

UNIVERSIDADE ESTADUAL DE PONTA GROSSA
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA APLICADA

AGNES THIANE PEREIRA MACHADO

DETERMINAÇÃO E REFINAMENTO DE ESTRUTURAS TRIDIMENSIONAIS DA
PROTEÍNA SurE de *Xylella fastidiosa*

PONTA GROSSA
2013

AGNES THIANE PEREIRA MACHADO

DETERMINAÇÃO E REFINAMENTO DE ESTRUTURAS TRIDIMENSIONAIS DA
PROTEÍNA SurE de *Xylella fastidiosa*

Dissertação apresentada para obtenção do
título de Mestre em Química Aplicada no
Programa de Pós-Graduação em Química
Aplicada da Universidade Estadual de Ponta
Grossa.

Orientador: Prof. Dr. Jorge Iulek

PONTA GROSSA
2013

Ficha Catalográfica Elaborada pelo Setor Tratamento da Informação BICEN/UEPG

M149

Machado, Agnes Thiane Pereira

Determinação e refinamento de estruturas tridimensionais da proteína surE de *Xylella fastidiosa* / Agnes Thiane Pereira Machado. Ponta Grossa, 2013.
87f.

Dissertação (Mestrado em Química Aplicada) – Universidade Estadual de Ponta Grossa .

Orientador: Prof. Dr. Jorge Iulek

1. SurE. 2. *Xylella fastidiosa*. 3. Estrutura tridimensional de proteínas. I. Iulek, Jorge. II. Universidade Estadual de Ponta Grossa. Mestrado em Química Aplicada. III.T

CDD: 572.67

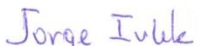
TERMO DE APROVAÇÃO

AGNES THIANE PEREIRA MACHADO

**"[DETERMINAÇÃO E REFINAMENTO DE ESTRUTURAS
TRIDIMENSIONAIS DA PROTEÍNA SurE DE *Xylella fastidiosa*]."**

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre no Programa de Pós-Graduação em Química Aplicada da Universidade Estadual de Ponta Grossa, pela seguinte banca examinadora.

Orientador:


Prof. Dr. Jorge Iulek
UEPG/PR


Profa. Dra. Mariza Boscacci Marques
UEPG/PR


Prof. Dr. Ricardo Aparício
UNICAMP/SP

Ponta Grossa, 12 de março de 2013.

AGRADECIMENTOS

A Deus pela vida, por todas as oportunidades e por sempre estar ao meu lado.

Ao Tiago, o amor da minha vida, e ao meu filho Gustavo, um pedacinho de mim, vocês são muito importantes na minha vida.

À minha mãe e ao Adevir por acreditarem que seria possível. Aos meus irmãos e aos meus avós por orarem por mim.

À Sr^a. Bete e ao Sr. Tito pelo apoio e por me ajudarem a cuidar do Gustavo.

Às minhas amigas Cristiane, Renata, Carla, que de alguma maneira contribuíram para esse projeto e aos amigos de laboratório por nossas discussões e ajuda recíproca.

Ao orientador Prof. Dr. Jorge Iulek pela ajuda a partir de seus conhecimentos. Aos Prof. Dr. Igor Polikarpov (IFSC-USP) e Prof. Dr. Ricardo Aparício (UNICAMP) por terem-nos disponibilizado as imagens de difração. Ao grupo de pesquisa do Prof. Dr. Ji-Rui Wang (Sichuan Agricultural University) por nos enviar em colaboração os plasmídeos com os genes codificantes dos inibidores.

Ao Prof. Dr. Marcio Silva (UTFPR - Ponta Grossa) por ajudar na transformação bacteriana e expressão dos inibidores. À Prof^a. Dr^a. Viviane Paula Martini (UNICENTRO) por ajudar significativamente nos experimentos e pela sua compreensão, amizade e paciência.

À CAPES, pelo auxílio financeiro, e ao Programa de Pós-Graduação em Química Aplicada da UEPG.

A todos que direta/indiretamente contribuíram para que este trabalho se realizasse.

“Fácil é sonhar todas as noites. Difícil é lutar por um sonho.”
Carlos Drummond de Andrade

RESUMO

No presente trabalho realizaram-se o processamento de sete conjuntos de imagens e a determinação e refinamento de duas estruturas da proteína SurE (stationary phases survival protein E) de *Xylella fastidiosa* (XfSurE). Os cristais foram obtidos pelo grupo de pesquisa do Prof. Dr. Ricardo Aparicio a partir da proteína produzida pelo grupo da Prof^a. Dr^a. Anete Pereira de Souza, ambos da Universidade Estadual de Campinas. XfSurE linhagem 9a5c é uma bactéria praga que causa uma enfermidade chamada Clorose Variegada dos Citros (CVC), grave em laranjeiras; foi à primeira bactéria que teve seu genoma completamente sequenciado no Brasil. Estudos revelam que essa proteína apresenta várias funções, dentre elas, de nucleotidase, capaz de desfosforilar vários ribo e desoxirribonucleicos 5-monofosfatados e ribonucleosídeos 3-monofosfatos em fosfato inorgânico e nucleosídeos. Dos sete conjuntos, quatro apresentaram anéis de difração por gelo no padrão de difração. Todas as estruturas apresentam grupo de espaço C2, mas há quatro celas unitárias diferentes. A substituição molecular foi realizada com o programa Phaser e apenas dois conjuntos de imagens tiveram suas estruturas tridimensionais refinadas em um processo iterativo com os programas Coot e Phenix, que foram posteriormente amplamente validadas. Uma apresenta um tetrâmero e a outra um dímero na unidade assimétrica. Durante o refinamento, modelaram-se íons Mn^{2+} , I^- e PO_4^{3-} . Comparações entre a XfSurE e CbSurE (a correspondente de *Coxiella burnetii*) indicam que XfSurE está em conformação aberta, ao contrário de CbSurE. A proteína apresenta o enovelamento característico de Rossmann-fold-like, com as alças de tetramerização e com projeção de C-terminal no contra-monômero.

Palavras-chaves: SurE, *Xylella fastidiosa*, Estrutura tridimensional de proteínas.

ABSTRACT

This dissertation presents the processing of seven image datasets and the determination and refinement of two protein structures of SurE (stationary phase survival protein E) from *Xylella fastidiosa* (XfSurE). The crystals were obtained by the research group of Prof. Dr. Ricardo Aparicio with the protein produced by the group of Prof^a. Dr^a. Anete Pereira de Souza, both from the State University of Campinas. *Xylella fastidiosa* strain 9a5c is a plague bacterium that causes a disease called Citrus Variegated Chlorosis (CVC) in orange grove; it was the first bacterium to have its genome completely sequenced in Brazil. Studies show that this protein has several functions, amongst them, of a nucleotidase able to dephosphorylate various ribo and deoxyribonucleic 5-monophosphates and ribonucleoside 3-monophosphate to inorganic phosphate and nucleosides. Seven different sets of diffraction images were processed, four of them presented rings from ice diffraction. All structures have space group C2, but there are four different unit cells. Molecular replacement was performed with Phaser program and only two sets of images had their three-dimensional structures refined in an iterative process with the programs Coot and Phenix, which were subsequently largely validated. One of them shows a tetramer and the other a dimer in the asymmetric unit. During refinement, Mn^{+2} , I^- and PO_4^{3-} ions were modeled. Comparisons between XfSurE and CbSurE (the correspondent protein from *Coxiella burnetii*) indicate that XfSurE is in open conformation, unlike CbSurE. The protein shows the characteristic Rossmann fold-like, with the tetramerization loop and C-terminal projection into the counter-monomer.

Key-words: SurE, *Xylella fastidiosa*, three-dimensional structure of proteins.

LISTA DE FIGURAS

Figura 1	- Microscopia eletrônica da <i>Xylella fastidiosa</i> nos vasos de xilema em laranjeiras.....	19
Figura 2	- Árvores contaminadas pela bactéria <i>Xylella fastidiosa</i> apresentam os sintomas da CVC em seus frutos e folhas.....	19
Figura 3	- Esquema de um arranjo Rossmann-fold-like em três camadas $\beta/\alpha/\beta$	21
Figura 4	- A) Modelo estrutural cristalográfico da TmSurE. B) Modelos estruturais para a proteína TmSurE adaptados por IWASAKI et al. 2007, onde aparecem as alças de tetramerização e o sítio ativo da família SurE.....	22
Figura 5	- Difração de raios X pela lei de Bragg.....	26
Figura 6	- A) Quando os raios X atingem diferentes planos, são refletidos em diferentes ângulos sendo detectado por um filme; B) Imagem de um padrão de difração	26
Figura 7	- Padrões de difração representativos dos conjuntos de imagens com anéis de gelo: A) XfG3AN2_NatMnaBA2; B) XfG3AN2_NatMnB_B5; C) XFG3A3ampD5_I; DI) imagens 1 -118 XFG3A3B_NATB4, DII) imagens 119 – 214 da XFG3A3B_NATB4.....	39
Figura 8	- Diagrama de fases para a água, mostrando o estado líquido, sólido dividido em (XI) gelo vítreo, (Ic) forma cúbica, (Ih) forma hexagonal.....	40
Figura 9	- Função de auto-rotação para κ igual a 180° . Contornos iniciam em altura relativa 20 e estratificam-se a cada 10 unidades, até o máximo de 100.....	45
Figura 10	- Alinhamento de sequência da proteína SurE de <i>Xylella fastidiosa</i> contra a de SurE de <i>Coxiella burnetti</i> (3TY2).....	47
Figura 11	- Gráfico de Ramachandran para as estruturas; A) XFG3A3oriC4cen e B) XFG3A3ampD5_I.....	50

Figura 12	- Qualidade do ajuste na densidade eletrônica representado em gráfico; AI, BI, CI e DI) estrutura XFG3A3oriC4cen; AII e BII) estrutura XFG3A3ampD5_I.....	53
Figura 13	- A) Diagrama da topologia da XfSurE; B) ilustração da estrutura para um monômero mostrando o enovelamento do tipo Rossmann-fold-like $\alpha/\beta/\alpha$	56
Figura 14	- Estrutura Tridimensional da XfSurE; A) XFG3A3oriC4cen; B) XFG3A3ampD5_I.....	58
Figura 15	- Regiões relativas de contato entre os monômeros no tetrâmero.....	59
Figura 16	- Sítio ativo em apenas um monômero da estrutura da XfSurE.....	61
Figura 17	- A), B), C) e D) Interação dos íons para cada monômero (A, B, C e D, respectivamente) da estrutura XFG3A3oriC4cen; e) e f) Interação dos íons respectivamente para cada monômero (A e B, respectivamente) da estrutura XFG3A3ampD5_I.....	62
Figura 18	- A) Sítios de interações para I- antes e após o estudo de ocupações para a estrutura XFG3A3oriC4cen. B) Sítios de interação para I- com ocupação de 100 % para a estrutura XFG3A3ampD5_I.....	64
Figura 19	-A) Superposição entre a estrutura XFG3A3oriC4cen (em vermelho) e XFG3A3ampD5_I (em azul); B) superposição de apenas um monômero mostrando cadeia lateral de alguns resíduos listados com algum ângulo chi diferente de 120°.....	68
Figura 20	- Superposição de um monômero da XfSurE (em verde) com a CbSurE (em azul), evidenciando as diferenças na cadeia principal, um destaque maior nas alças de tetramerização que estão deslocadas uma em relação à outra.....	71
Figura 21	- Superposição entre os sítios ativos, da XfSurE representado em verde e da CbSurE representado em azul.....	72
Figura 22	- Plasmídeo pPet30+ com os genes dos inibidores de α -amilase R9 e Ag2.....	84

Figura 23	- Placas com colônias de bactérias transformadas: a) colônias de bactérias com o gene da proteína R9; b) colônias de bactérias com o gene da proteína Ag2 (onde a posição do controle está trocada com da diluição 1:300).....	86
Figura 24	- Eletroforegrama SDS-PAGE para teste de expressão dos inibidores R9 e Ag2.....	87

LISTA DE TABELAS

Tabela 1	- Dados de difração de raios X para gelo hexagonal e cúbico.....	27
Tabela 2	- Resumo dos parâmetros das coletas de dados de difração referente aos sete cristais obtidos.....	38
Tabela 3	- Espessuras dos anéis de difração por gelo considerado como melhores para exclusão de cada conjunto de imagens separadamente.....	41
Tabela 4	- Resumo dos dados de processamento para as sete estruturas.....	43
Tabela 5	- Soluções de rotação encontradas pelo programa de substituição molecular Phaser); A) para a estrutura XFG3A3oriC4cen; B) para a estrutura XFG3A3ampD5_I.....	46
Tabela 6	- Dados estatísticos do refinamento inicial de corpo rígido, com restrições e simulação "annealing", para as duas estruturas refinadas.....	48
Tabela 7	- Valores estatísticos de refinamento para ambas as estruturas refinadas.....	49
Tabela 8	- A) Resumo da validação da estrutura XFG3A3oriC4cen; B) Resumo da validação da estrutura XFG3A3ampD5_I.....	51
Tabela 9	- Valores globais de RSR e RSCC por cadeia.....	55
Tabela 10	- Área de superfície e quantidades de resíduos em contatos intermonômeros de acordo com o programa PISA.....	59
Tabela 11	- Valores de ocupações e Bfactor médios dos resíduos e águas ao redor dos íons Mn^{2+} e PO_4^{3-} , para cada estrutura. A estrutura XFG3A3oriC4cen apresenta os valores das águas com ocupações refinadas alternadas com o PO_4^{3-}	63
Tabela 12	- B_{factor} do I's, Bfactor médios dos resíduos e das águas ao redor do I's.....	66
Tabela 13	- Comparação entre os picos dos mapas anômalos com as posições refinadas dos íons I- referentes as estrutura A) XFG3A3oriC4cen e B) XFG3A3ampD5_I.....	67

