactérias, possuem grande afinidade mas baixa especificidade a substratos peptídicos (GUYER *et al.*, 1985; HILLES e HIGGINS, 1986). Diferenças nas afinidades entre os diversos parólogos em *Bbu*, sugerem que podem existir distintas funções neste organismo para cada uma das OppAs (WANG *et al.*, 2005), enquanto que a existência de uma única cópia em *Xac* sugere que sua capacidade de captação seja mais ampla e menos específica. Nossos ensaios de ancoragem, assim como os dados estruturais da proteína OppA de *Xac*, permitirão novas diretrizes no direcionamento dos estudos *in vitro*.

Dados dos alinhamentos estruturais e sequenciais entre *Xac* e *Bsu*, indicavam fortemente que o comportamento da OppA de *Xac*, quanto a captação de peptídeo, deveria ser semelhante ao observado em gram-positivas, que possuem uma afinidade maior por cadeias longas de peptídeos. Os dados extraídos dos ensaios de ancoragem surpreendeu-nos mais uma vez revelando uma proteína com afinidade por oligopeptídeos descritos para gram-negativas, em especial as descritas para o ortólogo em *Stm*. Diferenças na especificidade entre as OppAs de *Stm* e *Lla* foram reportadas aonde a OppA de *Stm* é capaz de ligar di e tripeptídeos e com uma baixa afinidade em relação a pentapeptídeos do que por tetrapeptídeos, enquanto a OppA de *Lla* tem uma maior afinidade por penta e hexapeptídeos

(TYNKKYNEN *et al.*, 1993; JUILLARD *et al.*, 1998). Testamos estas diferenças na OppA de *Xac*. Estudos do mecanismo de captação realizados com OppA de *Lla* e nonapeptídeos, indicaram que os seis primeiros resíduos ficam imersos na cavidade formada pelos domínios, enquanto que a porção restante interage com a superfície externa da proteína ligadora (DETMERS *et al.*, 2001). Já na OppA de *Xac*, estes mesmos peptídeos acomodaram-se externamente com grande dificuldade, muitas vezes em regiões não permitidas da proteína. Em *Lla*, a translocação de peptídeos está sob controle de dois principais fatores: a carga dos peptídeos e o seu peso molecular, sendo que peptídeos hidrofóbicos possuem uma maior preferência em relação a resíduos hidrofílicos (JUILLARD *et al.*, 1998). Na OppA de *Xac*, observamos esta preferência por resíduos hidrofóbicos e obtivemos os melhores resultados de ancoragem na presença de aminoácidos neutros (serina e Treonina), seguido de aminoácidos apolares (prolina, valina e glicina) que possuem características hidrofóbicas com hidrocarbonetos em suas cadeias laterais. Estas características, juntamente com a presença de moléculas de água, ocultariam cargas eventualmente presentes no ligante, e explicaria a baixa especificidade do sistema já que peptídeos de composição diversa podem se acomodar na proteína (TAME *et al.*, 1994).

O ligante peptídico ficou completamente embebido na cavidade da proteína OppA de *Xac* quando a amplitude da extensão das cadeias peptídicas esteve na faixa de 3 a 5 resíduos. A cadeia lateral mostrou-se tão importante que permitiu que a OppA de *Xac* se ligasse a heptapeptídeos com afinidade semelhante ao observado para pentapeptídeos, quando estes, em sua composição, apresentavam apenas os resíduos citados acima. Desta forma, concluímos que a afinidade da OppA de *Xac* é influenciada por dois fatores: a cadeia lateral do resíduo e a extensão da cadeia peptídica. Em ambos, percebemos que o fator espaço influencia mais do que fatores químicos, já que na cavidade dos ortólogos das OppAs não existe um consenso sobre quais resíduos interagem com o ligante, mas sim uma gama diversa de resíduos que circundam o ligante e que poderiam interagir conforme a sua composição química. Esta característica permite que as OppAs interajam com seus substratos peptídicos com uma alta afinidade e baixa especificidade e para a OppA de *Xac* sua afinidade crescente por oligopeptídeos com 3 a 5 resíduos de extensão correlaciona-se perfeitamente com a barreira

promovida pela presença de uma membrana externa típica de bactéria gram-negativas.