LISTA DE SIGLAS

3-AMP	- Adenosina Monofosfato
5-AMP	- 5-monofosfatados
BLAST	- <i>Basic Local Alignment Search Tool</i>
bp	- Pares de bases
CASTp	- <i>Computed Atlas of Surface Topography of proteins</i>
CbSurE	- SurE de <i>Coxiela burnetti</i>
Coot	- Programa utilizado para modelagem de proteínas (do inglês <i>CrystallographicObject-Oriented Toolkit</i>)
CVC	- Clorose Variegada dos Citros
d	- Distância
5-dAMP	- 2-Deoxiadenosina 5-fosfato
5-dGMP	- 3,5-GMP Cíclico Fosfodiesterases
DNA	- Ácido desoxirribonucleico
EBI	- European Bioinformatics Institute
<i>E.coli</i>	- <i>Escherichia coli</i>
F_{cal}	- Fator de estrutura calculado
F_{obs}	- Fator de estrutura observado
MW	- Massa molar
NCBI	- <i>Nacional Center for Biological Information (USA)</i>
NCS	- Simetria não cristalográfica
PDB	- Protein Data Bank
PISA	- Interfaces and Assemblies
PCM	- O-metiltransferase
Pet 30a+	- Vetor de expressão com o gene R9 e outro com o gene Ag2
rmsba	- Desvio médio quadrático de ângulos de ligação
rmsbl	- Desvio médio quadrático de comprimentos de ligação

rmssc	- Desvio médio quadrático de restrições de centros quirais
RNA	- Ácido ribonucléico
RSCC	- Valor do coeficiente de correlação do espaço real
RSR	- Valor residual no espaço real
R_{free}	- Índice residual livre, calculado com 5 a 10 % dos dados de reflexão num conjunto a parte que não é utilizado no refinamento.
R_{factor}	- Índice residual, é usado para medir o ajuste de uma estrutura de proteína aos dados de difração
SCOp	- Structural Classification of Proteins
SurE	- Stationary Phase Survival Protein E
TEMED	-Tetrametiletenodiana
TLS	- Translação- libração- parafuso (<i>Translation – libration-screw</i>)
TfSurE	- SurE de <i>Termusthermophilus</i>
TMSurE	- SurE de <i>Thermotoga marítima</i>
orf XF0703	- Codifica uma stationaryphasesurvivalprotein E (<i>SurE</i>)
XfSurE	- SurE de <i>Xylella fastidiosa</i>
XFG3A3oriC4cen	- Código para conjunto de dados de difração da XfSurE
XFG3A3B_NATB4	- Código para conjunto de dados de difração da XfSurE
XFG3A3ampD5_I	- Código para conjunto de dados de difração da XfSurE
XfG3AN2_NatMnaBA2	- Código para conjunto de dados de difração da XfSurE
XFAMPreso24	- Código para conjunto de dados de difração da XfSurE
XfG3AN2_NatMnB_B5	- Código para conjunto de dados de difração da XfSurE
XFAMPreso29	- Código para conjunto de dados de difração da XfSurE

SUMÁRIO

1 INTRODUÇÃO	17
1.1 JUSTIFICATIVA	17
1.2 ESTADO DA ARTE (REVISÃO BIBLIOGRÁFICA)	18
1.2.1 Importância do estudo tridimensional das proteínas	18
1.2.2 Xylella fastidiosa 9a5c	18
1.2.2.1 Protein SurE (stationary phase survival protein E)	19
1.2.2.1.1 Funções das SurEs	24
1.3 FUNDAMENTOS DAS TÉCNICAS	25
1.3.1 Difração de Raios X	25
1.3.1.1 Lei de Bragg	25
1.3.1.2 Padrão de difração	26
1.3.1.2.1 Formação de anéis de difração por gelo	27
1.3.1.3 Processamento de imagens: indexação, integração e escalonamento	28
1.3.1.4 Fator de estrutura e transformada de Fourier	29
1.3.1.5 Problema das fases	29
1.3.1.5.1 Substituição molecular	31
1.3.1.6 Refinamento	31
1.3.1.7 Validação	32
1.3.1.7.1 Validação na densidade eletrônica	32
1.3.2 Comparações estruturais	33
2 OBJETIVOS	34
2.1 Objetivos gerais	34
2.2 Objetivos específicos	34
3 PARTE COMPUTACIONAL	35
3.1 Programas computacionais:	35
4 MÉTODOS	36
4.1 Determinação da estrutura tridimensional da SurE de Xylella fastidiosa (XfSurE)	36
4.1.1 Indexação, integração e escalonamento das intensidades:	36

4.1.2 Substituição molecular	36
4.1.3 Refinamento	36
4.1.4 Validação.....	37
4.1.5 Comparações estruturais	37
5 RESULTADOS E DISCUSSÕES	38
5.1 Determinação da estrutura tridimensional da SurE de <i>Xylella fastidiosa</i> (XfSurE)	38
5.1.1 Dados de difração e processamento	38
5.1.1.1 Função de auto-rotação	45
5.1.2 Substituição molecular	45
5.1.3 Refinamento das estruturas cristalográficas.....	48
5.1.3 Validação das estruturas refinadas	48
5.1.5 Análises estruturais	55
5.1.5.1 Sítio ativo.....	59
5.1.5.2 Íons fora do sítio ativo	64
5.1.6 Comparações Estruturais	67
5.1.6.1 Comparações entre as estruturas refinadas.....	67
5.1.6.2 Comparação da estrutura XfSurE com a SurE de <i>Coxiella burnetti</i> (CbSurE)	71
6 CONCLUSÕES E PERSPECTIVAS FUTURAS	73
7 REFERÊNCIAS.....	74
8 APÊNDICE.....	81
8.1 Teste de expressão dos inibidores de α -amilase presente no gênero <i>Triticum</i> ...	81
8.1.1 Inibidores de α-amilase do gênero <i>Triticum</i>	81
8.1.1.1 Importância dos inibidores proteicos de α -amilase.....	81
8.1.2 Materiais e reagentes:.....	83
8.1.3 Parte experimental	84

8.1.3.1 Introdução dos vetores de expressão em E.coli por eletroporação	84
8.1.3.2 Teste de Expressão dos inibidores de α -amilase R9 e Ag2	85
8.1.4 Resultados e discussões.....	85
8.1.4.1 Introdução dos vetores de expressão em E.coli por eletroporação	85
8.1.4.2 Teste de Expressão dos inibidores de α -amilase R9 e Ag2	86

1 INTRODUÇÃO

1.1 JUSTIFICATIVA

As proteínas são as macromoléculas mais versáteis nos sistemas vivos e servem para funções cruciais essencialmente em todos os processos biológicos; conhecer a estrutura tridimensional destas permite investigar as relações entre sua sequência, sua estrutura e sua função, com uma indiscutível importância específica para o desenho racional de drogas terapêuticas. A cristalografia de proteínas foi o primeiro método desenvolvido para determinar a estrutura das proteínas com detalhes atômicos. Essa técnica pode revelar as posições tridimensionais precisas da maioria dos átomos em uma molécula de proteína (WHITFORD, 2005; ALBERTS *et al.*, 2006; STRYER *et al.*, 2004).

No presente trabalho realizaram-se o processamento de sete conjuntos de imagens e a determinação e refinamento tridimensional de duas estruturas da proteína Sure de *Xylella fastidiosa* (XfSurE). Os cristais foram obtidos pelo grupo de pesquisa do Prof. Dr. Ricardo Aparicio a partir da proteína produzida pelo grupo da Prof^a. Dr^a. Anete Pereira de Souza, ambos da Universidade Estadual de Campinas. Esta proteína apresenta múltiplas funções, entre elas, a de maior evidência é a de nucleotidase. *Xylella fastidiosa* é uma bactéria praga que causa uma enfermidade economicamente importante em laranjeiras.

1.2 ESTADO DA ARTE (REVISÃO BIBLIOGRÁFICA)

1.2.1 Importância do estudo tridimensional das proteínas

As proteínas são, indubitavelmente, os compostos biológicos mais importantes. Do ponto de vista químico, elas são as moléculas estruturalmente mais complexas e funcionalmente mais sofisticadas que conhecemos (ALBERTS *et al.*, 2006; STRYER *et al.*, 2004). Praticamente quase todas as proteínas são formadas por 20 aminoácidos, cada um com propriedades químicas distintas, ligados por ligação peptídica.

Existem vários tipos de proteínas que realizam diversas funções: estrutural, catálise, movimento, transporte, hormônios, proteção, armazenamento, regulação, entre muitas outras.

O estudo tridimensional nos permite investigar as relações entre sequência com sua estrutura e função. Por exemplo, dado que elas sustentam todos os aspectos da atividade biológica, seu estudo exerce um grande impacto na medicina. Avanços na genética molecular revelam muitas doenças que decorrem de defeitos específicos nas proteínas uma das maneiras de se entender detidamente isso é pelo estudo tridimensional da estrutura proteica (WHITFORD, 2005). Pode-se acrescentar, ainda, que é desta forma que os pesquisadores identificam o sítio ativo e descobrem mais rapidamente novas drogas que podem ser usadas para tratamento de várias doenças.

1.2.2 Xylella fastidiosa 9a5c

A *Xylella fastidiosa* linhagem 9a5c é uma bactéria gram-negativa (Figura 1) que possui formato de bacilo com tamanho aproximado de 0,25µm x 0.9 µm, o qual exhibe rugosidades na superfície da parede bacteriana, e não apresenta flagelos embora tenham fímbrias filamentosas. Causa uma enfermidade chamada Clorose Variegada dos Citros (CVC) conhecido popularmente como “amarelinho”, grave em laranjeiras. A CVC foi registrada pela primeira vez no Brasil em 1987. Um dos sintomas dessa enfermidade é a produção de uma substância viscosa, que é uma das possíveis causas do entupimento de vasos de xilema, reduzindo o suprimento de

água e nutrientes para as plantas, deixando os frutos afetados pequenos e endurecidos sem nenhum valor comercial (Figura 2) (POOLER *et al.*, 1995; SARAIVA *et al.*, 2009).

Figura 1: Microscopia eletrônica da *Xylella fastidiosa* nos vasos de xilema em laranjeiras. Fonte: <<http://aeg.lbi.ic.unicamp.br/xf/home/mmachado>>. Acesso em 23/11/2012.

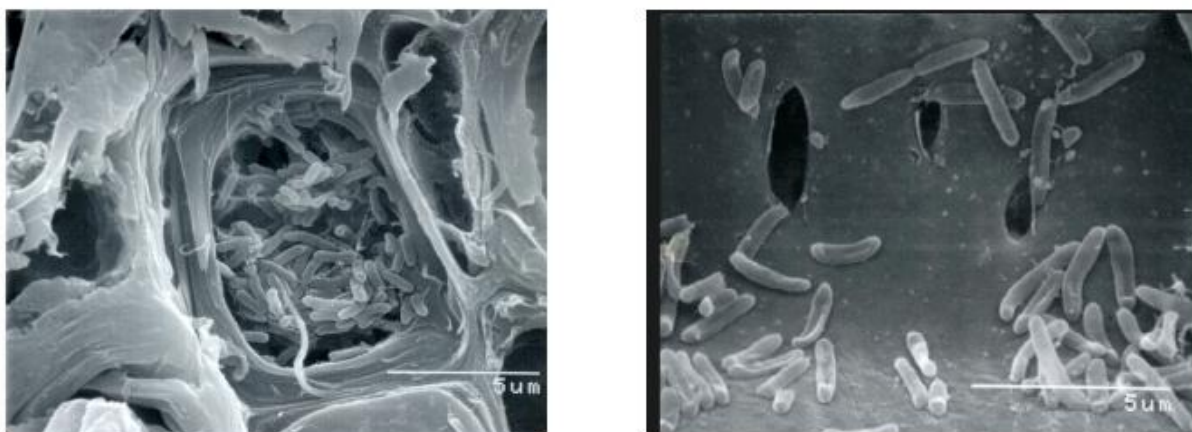


Figura 2: Árvores contaminadas pela bactéria *Xylella fastidiosa* apresentam os sintomas da CVC em seus frutos e folhas. <www.cnpmf.embrapa.br/publicacoes/produto_em_foco/citros_34.pdf>. Acesso em 14/01/2013.