Certamente, todo o conhecimento obtido durante a modelagem da OppA, servirá de subsídio para a obtenção de ensaios de captação de oligopeptídeos *in vivo* e para a modelagem dos componentes formadores do poro e energizantes, que possuem atualmente um número maior de proteínas moldes pertencentes ao sistema ABC de transporte como os captadores de maltose, B12, metionina, molibdênio. A modelagem dos componentes de membrana, será possível devido a elevada conservação entre estes componentes nas diversas famílias de substratos dentro dos transportadores ABC dados já descritos por Kadaba e colaboradores (2008) nos alinhamentos estruturais entre MetI, ModB MalG e MalF. Estamos seguros destas possibilidades, inclusive pelos resultados preliminares que possuímos com os componentes *oppBCDF*. A modelagem destes componentes, assim como ensaios de ancoragem entre a OppA e as permeases, permitirão compreender o mecanismo de interação entre todos os componentes do óperon. De forma semelhante, ensaios de ancoragem entre os componentes energizantes e adenosina trifosfato (ATP) permitirão inferir com maior confiabilidade se o óperon encontra-se funcional ou seria uma relíquia evolutiva em *Xac*.

6 CONCLUSÃO

O projeto permitiu elucidar as estruturas espaciais, por meio de técnicas de modelagem molecular, de duas proteínas ligadoras pertencentes a sistemas de transporte do tipo ABC do fitopatógeno *Xac*: as proteínas ModA e OppA

A proteína ModA, codificada pelo primeiro códon (*modA*) do óperon *mod*, foi encontrada em todas as espécies de *Xanthomonas* com genomas seqüenciados. A sequência de aminoácidos de ModA revela-se conservada (similaridade acima de 80%) nas diferentes espécies de *Xanthomonas*. Foi possível ainda definir 3 grupos filogenéticos a partir de análises de similaridade entre proteínas ModA encontradas em bactérias que interagem com plantas, bactérias de solo e bactérias entéricas. O modelo estrutural da proteína ModA permitiu a identificação dos resíduos que compõem a cavidade ligadora de molibdato. A precisão do modelo estrutural da proteína ModA foi confirmada com a determinação experimental da estrutura cristalográfica da proteína pelo grupo.

Com o modelo estrutural da proteína OppA de *Xac* foi possível realizar análises de afinidade com diferentes oligopeptídeos. Os resultados confirmam que a proteína OppA tem maior afinidade por oligopeptídeos com 3 a 5 resíduos de extensão. A afinidade da proteína aos substratos peptídicos depende também da composição de aminoácidos sendo que a presença de resíduos apolares ou polares neutros de cadeia curta aumentou a afinidade e o reconhecimento da proteína pelo oligopeptídeo.

Os modelos gerados no presente trabalho mostram-se particularmente úteis no direcionamento de estudos estruturais e funcionais dos sistemas de transporte Mod e Opp de *Xac*. Além disto, a presente tese representa uma contribuição importante para o conhecimento mais detalhado da fisiologia nutricional de *Xac*, um fitopatógeno com grande relevância econômica no país e no mundo.

REFERÊNCIAS*

AMES, G. F. Bacterial periplasmic transport systems: structure, mechanism, and evolution. **Annu. Rev. Biochem.**, v.55, p. 397-425, 1986.

AMES, G. F.; MIMURA, C. S.; HOLBROOK, S. R.; SHYAMALA, V. Traffic ATPases: a superfamily of transport proteins operating from *Escherichia coli* to humans. **Adv. Enzymol. Relat. Areas Mol. Biol.**, v. 65, p. 1-47, 1992.

ARNON, D. I.; STOUT, P. R. The essentiality of certain elements in minute quantity for plants with special reference to copper. **Plant Physiol.**, v. 14, p. 371-375, 1939.

BALAN, A.; SANTACRUZ-PÉREZ, C.; MOUTRAN, A.; FERREIRA, L. C.; NESHICH, G.; GONÇALVES, B. J. A. Crystallographic structure and substrate-binding interactions of the molybdate-binding protein of the phytopathogen *Xanthomonas axonopodis* pv. *citri*. **Biochim Biophys Acta.**, v. 1784 p. 393-399, 2008.

BALAN, A.; SOUZA, C. S.; MOUTRAN, A.; FERREIRA, R. C. C.; FRANCO, C. S.; RAMOS, C.; FERREIRA, L. C. S. Purification and in vitro characterization of the maltose-binding protein of the plant pathogen *Xanthomonas citri*. **Prot. Exp. Purif.** v. 2, p. 103-110, 2005.

BENDTSEN, J. D.; NIELSEN, H.; VON HEIJNE, G.; BRUNAK, S. Improved prediction of signal peptides: SignalP 3.0. **J. Mol. Biol.**, v. 340, p. 783-795, 2004.

BERENDSEN, H. J. C.; VAN DER SPOEL, D.; VAN DRUNEN, R. GROMACS: a new message-passing parallel molecular dynamics implementation. **Comp. Phys. Commun.**, v. 91, p. 43-56, 1995.

BEVERIDGE, T. J.; GRAHAM, L. L. Surface layers of bacteria. **Microbiol. Rev.**, v. 55, p. 684-705, 1991.

---

* De acordo com:

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6023**: Informação e documentação: referências: elaboração. Rio de Janeiro, 2002.

BLUNDELL, T. L.; CARNEY, D.; GARDNER, S.; HAYES, F.; HOWLIN, B.; HUBBARD, T.; OVERINGTON, J.; SINGH, D.; SIBANDA, B. L.; SUTCLIFFE, M. Knowledge-based protein modelling and design; 18th Sir Hans Krebs **Eur. J. Biochem.**, v. 173, p. 513-520, 1988.

BRAIBANT, M.; GILOT, P.; CONTENT, J. The ATP binding cassette (ABC) transport systems of *Mycobacterium tuberculosis*. **FEMS Microb. Rev.**, v. 24, p. 449-467, 2000.

BABU, S. Y.; EALICK, S. E.; BUGG, C.E.; MONTGOMERY, J. A; SECRIST, J. Structure-based design of inhibitors of purine nucleoside phosphorylase **Acta Cryst.**, v. 51, p. 529-535, 1995.

CDART (Conserved Domain Architecture Retrieval Tool) Available from: <<http://www.ncbi.nlm.nih.gov/>>. Acesso em: 10 mar. 2004.

CHAKRABURTTY, K. Functional interaction of yeast elongation factor 3 with yeast ribosomes. **Int. J. Biochem. Cell Biol.**, v.31, p. 163-173, 1999.

CHAN, J. W. Y. F.; GOODWIN, P. H. The molecular genetics of virulence of *Xanthomonas campestris*. **Biotechnol. Adv.**, v. 17, p. 489-508, 1999.

COHEN, S. **A guide to the polyamines**. New York: Oxford University Press, 1998. 543p.

CUFF, J. A.; CLAMP, M. E.; SIDDIQUI, A. S.; FINLAY, M.; BARTON, G. J. Jpred: a consensus secondary structure prediction server, **Bioinformatics**, v. 14, p. 892-893, 1998.

COPLIN, D. L.; COOK, D. Molecular genetics of extracellular polysaccharide biosynthesis in vascular phytopathogenic bacteria. **Mol. Plant Microbe Interact.**, v. 3, p. 271-279, 1990.

CSERZO, M.; WALLIN, E.; SIMON, I.; VON HEIJNE, G.; ELOFSSON, A. Prediction of transmembrane alpha-helices in procariotic membrane proteins: the dense alignment surface method; **Prot. Eng.**, v. 10, p. 673-676, 1997

DA SILVA, A. C. R.; FERRO, J. A.; REINACH, F. C.; FARAH, C. S.; FURLAN, L. R.; QUAGGIO, R. B.; MONTEIRO-VITORELLO, C. B.; VAN SLUYS, M. A.; ALMEIDA JR.; ALVES, L. M. C.; DO AMARAL, A. M.; BERTOLINI, M. C.; CAMARGO, L. E. A.; CAMAROTTE, G.; CANNAVAN, F.; CARDOZO, J.; CHAMBERGO, F.; CIAPINA, L. P.; CICARELLI, R. M. B.; COUTINHO, L. L.; CURSINO-SANTOS, J. R.; EL-DORRY, H.; FARIA, J. B.; FERREIRA, A. J. S.; FERREIRA, R. C. C.; FERRO, M. I. T.; FORMIGHIERI, E. F.; FRANCO, M. C.; GREGGIO, C. C.; GRUBER, A.; KATSUYAMA, A. M.; KISHI, L. T.; LEITE JR, R. P.; LEMOS, E. G. M.; LEMOS, M. V. F.; LOCALI, E. C.; MACHADO, M. A.; MADEIRA, A. M. B. N.; MARTINEZ-ROSSI, N. M.; MARTINS, E. C.; MEIDANIS, J.; MENCK, C. F. M.; MIYAKI, C. Y.; MOON, D. H.; MOREIRA, L. M.; NOVO, M. T. M.; OKURA, V. K.; OLIVEIRA, M. C.; OLIVEIRA, V. R.; PEREIRA JR., H. A.; ROSSI, A.; SENA, J. A. D.; SILVA, C.; DE SOUZA, R. F.; SPINOLA, L. A. F.; TAKITA, M. A.; TAMURA, R. E.; TEIXEIRA, E. C.; TEZZA, R. I. D.; TRINDADE DOS SANTOS, M.; TRUFFI, D.; TSAI, S. M.; WHITE, F. F.; SETUBAL, J. C.; KITAJIMA, J. P. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. **Nature**, v. 417, p. 459-63, 2002.