A *Xylella fastidiosa* foi o primeiro organismo com genoma completamente sequenciado no Brasil (MEIDANIS *et al.*, 2002). Sua transmissão é feita por mudas contaminadas ou por vetores, como as dessezete espécies de “cigarrinhas” que se alimentam da seiva do xilema (SIMPSON *et al.*, 2000). Atualmente essa bactéria se

encontra dissiminada em todas as regiões decitricultura do país, principalmente no estado de São Paulo, um dos mais importantes citricultores do mundo (LAMBIA Set *al.*, 2000). O controle da CVC é atualmente limitado à remoção de ramos infectados pela poda, à aplicação de inseticidas e à utilização de plantas saudáveis para novos pomares. Além da CVC, outras estirpes de *Xylella fastidiosa* causam uma gama de enfermidades para diversas outras plantas, entre elas, videira, alfafa, pêsegueiro, ameixeiras, pereiras, amendoeiras, plantas de café. Ao contrário de outros fitopatógenos, a *Xylella* é xilema-dependente, vivendo exclusivamente dos vasos de xilema das árvores (SIMPSON *et al.*, 2000).

Nessa bactéria existem 220 genes cujas funções têm sido previstas por meio de análises de comparações da sequência, embora eles não estejam classificados pela via metabólica (WWW. Xylella. Incc. br). (SIMPSON *et al.*, 2000; SARAIVA *et al.*, 2009). O seu genoma compreende um cromossomo circular de 2,7 Mb e dois plasmídeos, um com 51kb e outro com 1kb.

Muitos dos genes dessa bactéria foram sequenciados na Universidade Estadual de Campinas (UNICAMP), disponíveis no site www.lbi.ic.unicamp.br/xf. Para a identificação das famílias de transporte, cada sequência de gene da *Xylella* foi comparada com um banco de dados consistindo de todos os exemplos de proteínas de transporte, como relatado no site (WWW-biology.ucsd.edu/Msaier/transporte), constantemente atualizado e no qual cada categoria de proteína de transporte tem uma descrição detalhada, incluindo o modo de transporte, a topologia da proteína, a especificidade de substrato e as espécies, seguidas por uma revisão na literatura (MEIDANIS *et al.*, 2002).

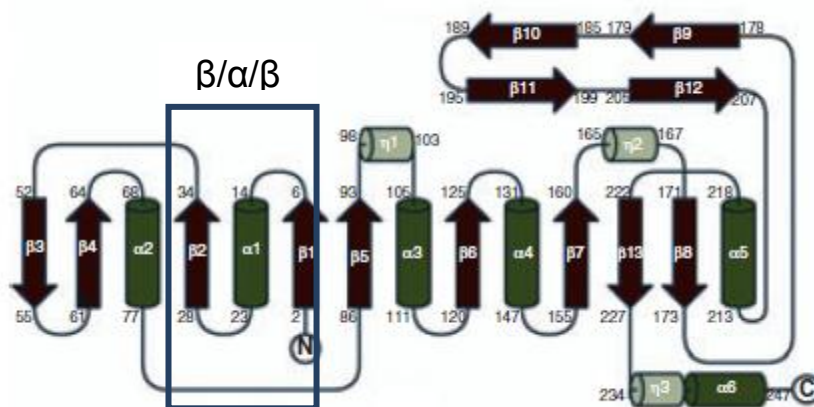
1.2.2.1 Protein SurE (stationary phase survival protein E)

Entre os genes da bactéria *Xylella fastidiosa*, a *orf* XF0703 codifica uma stationary phase survival protein E (SurE). A SurE é uma proteína codificada por genoma amplamente propagado entre as eubactérias e eucariontes e aparentemente é bem conservada. Esse gene apresenta 789 bp e codifica a proteína XfSurE que apresenta 262 aminoácidos com massa molecular de 28,3 kDa e ponto isoelétrico 5,23. Tem uma forte similaridade com outras SurE anteriormente estudadas (SARAIVA *et al.*, 2009).

O gene SurE foi primeiramente descoberto em *E.coli* e está localizado “upstream” ao gene pcm, o qual codifica a proteína L-isoaspartil e O-metiltransferase (PCM) (LI *et al.* 1994). Os genes SurE e PCM formam um operon bicistrônico em *E.coli* e em *Xylella*. Em células velhas, ocorrem danos espontâneos nas cadeias laterais de resíduos, a proteína PCM atua na via de sinalização da conversão de resíduos de L-isoaspartil em resíduos normais de L-aspartil. Estes genes são altamente expressos durante a fase estacionária em *E.coli*.

A estrutura cristalográfica de seis proteínas homólogas foi relatada: *Thermotoga marítima* (TmSurE) (1J9K) (LEE *et al.*, 2001), *Pyrobaculum aerophilum* (1L5X) (MURA *et al.*, 2003), *Thermophilus* (TtSurE) (2E6E) (IWASAKI *et al.*, 2007), *Aquifexaericus* (2PHJ) (ANTONYUK *et al.*, 2009) e *Salmonella typhimurium* (2V4N) (PAPPACHAN *et al.*, 2008) e *Coxiella burnetii* (CbSurE) (apenas estrutura depositada no PDB 3TY2). O estado oligomérico de proteínas SurE em solução sempre foi controverso, estudos revelam que esse estado depende do micro-organismo estudado, sendo que a SurE necessita estar em um arranjo tetramérico para se tornar ativa. Dímeros são as formas mais frequentemente relatadas na literatura. Observou-se que essa proteína existe como dímero e/ou como tetrâmero. A estrutura terciária das SurE essencialmente consiste em duas partes: um domínio globular arranjado como *Rossmann-fold-like* (Figura 3) em três camadas $\alpha/\beta/\alpha$ ou $\beta/\alpha/\beta$ e β -harpin, que representa as alças de tetramerização (PAPPACHAN *et al.*, 2008).

Figura 3: Esquema de um arranjo *Rossmann-fold-like* em três camadas $\beta/\alpha/\beta$, descrito por (PAPPACHAN *et al.*, 2008).



Estudos estruturais de TtSurE mostram que o sítio ativo compreende uma região de loop dos resíduos 34 ao 50 que, em conformação aberta, ocorre à ligação do substrato (AMP), este ausente em conformação fechada. Em cada monômero há um sítio de interação para o Mn^{2+} coordenado pelo glutamato 37, pelos aspartatos 8 e 9 e também pelo o íon PO_4^{3-} . Acredita-se que o glutamato seja fundamental para a ligação do cofator. Em experimentos realizados por mutação sítio-dirigida o mutante E37A não se liga ao cofator deixando a proteína sem atividade. O mecanismo de catálise ocorre por um ataque nucleofílico da água à ligação do fosfato ao substrato (IWASAKI *et al.*, 2007).

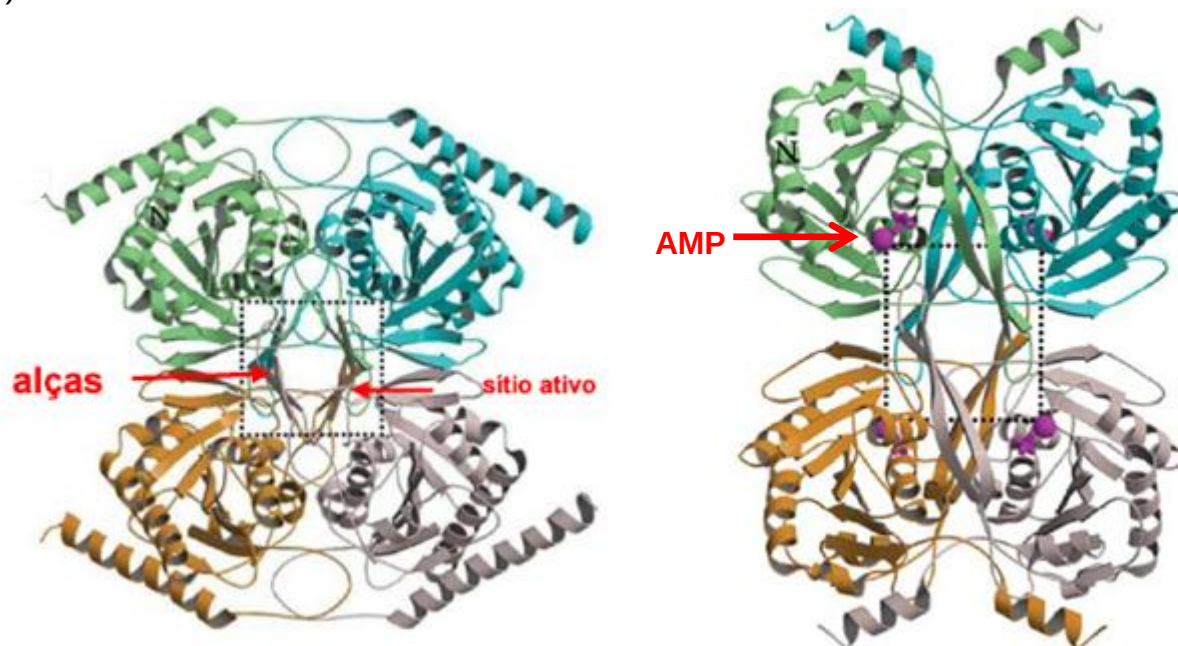
O modelo de tetrâmero da TmSurE (LEE *et al.*, 2001) (Figura 4) adaptado por IWASAKI *et al.*, 2007 mostra o sítio ativo composto pelas alças de tetramerização e os loops em conformação fechada, sem ligação do AMP; isso reforça a ideia que a mudança entre as conformações fechada e aberta é crítica para sua atividade (IWASAKI *et al.*, 2007).

Figura 4: A) Modelo estrutural cristalográfico da TmSurE. B) Modelos estruturais para a proteína TmSurE adaptados por IWASAKI *et al.* 2007, onde aparecem as alças de tetramerização e o sítio ativo da família SurE. Compara-se a estrutura em conformação aberta (cristalográfica), com interação com AMP (modelado), e fechada (modelada), sem interação com AMP. (LEE *et al.*, 2001; IWASAKI *et al.*, 2007).

A)



B)



Conformação sem interação do substrato Conformação com interação do substrato

Recentemente, o gene que codifica a proteína SurE de *Xylella fastidiosa* foi inserido num plasmídeo pET29a. Esse plasmídeo com o inserto foi clonado em *E. coli* BL21 e os ensaios de superexpressão apontaram bom resultado após 20 h de pós-indução usando lactose 5,6 mM. Aproximadamente 20 mg de proteína estavam na fração solúvel, à purificação foi realizada por cromatografia de afinidade usando níquel e SDS-Page identificou uma única banda com aproximadamente 30 kDa. (SARAIVA *et al.*, 2009).

Estudos realizados por SAXS revelarem que o estado oligomérico da proteína não depende da concentração da mesma, sendo que em solução ela assume forma tetramérica. (SARAIVA *et al.*, 2009).

Da proteína assim purificada, foram obtidos inicialmente dois cristais, denominados de isoforma I e isoforma II, ambas cristalizadas em presença de MnCl_2 e a isoforma II também na de 3-AMP (REIS *et al.*, 2012). O padrão de difração da isoforma II apresenta anéis de difração por gelo.

1.2.2.1.1 Funções das SurEs

Estudos revelam que essa proteína é uma nucleotidase ou nucleosídeofosfohidrolase (EC 3.1.3.5) capaz de desfosforilar vários ribo e desoxirribonucleicos 5'-monofosfatados e ribonucleosídios 3'-monofosfato, em fosfato inorgânico e nucleosídeos (PROUDFOOT *et al.*, 2004). Tal função contribui para o balanço correto dos *pools* de nucleotídeos da célula.

A manutenção correto dos “*pools*” de nucleotídeos é importante para que ocorra uma síntese precisa do DNA, seu desequilíbrio pode acarretar a redução da replicação do DNA e também o surgimento de mutações cromossômicas (KUNDZ *et al.*, 1994)

Apesar dos estudos, a função da proteína SurE não é completamente esclarecida, nem o seu mecanismo catalítico. As SurE são classificadas como nucleotidasas por causa ampla especificidade por nucleosídios, mas também podem estar envolvidas em uma via de regulação de síntese do DNA e RNA. Outra hipótese é que sejam enzimas “*house-clean*”, estando envolvidas no processo de sinalização da célula contra compostos perigosos, pois interceptam e hidrolisam nucleotídeos não canônicos, prevenindo o dano no DNA e suplementando o trabalho do sistema de reparo de DNA. (GONÇALVES *et al.* 2008; GALPERIN *et al.* 2006).

Essas proteínas apresentam uma maior afinidade em pH neutro e em presença de Mn^{2+} , seguido por Co^{2+} , Mg^{2+} e Ni^{2+} . Apresentam uma maior atividade em presença de 3-AMP, seguido por 5'-dGMP, 5'-dAMP e 5-AMP (PROUDFOOT *et al.*, 2004; SARAIVA *et al.*, 2009).