DASSA, E. ABC transport. In: _____. (Ed.). **Encyclopedia of microbiology**. New York: Academic Press, 2000 p. 1-12.

DAVIDSON, A. L.; MALONEY P. C. ABC transporters: how small machines do a big job. **Trends Microbiol.**, v. 15, p. 448-455 2007.

DELANO, W. L. **The PyMOL molecular graphics system on world wide web**. Available from: <http://www.pymol.org>. Accessed: 13 jun. 2005.

DETMERS, F. J.; LANFERMEIJER, F. C.; POOLMAN, B. Peptides and ATP binding cassette peptide transporters. **Res. Microbiol.**, v. 152, p. 245-258, 2001.

DOOLITTLE, R. F.; JOHNSON, M. S.; HUSAIN, I.; VAN HOUTEN, B.; THOMAS, D. C.; SANCAR, A. Domainal evolution of a prokaryotic DNA repair protein and its relationship to active-transport proteins. **Nature**, v. 323, p. 451-453, 1986.

ESTANISLAU, M. L. L. *et al.* **Informe agropecuário**, v. 22, p. 8-20, 2001.

ExPASy Molecular Biology Available from: <<http://www.expasy.org/>>. Acesso em: 10 mar. 2004.

FENNO, J. C.; TAMURA, M.; HANNAM, P. M.; WONG, G. W. K.; CHAN, R. A.; MCBRIDE, B. C. Identification of a *Treponema denticola* OppA homologue that binds host proteins present in the subgingival environment. **Infect. Immun.**, v. 68, p. 1884-1892, 2000.

FERREIRA, A. J. S. **Sequenciamento de parte do genoma de *Xylella fastidiosa* e análise do metabolismo de carboidratos no genoma anotado**. 116 f. Dissertação (Mestrado) – Instituto de Química, Universidade de São Paulo, São Paulo, 2000.

FUQUA, W. C.; WINANS, S. C.; GREENBERG, E. P. Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. **J. Bacteriol.**, v.176, p. 269-275, 1994.

GIBAS, C. **Desenvolvendo Bioinformática**. Rio de Janeiro: Campus, 2001. 440 p.

GIBSON, M. M.; PRICE, M.; HIGGINS, C. F. Genetic characterization and molecular cloning of the tripeptide permease (tpp) genes of *Salmonella typhimurium*. **J. Bacteriol.**, v. 160, p. 122-130, 1984.

GILSON, E.; NIKAIDO, H.; HOFNUNG, M. Sequence of the *malK* gene in *Escherichia coli* K12. **Nucleic Acids Res.**, v. 10, p. 7449-7458, 1982.

GOODELL, E. W.; HIGGINS, C. F. Uptake of cell wall peptides by *Salmonella typhimurium* and *Escherichia coli*. **J. Bacteriol.**, v. 169, p. 3861-3865, 1987.

GRAHAM R. D.; STANGOULIS J. R. C. Molybdenum and disease. In: DANTOFF, L.; ELMER, W.; HUBER, D. (Ed.). **Mineral nutrition and plant diseases**. St Paul: APS Press, 2005.

Greer, J. Comparative modeling methods: application to the family of the mammalian serine proteases. **Proteins**, v. 7, p. 317-334, 1990.

GRUNDEN, A. M.; RAY, R. M.; ROSENTEL, J. K.; HEALY, F. G.; SHANMUGAM, K. T. Repression of the *Escherichia coli modABCD* (molybdate transport) operon by ModE. **J. Bacteriol.**, v. 178, p. 735-744, 1996.

GRUNDEN, A. M.; SHANMUGAM, K. T. Molybdate transport and regulation in bacteria. **Arch. Microbiol.**, v. 168, p. 345-354, 1997.

GRUNDEN, A. M.; SELF, W. T.; VILLAIN, M.; BLALOCK, J. E.; SHANMUGAM, K. T. An analysis of the binding of repressor protein ModE to *modABCD* (Molybdate transport) operator/ promoter DNA of *Escherichia coli*, **J. Biol. Chem.**, v. 274 p. 24308–24315, 1999.

GUO, X.; FLORES, M.; MAVINGUI, P.; FUENTES, S. I.; HERNÁNDEZ, G.; DÁVILA, G.; PALACIOS, R. Natural genomic design in *Sinorhizobium meliloti*: novel genomic architectures. **Genome Res.**, v. 13, p. 1810–1817, 2003.

GUYER, C. A.; MORGAN, D. G.; OSHEROFF, N; STAROS, J. V. Purification and characterization of a periplasmic oligopeptide binding protein from *Escherichia coli*. **J. Biol. Chem.**, v. 260, p. 10812-10818, 1985

HEWITT, E. J.; BOLLE-JONES, E. W. Molybdenum as a plant nutrient. The influence of molybdenum on the growth of some *Brassica* crops in sand cultures. **J. Horticult. Sci.**, v.27, p. 245-256, 1952.

HIGGINS, C. F. ABC transporters: physiology, structure and mechanism – an overview. **Res. Microbiol.**, v. 152, p. 205-210, 2001.

HIGGINS, C.F. ABC transporters: from microorganisms to man. **Annu. Rev. Cell. Biol.**, v. 8, p. 67-113, 1992.

HIGGINS, C. F.; HARDIE, M. M.; JAMIESON, D.; POWELL, L. M. Genetic map of the *opp* (Oligopeptide permease) locus of *Salmonella typhimurium*. **J. Bacteriol.**, v. 153, p. 830-836, 1983.

HILLES I. D.; HIGGINS, C. F. Peptide uptake by *Salmonella typhimurium* the periplasmic oligopeptide-binding protein. **Eur. J. Biochem.**, v. 158, p. 561-567, 1986.

HILLE, R. Molybdenum and tungsten in biology, **Trends Biochem. Sci.**, v. 27, p. 360–367, 2002.

HOLLENSTEIN, K.; DAWSON, R. J.; LOCHER, K. P. Structure and mechanism of ABC transporter proteins. **Curr. Opin. Struct. Biol.**, v. 17, p. 412-418, 2007.

HU, Y.; RECH, S.; GUNSALUS, R. P.; REES, D. C. Crystal structure of the molybdate binding protein ModA. **Nat. Struct. Biol.**, v. 4, p. 703-707, 1997.

IMPERIAL, J.; HADI, M.; AMY, N. K. Molybdate binding by ModA, the periplasmic component of the *Escherichia coli mod* molybdate transport system. **Biochem. Biophys. Acta**, v. 1370, p. 337-346, 1998.

JOHANN, S.; HINTON, S. M. Cloning and nucleotide sequence of the *chlD* locus. **J. Bacteriol.**, v. 169, p. 1911-1916, 1987.

KADABA, N. S.; KAISER, J. T.; JOHNSON, E.; LEE, A.; REES, D. C. The high-affinity *E. coli* methionine ABC transporter: structure and allosteric regulation. **Science**, v. 321, p. 250-253, 2008.

KAISER, B. N.; GRIDLEY, K. L.; BRADY, J. N.; PHILLIPS, T.; TYERMAN, S. D. The role of molybdenum in agricultural plant production. **Ann. Bot.**, v. 96, 745-754, 2005.

KEGG. Available from: <<http://www.genome.ad.jp>>. Acesso em: 10 mar. 2004.

KISKER, C.; SCHINDELIN, H.; REES, D. C. Molybdenum-cofactor-containing enzymes: structure and mechanism. **Annu. Rev. Biochem.**, v. 66, p. 233-267, 1997.

KOONIN, E. V.; TATUSOV, R. L.; GALPERIN, M. Y. Beyond complete genome: from sequence to structure and function. **Curr. Opin. Struct. Biol.**, v. 8, p. 355-363, 1998.

KUPER, J.; BERSTENHORST, S. M.; VÖDISCH, B.; MENDEL, R. R.; SCHWARZ, G.; BOXER, D. H. In vivo detection of molybdate-binding proteins using a competition assay with ModE in *Escherichia coli*. **FEMS Microbiol. Lett.**, v. 218, p. 187-193, 2003.

LASKOWSKI, R. A. From protein structure to biochemical function. **J. Struct. Funct. Genomics**, v. 4, p. 167-177, 2003.