A especificidade por 3' - monofosfatados indica que essa proteína pode estar envolvida na regulação de 3' mononucleotídeos originados da degradação de RNAs, portanto, trata-se de uma enzima multifuncional com atividade de 5'-nucleotidase, 3'-nucleotidase e exopolifosfatase (PROUDFOOT *et al.*, 2004).

1.3 FUNDAMENTOS DAS TÉCNICAS

1.3.1 Difração de Raios X

1.3.1.1 Lei de Bragg

Estudos por difração de raios X têm permitido entender melhor as propriedades das proteínas, por exemplo, estudar a nível atômico os centros catalíticos de enzimas e compreender a especificidade das reações que essas catalisam (ROMÃO, 1996).

O fenômeno de difração de raios X é explicado pela Lei de Bragg, a qual demonstra que os ângulos com que emergem os feixes difratados (Figura 5) poderiam ser calculados tratando-se o fenômeno da difração como reflexões por vários conjuntos de planos paralelos e equivalentes, definidos pelos átomos de um cristal. Assim, cada ponto no padrão de difração se designa por reflexão. Bragg demonstrou que a difração ocorre quando a diferença da distância percorrida pelos raios é igual a um múltiplo do comprimento de onda dos feixes difratados; a diferença de distância depende do ângulo de reflexão e do espaçamento entre os planos, dado pela lei de Bragg (Equação 1) (CULLITY, 1978; BLOW, 2005):

$$2 d \sin \theta = n \lambda \quad \text{Equação 1}$$

Onde:

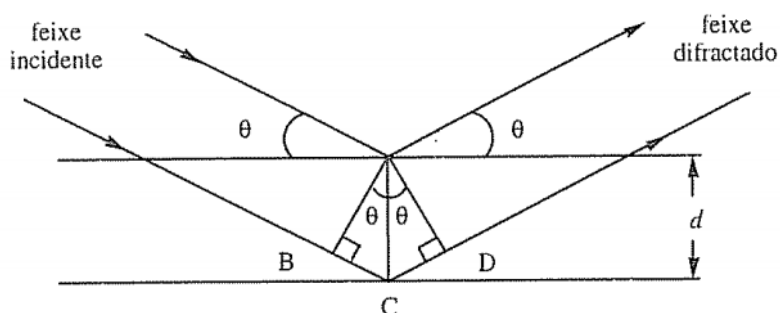
d = distância entre os planos

θ = ângulo de difração

λ = comprimento de onda.

n = número inteiro

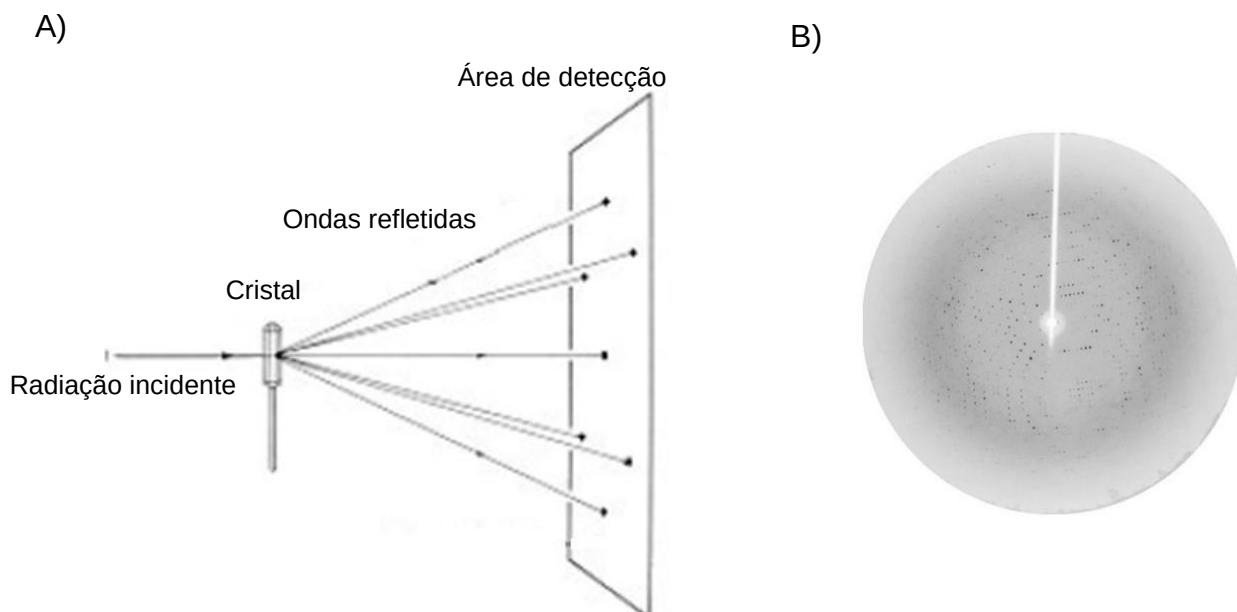
Figura 5: Difração de raios X pela lei de Bragg. (ROMÃO, 1996; CULLITY, 1978)



1.3.1.2 Padrão de difração

Um padrão de difração consiste em reflexões arranjadas de maneira ordenada. Cada ponto é o resultado do espalhamento das ondas refletidas para um grupo de planos reticulares paralelos. Durante o giro do cristal em um feixe de raios X incidente, a orientação de cada um dos planos reticulares move-se e as ondas que são refletidas em fase mudam consequentemente, mudando as posições das reflexões em sucessivas imagens (Figura 6)(BLOW, 2005).

Figura 6:A) Quando os raios X atingem diferentes planos, são refletidos em diferentes ângulos sendo detectado por um filme; B) Imagem de um padrão de difração. Disponível em <www.azevedolab.dominiotemporario.com/doc/diffraction1.pdf>. Acesso em 17 fev. 2012



No padrão de difração existe uma relação inversa simples entre o espaçamento das celas elementares de um cristal chamado espaço real e a distância entre as reflexões no reticulado registrado no filme chamado espaço recíproco. Assim, podem-se calcular as dimensões da cela unitária a partir do espaçamento observado no espaço recíproco, na imagem obtida por difração de raios X (BLOW, 2005). Cada ponto (reflexão) pode ser identificado por seus índices de Miller (h, k, l), a distância entre os planos e os respectivos ângulos de incidência (θ).

1.3.1.2.1 Formação de anéis de difração por gelo

Experimentos de difração de raios X de cristalografia de proteínas ocorrem amiúde a baixas temperaturas, técnicas de crioproteção são usadas rotineiramente em muitos laboratórios. Todavia, eventualmente anéis correspondendo à difração por gelo finamente dividido são observados em padrões de difrações; muitos estudos têm sido realizados para caracterizar a cristalinidade deste gelo. Geralmente existem três possibilidades de fases de água a baixa temperatura e pressão: fase hexagonal ou gelo I, fase cúbica e fase vítreo. A presença dos anéis de gelo prejudica a estimativa das intensidades durante o processamento dos conjuntos de imagens, para isso, é importante a exclusão desses anéis, o que por outro lado leva à perda de reflexões da proteína que ficaram próximas ou até mesmo superpostas aos anéis. Com perda dessas reflexões deterioramos o conjunto de dados, assim quando possível devemos evitar esses anéis ou então, na impossibilidade de evitá-los, no processamento minimizar os efeitos destes na estimativa do conjunto de intensidades (ELSPETH *et al.*, 1997).

A fase cúbica raramente aparece nos padrões de difração já que há fácil transição para a fase hexagonal (DOWELL *et al.*, 1960). A Tabela 1 mostra os planos, as distâncias entre estes e as intensidades relativas para as fases hexagonal e cúbica (DOWELL *et al.*, 1960; KOHL *et al.*, 2000; KIM C.U *et al.*, 2008; CHAPMAN M. S., 2010). Esses dados podem guiar os cristalógrafos na determinação da espessura de exclusão dos anéis de difração por gelo nas imagens antes do processamento (DOWELL *et al.*, 1960).

Tabela 1: Dados de difração de raios X para gelo hexagonal e cúbico (DOWELL *et al.*, 1960).

Hexagonal			Cúbico		
hkl	d (Å)	I _{rel}	hkl	d (Å)	I _{rel}
100	3,897	100			
002	3,669	75	111	3,670	100
101	3,441	53			
102	2,671	17			
110	2,249	39	220	2,248	40
103	2,072	30			
200	1,948	4			
112	1,918	18	311	1,917	15
201	1,883	3			
202	1,721	2			

1.3.1.3 Processamento de imagens: indexação, integração e escalonamento.

O próximo passo, após a coleta de dados, é a conversão da informação contida nas imagens em fatores de estrutura. Para isto, há diversos programas, como XDS (KABSCH, 1994) e MOSFLM (LESLIE, 2006). O programa utilizado neste trabalho foi o XDS, que requer o uso de imagens coletadas a partir da geometria de oscilação-rotação e radiação monocromática.

Na etapa de indexação, determinam-se os possíveis grupos de espaço, as dimensões e a orientação da célula unitária. Cada imagem em um mesmo conjunto de dados pode ter um valor de escala diferente devido à variação na radiação incidente, absorção dos raios difratados ou destruição do cristal, portanto, esses valores são necessários para colocar as intensidades em uma base comum (DRENTH, 1994). Também, um valor de variação de acordo com a resolução é calculado para cada imagem. Assim, na etapa do escalonamento, são determinados esses fatores de escala entre imagens/reflexões, tal que as intensidades relacionadas por simetria possam ser promediadas, de modo a se obter um conjunto final de reflexões únicas. O escalonamento é realizado por meio do programa SCALA ou XSCALE.

Os pontos nas imagens de difração apresentam intensidade proporcional ao quadrado do módulo do fator de estrutura (BLOW, 2005). As intensidades podem geralmente ser medidas até certo limite de resolução, que é dado pela área de detecção, à distância e orientação desta em relação ao cristal, os parâmetros de cela e a qualidade do cristal. É desejável produzir mapas de densidade eletrônica com alta qualidade, consequência da obtenção de fatores de estrutura a alta resolução.

A completeza do conjunto de imagens de difração é definida pelo número de reflexões únicas mensuradas, divididas pelo número total de reflexões únicas.

A multiplicidade ou redundância das intensidades medidas é definida como o número total de intensidades medidas dividido pelo número total de reflexões únicas medidas, ou seja, informa o número médio de vezes em que o ponto de difração ou seus equivalentes são medidos (ROMÃO, 1996; BLOW, 2005).

O fator R_{sym} determina a concordância entre os dados simetricamente equivalentes medidos. Erros nas medidas de intensidades isso aumentarão o R_{sym} . Esse fator varia com a resolução e é costumeiramente calculado para um pequeno número de faixas de resolução. Ele é calculado pela seguinte equação (Equação 2):

$$R_{sym} = \sum hkl \frac{\sum j |I_j(hkl) - [I(hkl)]|}{\sum j I_j} \quad \text{Equação 2:}$$

A relação sinal/ruído também é um fator importante no processamento das imagens. Ela mostra a qualidade dos dados medidos em relação ao desvio padrão estimado dentro de uma faixa de resolução.

O Wilson Plot permite estimar um fator de desordem global B da estrutura, em boa parte devido à vibração térmica, e colocar os fatores de estrutura em escala aproximadamente absoluta, o que significa que o valor numérico do fator de estrutura é aproximadamente igual ao valor de elétrons equivalentes naquela direção.

1.3.1.4 Fator de estrutura e transformada de Fourier

No padrão de difração, cada ponto na imagem gerada é produzido por um feixe difratado (que é a soma das ondas "refletidas" por um grupo de planos reticulares) que tem amplitude e fase. A expressão (Equação 3) abaixo é conhecida como uma transformada de Fourier porque é a soma de coeficientes de Fourier das ondas:

$$F_{hkl} = \sum_{j=1}^n f_j e^{2\pi i(hx_j + ky_j + lz_j)} \quad \text{Equação 3}$$

onde f_j é o fator de espalhamento atômico e o somatório é feito com todos os átomos na cela unitária (BLOW, 2005). O fator de espalhamento traz o quanto de radiação um átomo espalhará numa determinada direção. Quando as fases e as amplitudes de todos os fatores de estrutura são conhecidas, é possível calcular a transformada de Fourier inversa, que é a densidade eletrônica do cristal indicada na expressão (Equação 4) abaixo (BLOW, 2005; DRENTH, 1994):

$$\rho(x, y, z) = \frac{1}{V} \sum_{hkl} F(hkl) \exp [-2\pi i(hx + ky + lz)] \quad \text{Equação 4}$$

1.3.1.5 Problema das fases

Na coleta dos dados de difração obtemos apenas as amplitudes enquanto que as fases são perdidas, o que nos impede de calcular diretamente as densidades. Para se resolver esse problema existem métodos que permitem obter aproximadamente as fases, tais como substituição molecular, que será discutido adiante, substituição isomórfica e dispersão anômala, esses dois últimos métodos aplicados quando não se dispõe de um modelo estrutural (BLOW, 2005).