LASKOWSKI, R. A.; MACARTHUR, M. W.; MOSS, D. S.; THORNTON, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. **J. Appl. Cryst.**, v. 26, p. 283-291, 1993.

LAWSON, D. M.; WILLIAMS, C. E. M.; MITCHENALL, L. A.; PAU, R. N. Ligand size is a major determinant of specificity in periplasmic oxyanion-binding proteins: the 1.2m Å resolution crystal structure of *Azotobacter vinelandii* ModA. **Structure**, v. 6, p. 1529-1539, 1998.

LAWSON, D. M.; WILLIAMS, C. E.; WHITE, D. J.; CHOAY, A. P.; MITCHENALL, L. A.; PAU, R. N. Proteins ligand for molybdate, specificity and charge stabilization at the anion-binding sites of periplasmic and intracellular molybdate-binding proteins of *Azotobacter vinelandii*. **J. Chem. Soc. Dalton Trans.**, v. 21, p. 3981-3984, 1997.

LEE B. M.; PARK Y. J.; PARK D. S.; KANG H. W.; KIM J. G.; SONG E. S.; PARK I. C.; YOON U. H.; HAHN J. H.; KOO B. S.; LEE G. B.; KIM H.; PARK H. S.; YOON K. O.; KIM J. H.; JUNG C. H.; KOH N. H.; SEO J. S.; GO S. J. The genome sequence of *Xanthomonas oryzae pathovar oryzae* KACC10331, the bacterial blight pathogen of rice. **Nucleic Acids Res.**, v. 33 p. 577-586, 2005.

LEVEDIKOV, V. M.; BLAGOVA, E. V.; BRANNIGAN, J. A.; WRIGHT, L.; VAGIN, A. A.; WILKINSON, A. J. The structure of the oligopeptide-binding protein, AppA, from *Bacillus subtilis* in complex with a nonapeptide. **J. Mol. Biol.**, v. 345, p. 879-892, 2005.

LINTON, K. J.; HIGGINS, C. F. The *Escherichia coli* ATP-binding cassette (ABC) proteins. **Mol. Microbiol.**, v. 28, p. 5-13, 1998.

LUGTENBERG, B.; VAN ALPHEN, L. Molecular architecture and functioning of the outer membrane of *Escherichia coli* and other gram-negative bacteria. **Biochim. Biophys. Acta**, v. 737, p. 51-115, 1983.

MAKINO, D. L. **Estudos estruturais de DM43: um inibidor de metaloprotease de veneno de serpente extraído do soro do gambá *Didelphis marsupialis*** 151 f. Dissertação (Mestrado) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2000.

MANSON, M. D.; BLANK, V.; BRADE, G.; HIGGINS, C. F. Peptide chemotaxis in *E. coli* involves the Tap signal transducer and the dipeptide permease. **Nature**, v. 321, p. 253-256, 1986.

MAUPIN-FURLOW, J. A.; ROSENTEL, J. K.; LEE, J. H.; DEPPENMEIER, U.; GUNSALUS, R. P.; SHANMUGAM, K. T. Genetic analysis of the *modABCD* (molybdate transport) operon of *Escherichia coli*. **J. Bacteriol.**, v. 177, p. 4851-4856, 1995.

MCDOWELL, L. R. Copper and molybdenum. In: CUNHA, T. J. (Ed.). **Minerals in animal and human nutrition**. San Diego: Academic Press, 1992. p. 176-204.

MONNET, V. Bacterial oligopeptide-binding proteins. **Cell Mol. Life Sci.**, v. 60, p. 2100-2114, 2003.

MOUNCEY, N. J.; MITCHENALL, L. A.; PAU, R. N. Mutational analysis of genes of the *mod* locus involved in molybdenum transport, homeostasis, and processing in *Azotobacter vinelandii*. **J. Bacteriol.**, v. 177, p. 5294-5302, 1995.

MOUREZ, M.; HOFNUNG, J.; DASSA, E. Subunit interactions in ABC transporters: a conserved sequence in hydrophobic membrane proteins of periplasmic permeases defines an important site of interaction with the ATPase subunits. **EMBO J.**, v. 4, p. 2287-2293, 1997.

MOUTRAN, A.; QUAGGIO, R. B.; BALAN, A.; FERREIRA, L. C. S.; FERREIRA, R. C. C. The oligopeptide permease (Opp) of the plant pathogen *Xanthomonas axonopodis* pv citri. **Curr. Microbiol.**, v. 48, p. 354-359, 2004.