1.3.1.5.1 Substituição molecular

Antes da substituição molecular é necessária a realização do alinhamento da sequência de aminoácido da proteína de interesse com todas as sequências de aminoácidos das estruturas depositadas no PDB (Protein Data Bank); esse alinhamento é uma forma de organizar e comparar regiões de similaridade e identidade entre as proteínas e pode ser realizado pelo programa BLAST ("Basic Local Alignment Search Tool"), (PARK *et al.*, 2012). Ao identificar a estrutura com uma maior similaridade e identidade, então se realiza a preparação de um modelo estrutural inicial para a substituição molecular (FILHO *et al.*, 2003). Este modelo deve ser colocado dentro da cela unitária de modo que fases iniciais aproximadas possam ser calculadas para o padrão de difração da proteína desconhecida. A estimativa da orientação e posição do modelo é decomposta em dois componentes: a rotação tridimensional da molécula na orientação correta composta por ângulos e a translação tridimensional da mesma orientada corretamente dentro da unidade assimétrica composta por um conjunto de coordenadas fracionárias x, y, z. Assim, os fatores de estrutura são calculados para esta posição (F_{cal}) e comparados com os fatores de estrutura observados (F_{obs}), assim como fases iniciais são estimadas (BLOW, 2005).

1.3.1.6 Refinamento

O refinamento é um processo de ajuste do modelo que visa, entre vários objetivos, encontrar um melhor acordo entre o fator de estrutura calculado (F_{calc}) e o fator de estrutura observado (F_{obs}), variando-se as coordenadas atômicas e o fator B de temperatura (B_{factor}), entre outros parâmetros. Vários métodos têm sido desenvolvidos e aplicados, eles normalmente diminuem o valor dos índices residuais R_{factor} e R_{free} , definidos abaixo (Equação 5):

$$R_{factor} = \frac{\sum_h (|F_{obs}| - |F_{cal}|)}{\sum_h |F_{obs}|} \quad \text{Equação 5}$$

O R_{free} é calculado da mesma forma que o R_{factor} , só que para um conjunto de reflexões, normalmente de 5 a 10% do número total, excluídas do refinamento,

enquanto que o R_{factor} usa as demais; isto permite indicar sem tendenciosidade a probabilidade de que o modelo esteja mais correto ou não a cada ciclo de refinamento.

À medida que os átomos são movidos corretamente durante o refinamento, as fases calculadas do modelo estarão mais próximas das fases reais das ondas difratadas e a densidade eletrônica melhorará. Idealmente, num bom refinamento deve-se usar um número de parâmetros conhecidos várias vezes superiores ao número de variáveis a serem refinadas.

1.3.1.7 Validação

Na validação deve-se avaliar a qualidade de todo o modelo estimando os possíveis erros estruturais em regiões localizadas, levando-se em consideração os parâmetros estereoquímicos. Um programa bastante utilizado na avaliação dos parâmetros estereoquímicos é o PROCHECK, que avalia os comprimentos de ligações, os ângulos planos, a planaridade dos anéis de cadeias laterais, a quiralidade, as conformações de cadeias laterais, o impedimento estérico entre pares de átomos não ligados e a qualidade do gráfico de Ramachandran. Esse gráfico é particularmente útil porque define os resíduos que se encontram nas regiões energeticamente mais favoráveis e desfavoráveis e orienta a avaliação da qualidade de modelos teóricos ou experimentais de proteínas. É necessário avaliar também as interações entre as estruturas modeladas e o meio, neste tipo de análise, pode-se usar o programa WHATCHECK, que traz informações sobre a formação de regiões centrais hidrofóbicas (FILHO *et al.*, 2003).

1.3.1.7.1 Validação na densidade eletrônica

O valor do residual do espaço real (RSR) é uma medida da similaridade entre o mapa de densidade eletrônica calculada a partir de um modelo estrutural e um calculado a partir de dados experimentais. Esse coeficiente pode ser calculado pela Equação 6:

$$RSR = \frac{\sum \rho_{obs} - \rho_{cal}}{\sum \rho_{obs} + \rho_{cal}} \quad \text{Equação 6}$$

Um outro índice para indicar a qualidade do acerto do modelo na densidade eletrônica é o valor de coeficiente de correlação no espaço real (RSCC), calculado pela Equação 7. Ambos podem ser usados para indicar qualidade de ajuste de forma global ou localizada. A vantagem do RSCC é que a densidade calculada e a observada não precisam ser escalonadas conjuntamente, mas uma pequena desvantagem é que uma densidade baixa que tenha uma distribuição de intensidade adequada levará a um valor alto de RSCC.

$$RSCC = \frac{\sum |\rho_{obs} - (\rho_{obs})| \sum |\rho_{obs} - (\rho_{cal})|}{(\sum |\rho_{obs} - (\rho_{cal})|^2 \sum |\rho_{obs} - (\rho_{cal})|^2)^{1/2}} \quad \text{Equação 7}$$

1.3.2 Comparações estruturais

O processo de comparação de estruturas de proteínas permite indicar diferenças que justificam propriedades distintas entre estas. Pode-se comparar de modo mais global ou específico a certas regiões de maior interesse. A maior parte das proteínas reportadas com suas estruturas tridimensionais encontra-se depositada na base de dados pública PDB, que também armazena informações físico-químicas de cada estrutura, a forma com que foram determinados, os autores do trabalho e outros dados. Tal base de dados é importante porque permite que pesquisadores de diferentes centros de pesquisas espalhados pelo mundo utilizem informações estruturais e funcionais de proteínas conhecidas. Por meio destas informações, há maiores chances de identificar funções e características de novas proteínas com o uso de diversas técnicas de comparação, entre elas a superposição e o alinhamento estrutural.

2 OBJETIVOS

2.1 Objetivos gerais

- Determinar a estrutura tridimensional da proteína SurE de *Xylella fastidiosa*, a fim de elucidar sua interação com o cofator Mn^{2+} e realizar comparações estruturais.

2.2 Objetivos específicos

- Processar sete conjuntos de imagens da proteína SurE de *Xylella fastidiosa* com os programas XDS e XSCALE;
- Refinar duas estruturas, cada uma com cela unitária diferente;
- Determinar as diferenças entre as estruturas refinadas da SurE de *Xylella fastidiosa* e fazer outras comparações estruturais;

3 PARTE COMPUTACIONAL

3.1 *Programas computacionais:*

Os programas e servidores utilizados foram: a) para alinhamento das sequências, ClustalW2 (Multiple Sequence Alignment) e BLAST (PARK *et al.*, 2012); b) para representação gráfica, Aline (BOND *et al.*, 2009) e Pymol (DeLANO, 2002); c) para o trabalho cristalográfico, XDS (X-ray Detector Software) (KABSCH, 2009), Chainsaw (STEIN, 2008), Phaser (McCOY *et al.*, 2007), Coot (EMSLEY *et al.*, 2004) e Phenix.refine (AFONINE *et al.*, 2009); d) para validação das estruturas refinadas, Procheck (LASKOWSKI *et al.*, 1993), Validation PDB (WESTBOOK *et al.*, 2003), Mapmam (KLEYWEGT *ET al.*, 1996), Whatcheck (HOOFT *et al.*, 1996) e Molprobity (CHEN *et al.*, 2012); e) para análises e comparações estruturais, PISA (KRISSINEL *et al.*, 2007) em EBI (European Bioinformatics Institute www.ebi.ac.uk/), Superpose (KRISSINEL *et al.*, 2004), SCOP (Structural Classification of Proteins) (CONTE *et al.*, 2002), PDBsum (LASKOWSKI, 2009) e CASTp (DUNDAS *et al.*, 2006); f) para teste de TLS: TLMSD (PAINTER *et al.*, 2006).

4 MÉTODOS

4.1 Determinação da estrutura tridimensional da SurE de Xylella fastidiosa (XfSurE)

4.1.1 Indexação, integração e escalonamento das intensidades:

Foram processados sete conjuntos de imagens da proteína XfSurE gentilmente cedidos pelo Prof. Dr. Igor Polikarpov (IFSC-USP) e o Prof. Dr. Ricardo Aparicio (UNICAMP). O pacote utilizado para indexação, integração e escalonamento das imagens foi o XDS(KABSCH, 2009), com auxílio *scripts* elaborados pelo Prof. Dr. Jorge Iulek.

4.1.2 Substituição molecular

Realizou-se uma busca pelo programa BLAST (PARK *et al.*, 2012) do site NCBI para identificar quais estruturas do PDB apresentavam maiores similaridades e identidades com a proteína XfSurE. Para a preparação do modelo para a substituição molecular foi utilizado o programa Chainsaw (STEIN, 2008) e para encontrar a rotação e translação da molécula foi utilizado o programa Phaser (McCOY *et al.*, 2007).

4.1.3 Refinamento

Dos sete conjuntos de imagens, dois conjuntos selecionados e tiveram suas estruturas refinadas e modeladas num processo iterativo com os programas Phenix.refine (AFONINE *et al.*, 2009) e Coot (EMSLEY *et al.*, 2004). Os mapas utilizados para densidade eletrônica foram 2mFo-DFc no mapa Fourier diferença foi mFo-DFc.

4.1.4 Validação

Na validação dos modelos foram utilizados os programas Procheck (LASKOWSKI *et al.*, 1993), Whatcheck (HOOFT *et al.* 1996), Molprobity (CHEN *et al.*, 2012), ValidationPDB (WESTBOOK *et al.*, 2003) e Mapmam (KLEYWEGT ET *et al.*, 1996).

4.1.5 Comparações estruturais

Foram realizadas comparações entre as duas estruturas refinadas nesse trabalho e delas com a proteína SurEde *Coxiella burnetti* (PDB 3TY2), baseadas em similaridade estrutural e área e volume de sítio ativo.

5 RESULTADOS E DISCUSSÕES

5.1 Determinação da estrutura tridimensional da SurEde *Xylella fastidiosa* (XfSurE)

5.1.1 Dados de difração e processamento

A cristalização e a coleta dos dados de difração de dois conjuntos de imagens estão relatados em REIS *et al.*, (2012). No presente trabalho foram processados sete conjuntos de imagens, a Tabela 2 mostra um resumo de parâmetros das coletas em que se utiliza uma designação própria para cada conjunto de dados/cristal. Esses parâmetros são dados importantes para o início do processamento dos conjuntos de imagens. No Síncrotron utilizado, LNLS em Campinas-SP, há duas linhas para coleta de dados de difração por cristais de proteínas, MX1 e MX2, que diferem em alguns aspectos, do qual destacam-se a área de detecção e a intensidade da radiação incidente, maior para a última.

Tabela 2: Resumo dos parâmetros das coletas de dados de difração referente aos sete cristais obtidos.

Cristal	Comprimento de onda (Å)	Distância detector-cristal (mm)	Tempo de exposição por imagem(s)	Linha do Síncrotron (LNLS)
XFG3A3oriC4cen	1,437	180	60	MX2
XFG3A3B_NATB4	1,428	90	180	MX1
XFG3A3ampD5_I	1,437	180	60	MX2
XfG3AN2_NatMnaBA2	1,437	90	180	MX1
XFAMPreso24	1,428	120	180	MX2
XfG3AN2_NatMnB_B5	1,437	85	240	MX1
XFAMPreso29	1,428	150	90	MX2

Antes do processamento, observou-se que quatro conjuntos de imagens apresentavam anéis de difração por gelo, como mostra a Figura 7, em alguns casos

bem intensos, a formação desses anéis deveu-se à ineficiência de crioproteção, por causa da formação de muitos microcristais de gelo.

Figura 7: Padrões de difração representativos dos conjuntos de imagens com anéis de gelo: A) XfG3AN2_NatMnaBA2; B) XfG3AN2_NatMnB_B5; C) XFG3A3ampD5_I; DI) imagens 1 -118 XFG3A3B_NATB4, DII) imagens 119 – 214 da XFG3A3B_NATB4.

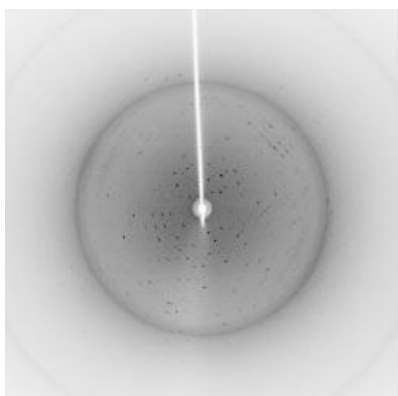
A)



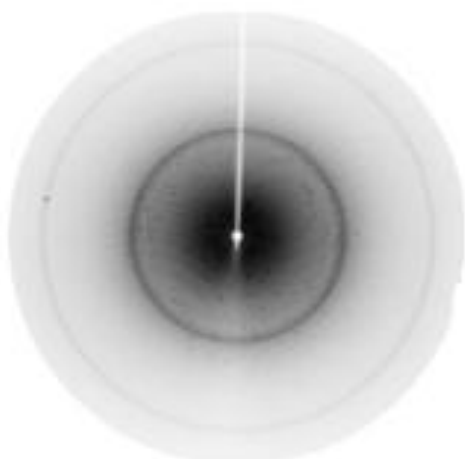
B)



C)



DI)

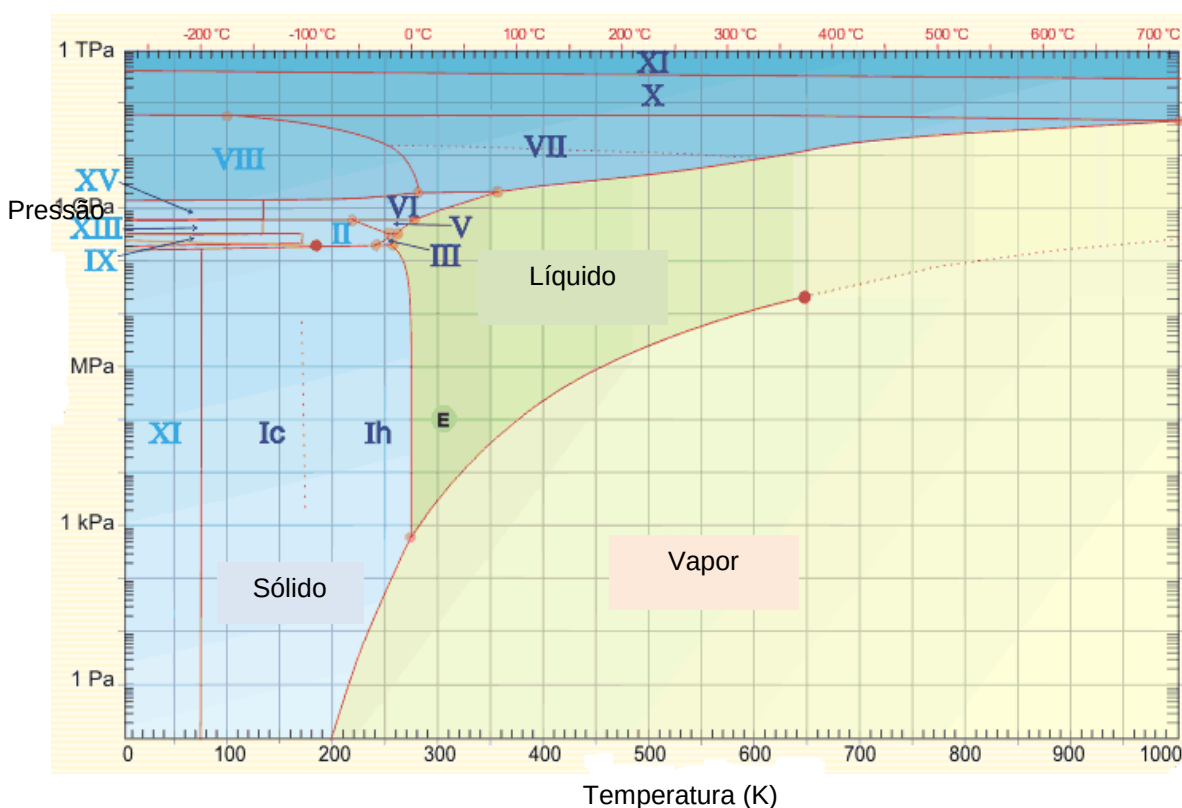


DII)



Particularmente, para o conjunto XFG3A3B_NATB4 observamos que as imagens 1 até 118 apresentam 3 anéis de difração por gelo e as imagens 119 até 214 apresentam 6 anéis, Podemos citar duas explicações, uma das possibilidades seria a transição da forma de gelo cúbica para a forma hexagonal, como comentado na introdução deste trabalho, o diagrama de fases da água (Figura 8) mostra que é possível essa transição próximo à temperatura de 170K, outra explicação seria que no início da coleta dos dados de difração ocorreu a formação do gelo de forma hexagonal não tão bem formado e com isso observamos 3 anéis mais intensos no padrão de difração e após um certo tempo de exposição do cristal a difração os cristais hexagonais ficaram mais evidentes e difrataram 6 anéis no padrão de difração.

Figura 8: Diagrama de fases para a água, mostrando o estado líquido, sólido dividido em (XI) gelo vítreo, (Ic) forma cúbica, (Ih) forma hexagonal. Fonte: <<http://www.lsbu.ac.uk/water/phase.html>>. Acesso em 21/04/2013 >.



Para se obter bons dados foi preciso excluí-los durante o processamento e escalonamento das imagens, com uma espessura ideal que dependia da intensidade relativa e da quantidade de gelo presente. Realizou-se uma busca na literatura para identificar as localizações e intensidades relativas dos anéis como mostrada na Tabela 1 e assim deduziu-se que convém excluir a difração pelas formas cúbica e hexagonal em imagens específicas. Assim, com esses dados, *scripts* foram preparados para realizar o estudo de espessura ideal (dados não mostrados) a ser excluída em cada conjunto, levando-se em consideração que a ideia é excluir o mínimo possível de dados com boa qualidade. A Tabela 3 mostra quais as regiões finalmente excluídas de cada conjunto de dados.

Tabela 3: Espessuras dos anéis de difração por gelo considerado como melhores para exclusão de cada conjunto de imagens separadamente.

Conjunto de imagens	Imagens	Espessura dos anéis de difração por gelo (Å)	Forma de gelo
XFG3A3B_NATB4	1 – 118	3,91 - 3,86	Cúbico
		3,68 – 3,64	
		2,25 – 2,21	
	119 - 214	3,91 - 3,86	Hexagonal
		3,68 – 3,64	
		3,47 – 2,63	
		2,25 – 2,21	
		2,07 – 2,03	
	1 - 289	3,93 – 3,86	Cúbico
		3,68 – 3,64	
		2,25 – 2,21	
XfG3AN2_NatMnaBA2	1 - 253	3,97 – 3,85	Hexagonal
		3,70 – 3,60	
		3,48 – 3,39	
		2,69 – 2,65	
		2,25 – 2,23	
		2,07 – 2,05	
XfG3AN2_NatMnB_B5	1 - 197	3,97 - 3,85	Hexagonal
		3,72 – 3,61	
		3,51 – 3,41	
		2,68 – 2,62	
		2,25 – 2,23	
		2,07 – 2,05	

A indexação mostrou que estes cristais pertencem ao mesmo grupo espacial C2, todavia, vários com parâmetros de cela diferentes. A Tabela 4 apresentamos relevantes obtidos do processamento. A partir do conjunto original de reflexões de cada estrutura, aproximadamente 5% das reflexões foram marcadas para serem utilizadas no cálculo de R_{free} .

Tabela4: Resumo dos dados de processamento para as sete estruturas.

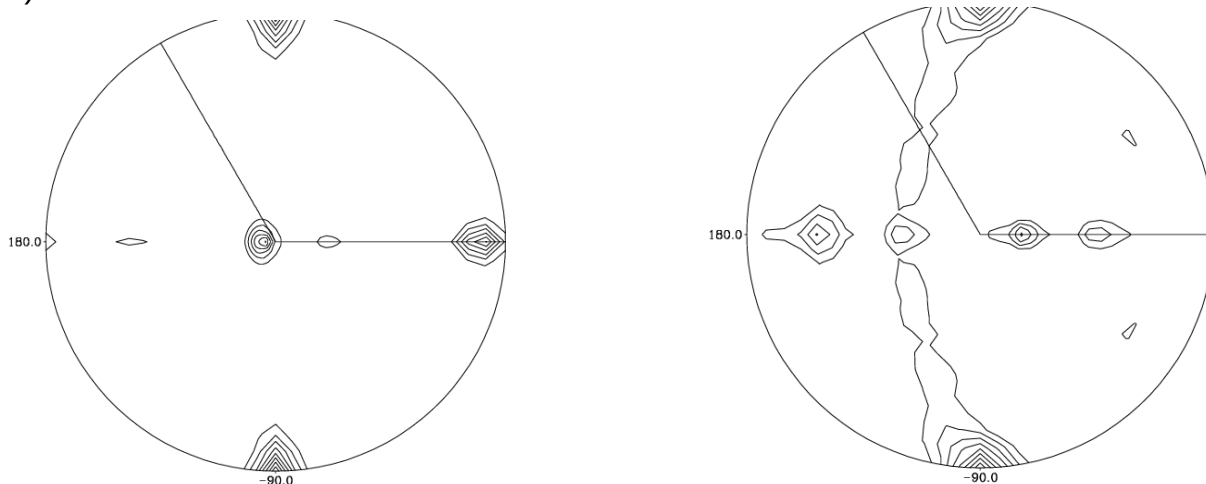
	4 monômeros na unidade assimétrica				2 monômeros na unidade assimétrica		
	XFAMPreso24	XFAMPreso29	XFG3A3B_NATB4	XFG3A3oriC4cen	XFG3A3ampD5_I	XfG3AN2_NatMnaBA2	XfG3AN2_NatMnB_B5
N° de imagens	256	172	214	330	289	253	197
Porcentagem de solvente	44,26	44,26	45,37	55,69	53,32	43,30	43,30
Coeficiente de Mathews	2,21	2,21	2,25	2,77	2,63	2,17	2,17
Resolução Å	81,63/2,76	81,59/2,94	81,59/2,02	86,74/1,96	76,37/2,62	71,18/2,70	71,16/2,92
Faixa de resolução	2,84/2,76	3,03/2,94	2,08/2,02	2,02/1,96	2,71/2,63	2,82/2,74	2,93/2,85
Grupo de espaço	C2	C2	C2	C2	C2	C2	C2
Parâmetros de cela							
a/ Å	159,46	158,10	159,43	172,42	92,61	88,03	86,84
b/ Å	78,19	78,80	78,79	84,20	85,22	81,18	79,85
c / Å	82,03	82,80	82,03	87,37	71,36	72,06	71,33
$\alpha / ^\circ = \gamma / ^\circ$	90	90	90	90	90	90	90
$\beta / ^\circ$	96,89	95,52	96,89	96,58	94,32	95,95	95,51
N° reflexões observadas	136682	76204	236002	31660	123383	56674	36598
N. reflexões únicas	24993	20737	55515	88536	17166	11091	9802
Completeza %	97,4 (92,1)	97,8 (94,0)	95,9 (87,9)	95,7 (99,5)	96,7 (97,4)	97,2 (81,7)	88,5 (82,5)
Redundância	5,46	3,67	4,05	4,76	7,18	5,08	3,76
Rsym/%	9,6 (57,4)	45,3 (7,8)	7,8 (49,7)	4,9 (33,8)	6,9 (53,6)	9,5 (57,5)	10,9 (47,7)
Rmeas /%	10,6 (63,7)	9,1 (53,4)	9,0 (61,8)	5,5 (38,5)	7,5 (58,1)	10,7 (63,6)	13,1 (51,0)
$\langle I/\sigma(I) \rangle$	3,40 (16,58)	3,14 (15,58)	2,22 (11,75)	4,80 (17,27)	4,11 (19,05)	2,48 (12,99)	3,25 (8,67)
Fator B do gráfico de Wilson/ Å ²	60,66	65,67	35,42	29,71	63,82	65,93	66,36
N° parâmetros refinados	31965	31413	31505	34665	16341	15301	15445

5.1.1.1 Função de auto-rotação

A Figura 9 mostra mapas de contorno da função de auto-rotação para a estrutura tetramérica (a) e dimérica (b) com apenas κ (kappa) igual a 180° (único com valor relevante), para revelar existência e direção das simetrias não cristalográficas. Observamos picos relevantes e localização (θ , Φ) aproximada: a) para o tetrâmero, com alturas relativas de 70,7, a $83,9^\circ$ e $0,0^\circ$ e a $5,9^\circ$ e $180,0^\circ$; b) para o dímero, com altura relativa de 51,2, a $69,4^\circ$ e $180,0^\circ$.

Figura 9: Função de auto-rotação para κ igual a 180° . Contornos iniciam em altura relativa 20 e estratificam-se a cada 10 unidades, até o máximo de 100. A) Para o modelo de tetrâmero; B) para o modelo de dímero.

A)B)



5.1.2 Substituição molecular

A estrutura de referência para a substituição molecular foi a SurE de *Coxiella burnetti* (CbSurE) (código PDB 3TY2) que apresenta 58% de identidade e 74% de similaridade. A Figura 10 mostra o alinhamento realizado pelo programa Clustalw2 e representado no programa Aline (BOND et al., 2009). Para a estrutura XFG3A3oriC4cen, encontraram-se 4 soluções de função de rotação com bom contraste e escore, e para a estrutura XFG3A3ampD5_I, 2 soluções, as quais produziram depois soluções unívocas de translação, confirmando sua correção.

A Tabela 5 mostra as melhores soluções sugeridas para rotação, algumas acrescentadas para realçar o contraste.

Tabela 5: Soluções de rotação encontradas pelo programa de substituição molecular Phaser (AFONINE *et al.*, 2009); A) para a estrutura XFG3A3oriC4cen; B) para a estrutura XFG3A3ampD5_I.