MORRIS, A. L.; MACARTHUR, M. W.; HUTCHINSON, E. G.; THORNTON, J. M. Stereochemical quality of protein structure coordinates. **Proteins**, v. 12, p. 345-364, 1992.

NCBI National Center of Bioinformatics Available from: <<http://www.ncbi.nlm.nih.gov/>>. Acesso em: 10 mar. 2004.

NESHICH, G.; TOGAWA, R. C.; MANCINI, A. L.; KUSER, P. R. STING Millennium: A web-based suite of programs for comprehensive and simultaneous analysis of protein structure and sequence. **Nucleic Acids Res.**, v. 31, p. 3386-3392, 2003.

NEVES, E. M.; DAYOUB, M.; DRAGONE, D. S.; NEVES M. F. Citricultura brasileira: efeitos econômicos-financeiros, **Rev. Bras. Frutic.**, v. 23, p. 432-436, 2000.

NICHOLAS, K. B.; NICHOLAS, H. B.; DEERFIELD, D. W. II. GeneDoc: analysis and visualization of genetic variation, **EMBNEW NEWS**, v. 4, p. 14, 1997.

PAYNE, J. W.; GILVARG, C. Size restriction on peptide utilization in *Escherichia coli*. **J. Biol. Chem.**, v. 243, p. 6291-6299, 1968.

PDB. Available from: <http://www.rcsb.org/pdb/>. Acesso em: 10 mar. 2004.

PICON, A.; KUNJI, E.R.S.; LANFERMEIJER, F. C.; KONINGS, W. N.; POOLMAN, B. Specificity mutants of the binding protein of the oligopeptide transport system of *Lactococcus lactis*. **J. Bacteriol.**, v. 182, p. 1600-1608, 2000.

PICON, A.; WELY, K. H. M. V. Peptide binding to the *Bacillus subtilis* oligopeptide-binding proteins OppA and AppA. **Mol. Biol. Today**, v. 2, p. 21-25, 2001.

POPLAWSKY, A. R.; URBAN, S. C.; CHUN, W. Biological role of xanthomonadin pigments in *Xanthomonas campestris* pv. *campestris*. **Appl. Environ. Microbiol.**, v. 66, p. 5123-5127, 2000.

POXTRON, I. R. Prokaryote envelope diversity. **J. Appl. Bacteriol.**, v. 74, p. 1-11, 1993.

RAJAGOPALAN, K. V.; JOHNSON, J. L. The pterin molybdenum cofactors. **J. Biol. Chem.**, v. 267, p. 10199-10202, 1992

RAMACHANDRAN, G. N.; SASISEKHARAN, V. Conformation of polypeptides and proteins. **Adv. Prot. Chem.**, v. 23, p. 283-438, 1968.

RECH, S.; WOLIN, C.; GUNSALUS, R. P. Properties of the periplasmic ModA molybdate-binding protein of *Escherichia coli*. **J. Biol. Chem.**, v. 271, p. 2557-2562, 1995.

SALI, A.; BLUNDELL, T. L. Comparative protein modelling by satisfaction of spatial restraints. **J. Mol. Biol.**, v. 234, p. 779-815, 1993.

SÁNCHEZ, R.; SALI, A. Advances in comparative protein-structure modelling **Curr. Opin. Struct. Biol.**, v. 7, p. 206-214, 1997.

SÁNCHEZ, R.; PIEPER, U.; MELO, F.; ESWAR, N.; MARTÍ-RENOM, M. A.; MADHUSUDHAN, M. S.; MIRKOVIĆ, N.; SALI, A. Protein structure modeling for structural genomics. **Nat. Struct. Biol.**, v.7, p.986-990, 2000.

SANTACRUZ, C. P.; BALAN, A.; FERREIRA, L. C. S.; BARBOSA, J. A.; Crystallization, data collection and phasing of the molybdate-binding protein of the phytopathogen *Xanthomonas axonopodis* pv. *citri*, **Acta Crystallogr. Sect. F. Struct. Biol. Crystal. Commun.**, v. 62, p. 289–291, 2006.

SAURIN, W.; KÖSTER, W.; DASSA, E. Bacterial binding protein-dependent permeases: characterization of distinctive signatures for functionally related integral cytoplasmic proteins. **Mol. Microbiol.**, v. 12, p. 993-1004, 1994.

SELF, W. T.; GRUNDEN, A. M.; HASONA, A.; SHANMUGAM, K. T. Molybdate transport. **Res. Microbiol.**, v. 152 p. 311–321, 2001.