A)

Soluções	Ângulo de rotação (α , β , γ).	Escore Z
1°	88,480/ 86,987/ 214,542	7,95
2°	268,629/ 85,538/ 229,198	7,71
3°	91,666/ 82,260/ 216,007	7,56
4°	270,057/ 82,833/ 226,692	7,44
5°	269,432/ 86,803/ 49,254	5,55
6°	267,555/ 85,812/ 235,001	4,65
7°	91,027/ 87,706/ 26,925	4,62

B)

Soluções	Ângulo de rotação (α , β , γ).	Escore Z
1°	93,584/ 85,147/ 149,006	9,05
2°	265,559/ 87,501/ 293,637	7,63
3°	99,892/ 87,824/ 148,026	7,12
4°	262,769/ 87,824/ 285,167	5,68
5°	93,206/ 87,752/ 323,771	5,45
6°	102,915/ 88,250/ 150,765	5,43
7°	80,954/ 87,093/ 334,267	5,11

Figura 10: Alinhamento de sequência da proteína SurE de *Xylella fastidiosa* contra a de SurE de *Coxiella burnetii* (3TY2), os cilindros representam as α -hélices e as setas são as folhas- β da última.

SurE de <i>Xylella fastidiosa</i>	1	5	10	15	20	25	30	35																															
	M	R	V	L	V	S	N	D	D	G	V	D	A	P	G	I	K	I	L	A	D	A	L	R	N	A	G	H	E	V	M	V	V	A	P	D	R	37	
SurE de <i>Coxiella burnetii</i>	K	L	R	L	L	S	N	D	D	G	V	Y	A	K	G	L	A	I	L	A	K	T	L	A	D	L	G	.	E	V	D	V	V	A	P	D	R	37	
SurE de <i>Xylella fastidiosa</i>	40	45	50	55	60	65	70	75																															
	D	R	S	G	A	S	N	S	L	T	L	D	T	P	I	R	A	K	Q	I	D	M	H	T	Y	S	V	A	G	T	P	T	D	C	V	H	L	A	75
SurE de <i>Coxiella burnetii</i>	N	R	S	G	A	S	N	S	L	T	L	N	A	P	L	H	I	K	N	L	E	N	G	M	I	S	V	E	G	T	P	T	D	C	V	H	L	A	75
SurE de <i>Xylella fastidiosa</i>	80	85	90	95	100	105	110																																
	L	T	G	L	L	N	Y	D	P	D	I	V	V	S	G	I	N	N	T	G	N	L	G	D	D	V	I	Y	S	G	T	V	S	A	A	M	E	G	113
SurE de <i>Coxiella burnetii</i>	I	T	G	V	L	P	E	M	P	D	M	V	V	A	G	I	N	A	G	P	N	L	G	D	D	V	W	Y	S	G	T	V	A	A	M	E	G	113	
SurE de <i>Xylella fastidiosa</i>	115	120	125	130	135	140	145	150																															
	R	F	L	G	L	P	A	V	A	V	S	L	V	T	L	Y	R	E	G	Q	Q	A	P	Q	Y	E	T	A	A	H	A	A	I	N	I	V	A	Q	151
SurE de <i>Coxiella burnetii</i>	R	F	L	G	L	P	A	L	A	V	S	L	G	.	.	.	.	G	E	L	F	R	Y	Y	E	T	A	A	K	V	V	Y	Q	L	I	Q	R	146	
SurE de <i>Xylella fastidiosa</i>	155	160	165	170	175	180	185																																
	L	K	T	D	P	L	P	A	D	T	I	L	N	V	N	V	P	D	V	T	W	Q	Q	M	R	G	F	K	V	T	R	L	G	N	R	H	R	S	189
SurE de <i>Coxiella burnetii</i>	I	E	K	D	P	L	P	P	S	T	I	L	N	I	N	V	P	D	L	P	Y	E	E	L	K	G	F	E	V	T	R	L	G	T	R	H	R	A	184
SurE de <i>Xylella fastidiosa</i>	190	195	200	205	210	215	220	225																															
	A	P	C	L	T	Q	T	D	P	R	G	H	T	I	Y	W	I	G	P	A	G	P	E	Q	D	A	G	P	G	T	D	F	D	A	V	R	N	T	227
SurE de <i>Coxiella burnetii</i>	E	P	T	I	R	Q	I	D	P	R	G	H	P	I	Y	W	V	G	A	A	G	P	E	Q	D	S	G	P	G	T	D	F	F	A	M	N	H	H	222
SurE de <i>Xylella fastidiosa</i>	230	235	240	245	250	255	260																																
	Y	I	S	I	T	P	I	H	V	D	L	T	R	Y	Q	A	L	E	N	V	T	R	W	T	D	R	L	T	A	H	M	D	W	P	T		262		
SurE de <i>Coxiella burnetii</i>	C	V	S	I	T	P	L	R	V	D	L	T	H	Y	E	A	F	D	Q	L	A	S	W	V	K	R	L	E	M	.	.	.	.	.	.		251		

5.1.3 Refinamento das estruturas cristalográficas

Após a substituição molecular, refinaram-se as estruturas dos conjuntos XFG3A3oriC4cen e XFG3A3ampD5_I. Inicialmente, realizou-se um refinamento de corpo rígido, seguido de um refinamento com restrições e para retirar qualquer tendenciosidade do modelo provindo da substituição molecular realizou-se um refinamento com Simulação "annealing". Os valores de R e R_{free} para estes passos estão relatado na Tabela 6.

Tabela 6: Dados estatísticos do refinamento inicial de corpo rígido, com restrições e simulação "annealing", para as duas estruturas refinadas.

Estruturas	Refinamento de corpo rígido		Refinamento com restrições		Refinamento Simulação "annealing"	
	R	R_{free}	R	R_{free}	R	R_{free}
XFG3A3oriC4cen	0,3282	0,3580	0,2704	0,2862	0,2147	0,2558
XFG3A3ampD5_I	0,3043	0,3652	0,2496	0,3003	0,2104	0,2911

Durante o processo de refinamento foram otimizados os pesos relativos aos dados experimentais e restrições geométricas e a região de máscara para o "bulk solvent". Também, vários modelos para TLS (translação, libração, parafuso) e NCS (simetria não cristalográfica) foram testados. Decidiu-se usar para a estrutura XFG3A3oriC4cen um único domínio para tensores de TLS compreendendo para cada cadeia polipeptídica como corpo rígido mediante estudo feito no servidor TLMSD (PAINTER *et al.*, 2006), também se decidiu não usar NCS. Já para a estrutura XFG3A3ampD5_I decidiu-se usar para cada cadeia dois domínios para TLS, incluindo-se região do C-terminal na contra cadeia para a qual ela se projeta. No tocante ao NCS, para essa estrutura, restringiu-se apenas a cadeia principal das regiões que se verificou que eram semelhantes. A Tabela 7 mostra os valores estatísticos da estrutura final refinada bem como outros parâmetros utilizados.

Tabela 7: Valores estatísticos de refinamento para ambas as estruturas refinadas.

	XFG3A3oriC4cen	XFG3A3ampD5_I
rmsbl / Å	0,015	0,011
rmsba / °	1,387	1,347
rmscc / Å ³	0,096	0,084
Fator B médio / Å	15,58	24,38
R_{free}	0,2432	0,2605
R	0,2010	0,1886
N° resíduos	1026	509
N° moléculas de água	733	169

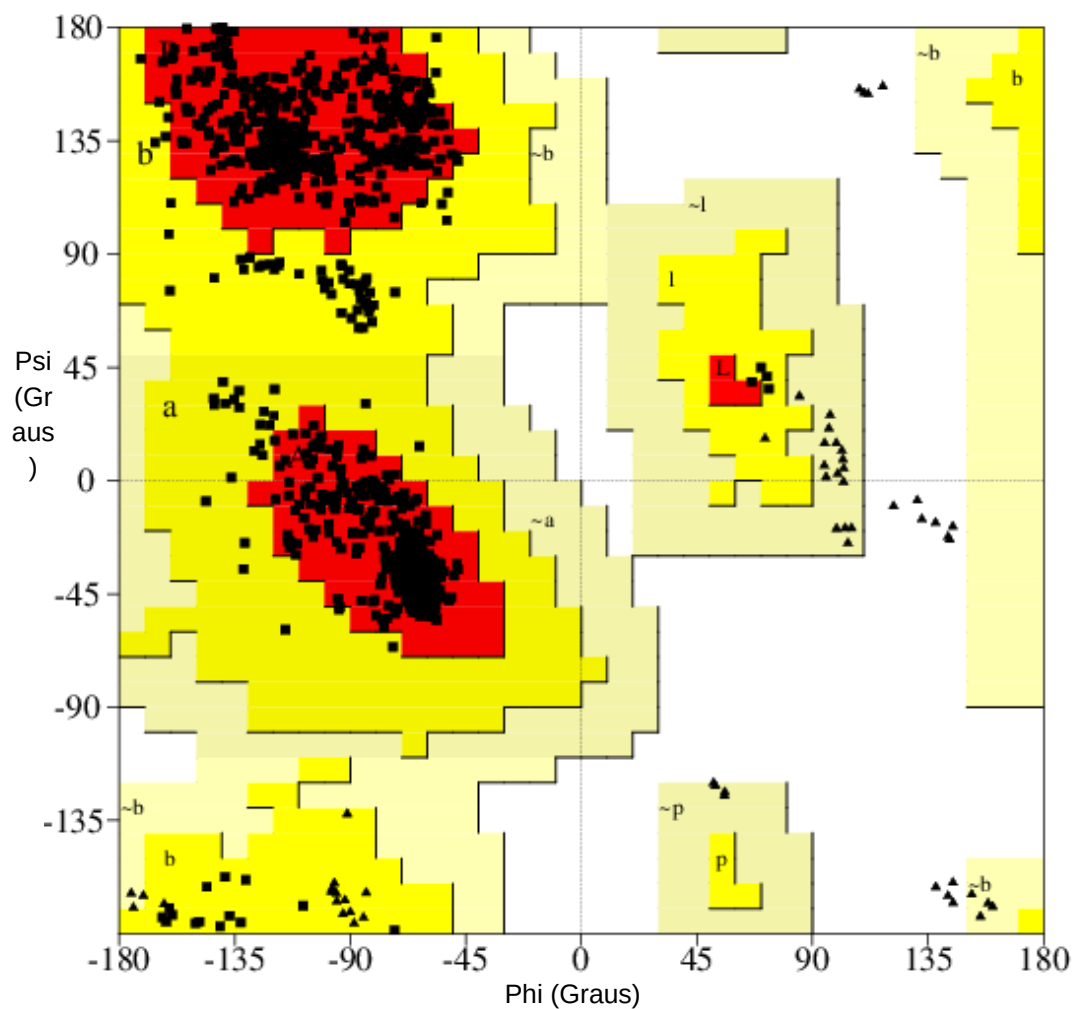
rmsbl: desvios quadráticos de comprimentos de ligação; *rmsba*: desvios médios quadráticos de ângulos de ligação; *rmscc*: desvios médios quadráticos de centros quirais.

5.1.3 Validação das estruturas refinadas

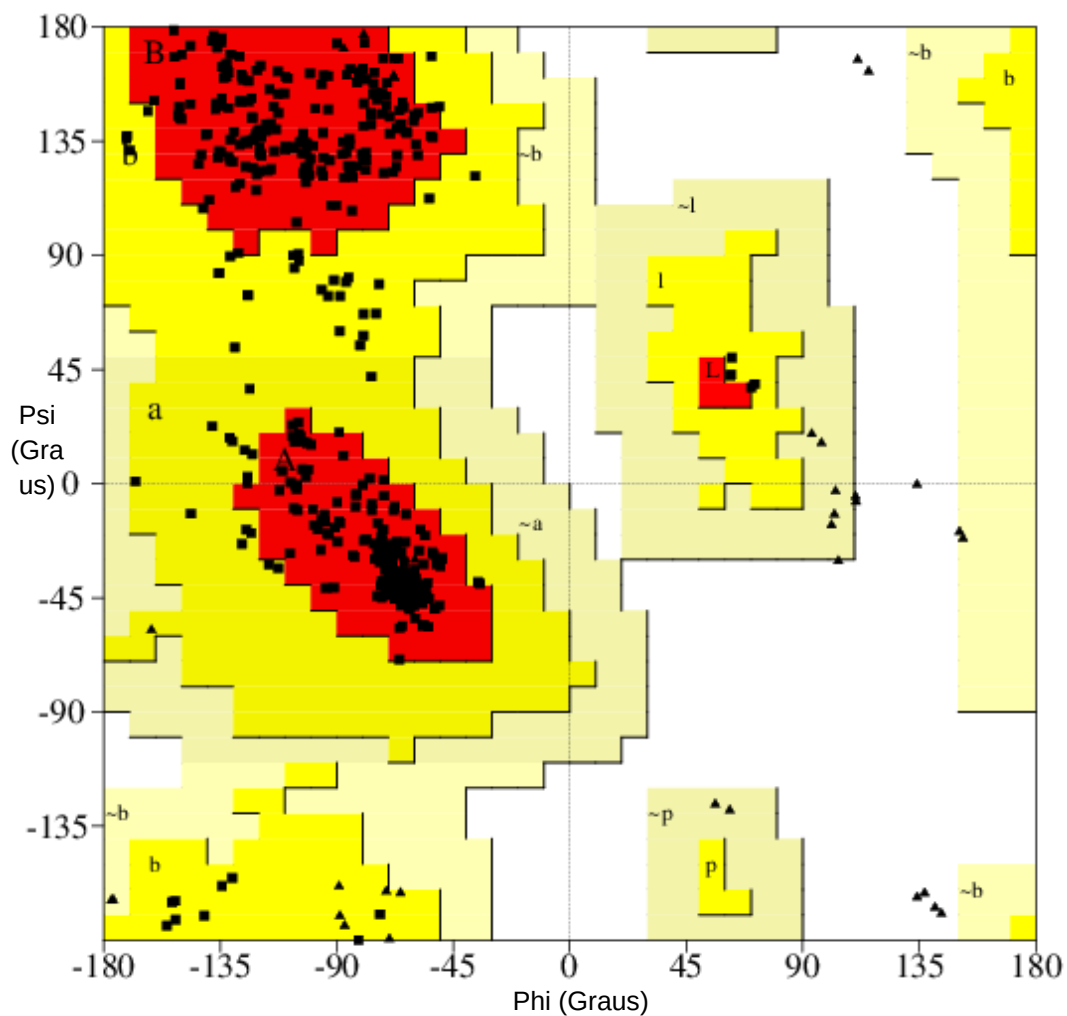
A qualidade estereoquímica do modelo foi estimada utilizando-se os programas Procheck, Whatcheck, Molprobity e ValidationPDB. Os gráficos de Ramachandran (Figura 11), como calculado pelo procheck, para as estruturas XFG3A3oriC4cen, XFG3A3ampD5_I apresentam, respectivamente, 88,9% e 86,3% dos resíduos nas regiões favorável e permitida e os demais na região generosamente permitida.