SHINDYALOV, I. N.; BOURNE, P. E. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. **Protein Eng.**, v. 11, p. 739-747, 1998.

SHUMAN, H. A. The genetics of active transport in bacteria. **Annu. Rev. Genet.**, v. 21, p. 155-177, 1987.

SLEYTR, U. B.; MESSNER, D. P.; SÁRA, M. Crystalline bacterial cell surface layers. **Mol. Microbiol.**, v.10, p. 911-916, 1993.

TAME, J. R. H.; DODSON, E. J.; MURSHUDOV, G. N.; HIGGINS, C. F. The crystal structures of the oligopeptide-binding protein OppA complexed with tripeptide and tetrapeptide ligands. **Structure**, v. 3, p. 1395-1406, 1995.

TAME, J. R. H.; WILKINSON, A. J.; SLEIGH, S. H.; LADBURY, J. E. the role of water in sequence-independent ligand binding by an oligopeptide transport protein. **Nat. Struct. Biol.**, v. 3, p. 998-1001, 1996.

TAYLOR, D. L.; WARD, P. N.; RAPIER, C. D.; LEIGH, J. A.; BOWLER, L. D. Identification of a differentially expressed oligopeptide binding protein (OppA2) in *Streptococcus uberis* by representational difference analysis of cDNA. **J. Bacteriol.**, v. 185, p. 5210-5219, 2003.

THOMPSON, J. D.; GIBSON, T. J.; PLEWNIAC, F.; JEANMOUGIN, F.; HIGGINS, D. G. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. **Nucleic Acids Res.**, v. 25, p. 4876-4882, 1997.

TOMII, K. E; KANEHISA, M. A comparative analyses of ABC transporters in complete microbial genomes. **Genome Res.**, v. 8, p. 1048 –1059, 1998.

TOVCHIGRECHKO A.; VAKSER I. A. Development and testing of an automated approach to protein docking. **Proteins**, v. 60, p. 296-301, 2005.

TWING, W. M.; SWINGS, J. G. The host of *Xanthomonas*. In: SWINGS, T; EIVEROLO, E. L. (Ed.). **Xanthomonas**. London: Chapman & Hall, 1993. p. 51-54.

VAN DER SPOEL, D.; LINDAHL, E.; HESS, B.; GROENHOF, G.; MARK, A. E.; BERENDSEN, H. J. C. GROMACS: Fast, Flexible, and Free. **J. Comput. Chem.**, v. 26, p. 1701-1718, 2005.

VAUTERIN, L.; YANG, P.; ALVAREZ, A.; TAKIKAWA, Y.; ROTH, D. A.; VIDAVER, A. K.; STALL, R. E.; KERSTERS, K.; SWINGS, J. Identification of non-pathogenic *Xanthomonas* strains associated with plants. **Syst. Appl. Microbiol.**, v. 19, p. 96-105, 1996.

WANG, C. H.; LIN, C. Y.; LUO, Y. H.; TSAI, P. J.; LIN, Y. S.; LIN, M. T.; CHUANG, W. J.; LIU, C. C. and WU, J. J. Effects of oligopeptide permease in group A Streptococcal infection. **Infect. Immun.**, v. 73, p. 2881-2890, 2005.

WILLIAMS, R. J. P.; FRÁUSTO DA SILVA, J. J. R. The involvement of molybdenum in life. **Biochem. Biophys. Res. Commun.**, v. 292, p. 293-299, 2002.

YOUNG, J.; HOLLAND, I. B. ABC transporters: bacterial exporters-revisited five years on. **Biochim. Biophys. Acta**, v. 1461, p. 177-200, 1999.

ZHANG, Y.; GLADYSHEV, V. N. Molybdoproteomes and evolution of molybdenum utilization. **J. Mol. Biol.**, v. 379, p. 881-899, 2008.

ANEXOS A - Structural model and ligand interactions of the *Xanthomonas axonopodis* pv. *citri* oligopeptide-binding protein

ANEXOS B - Crystallographic structure and substrate-binding interactions of the molybdate-binding protein of the phytopathogen *Xanthomonas axonopodis* pv. *citri*

ANEXOS C - The molybdate-binding protein (ModA) of the plant pathogen *Xanthomonas axonopodis* pv. *citri*

ANEXOS D - Purification and in vitro characterization of the maltose-binding protein of the plant pathogen *Xanthomonas citri*

ANEXOS E – Estágios Realizados