Figura 11: Gráfico de Ramachandran para as estruturas; A) XFG3A3oriC4cen e B) XFG3A3ampD5_I. As regiões em vermelho são as permitidas, as amarelas adicionalmente permitidas, em amarelo claro generosamente permitido e em branco proibidas.

A)



B)



A Tabela 8 mostra o resumo da validação geométrica das duas estruturas obtida com o programa MolProbity.

Tabela8:A) Resumo da validação da estrutura XFG3A3oriC4cen; B) Resumo da validação da estrutura XFG3A3ampD5_I.

A)

Geometria da proteína	Rotâmeros pobres	2.13%	Meta: <1 %
	Ramachandran proibido	0.00%	Meta: <0,2 %
	Ramachandran favorecido	97.43%	Meta: >98 %
	C β desvios >0,25 Å	4	Meta: 0
	Resíduos com mas ligações	0.00%	Meta: 0 %
	Resíduos com maus ângulos	0.29%	Meta: <0,1 %

B)

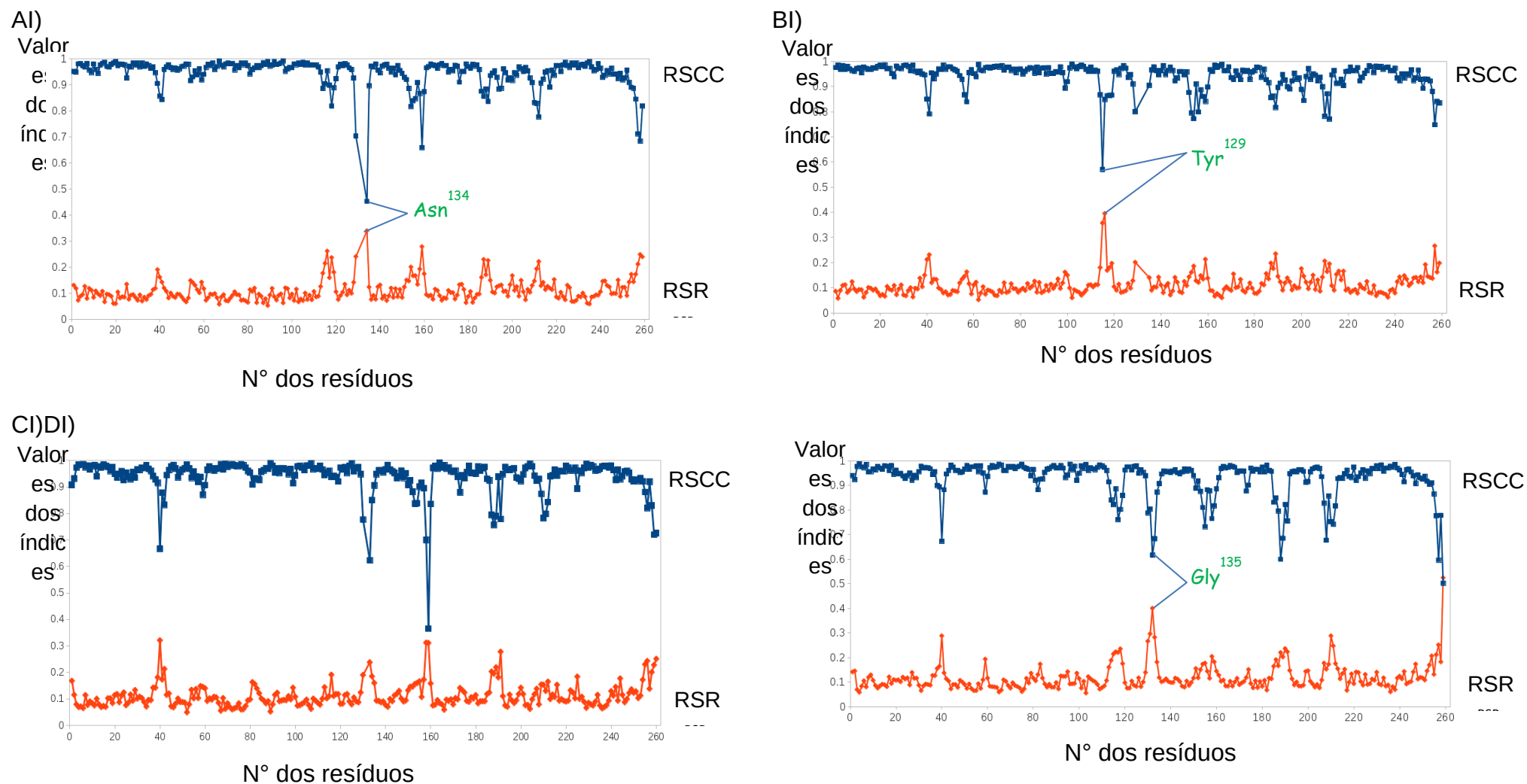
Geometria da proteína	Rotâmeros pobres	3.80%	Meta: <1%
	Ramachandran proibido	0.40%	Meta: <0,2 %
	Ramachandran favorecido	95.41 %	Meta: >98 %
	C β desvios > 0,25 Å	3	Meta: 0
	Resíduos com mas ligações	0.00%	Meta: 0 %
	Resíduos com maus ângulos	0.00%	Meta: <0,1 %

Em nem todos os critérios o objetivo (goal), como considerado pelo molprobit, foi alcançado. Vale lembrar que os pesos entre restrições geométricas e dados experimentais foram otimizados, mas há uma considerável quantidade de ruído em várias regiões dos mapas. Todos os "outliers" foram novamente verificados e procurou-se fazer o melhor. Particularmente, para os rotâmeros "outliers" observa-se que em ambas as estruturas são praticamente os mesmos resíduos, exemplo o Asp⁸, que coordena o íon Mn²⁺, com densidade razoável, situação parecida para resíduo Gln²⁴² que coordena o íon I⁻. Já a Arg¹¹⁴ está localizada em uma região ruidosa, porém a densidade na cadeia lateral é relativamente forte para todos os seis monômeros, indicando tal rotâmero incomum.

Sobre os ângulos ruins de XFG3A3oriC4cen, dos resíduos Gly¹¹⁷, Pro¹³⁶ e Pro¹⁹¹ observaram ruído ao redor além da flexibilidade, embora se argumente no sentido de se permitir a modelagem, já para a Pro¹⁹¹, a densidade corrobora a posição (e ângulo) do resíduo.

Para cada estrutura realizaram-se cálculos dos índices RSR (Valor residual no espaço real) e RSCC (Valor do coeficiente de correlação do espaço real) com o programa Mapmam (KLEYWEGT ET *et al.*, 1996). A Figura 12 mostra gráficos para cada monômero de ambas as estruturas e por resíduo, enquanto que a Tabela 9 trazos valores globais de RSR e RSCC para cada cadeia.

Figura 12: Qualidade do ajuste na densidade eletrônica representado em gráfico; AI, BI, CI e DI) estrutura XFG3A3oriC4cen para cada monômero, nos gráficos as setas indicam quais resíduos em regiões de alta flexibilidade; AII e BII) estrutura XFG3A3ampD5_I.



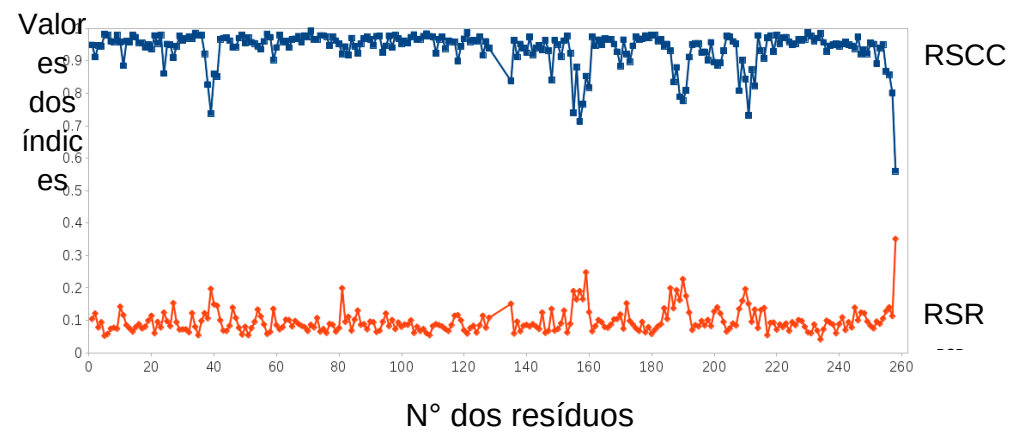
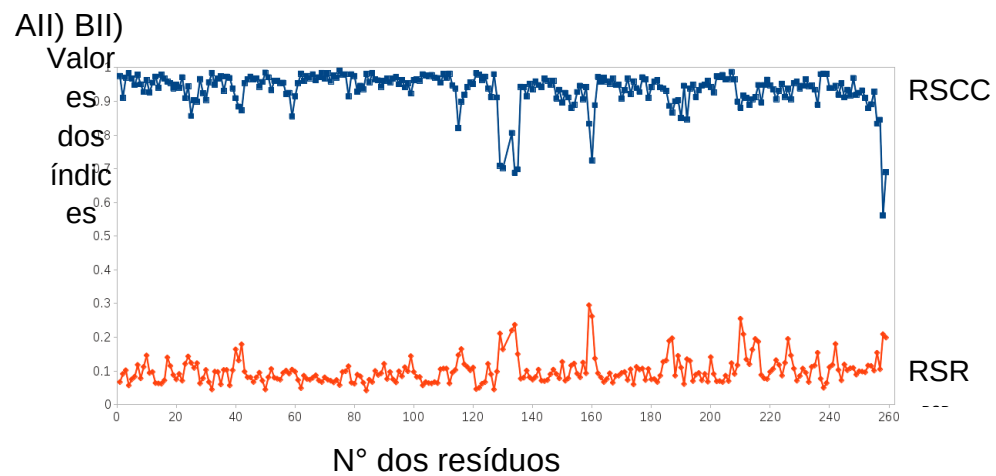


Tabela 9: Valores globais de RSR e RSCC por cadeia

Cadeia	XFG3A3oriC4cen		XFG3A3ampD5_I	
	RSCC	RSR	RSCC	RSR
A	0,9347	0,1370	0,9330	0,1248
B	0,9360	0,1374	0,9352	0,1219
C	0,9310	0,1403	-	-
D	0,9229	0,1467	-	-
Todas as cadeias	0,9317	0,1403	0,9348	0,1229

Nos gráficos observam-se regiões onde os valores de RSR apresentam-se altos e os de RSCC apresentam-se baixos; estão relacionadas às regiões flexíveis da proteína, onde se procurou fazer a melhor modelagem possível. Particularmente, entre os resíduos 130 a 134 a flexibilidade é ainda maior, o que levou na maior parte dos casos a não modelá-los, variando em um ou dois resíduos a mais ou a menos conforme a estrutura, exceção apenas o monômero D da estrutura XFG3A3oriC4cenque, embora a flexibilidade apresente pelo menos uma fraca continuidade da cadeia principal que foi, então, modelada.

Para delimitar essas regiões de alta flexibilidade e decidir até onde se modelou, utilizaram-se os procedimentos principais:

- Realizou-se um “*annealing computacional*” dando energia e permitindo mobilidade para um possível ajuste desses resíduos;

- Tentou-se colocar os resíduos faltantes nas regiões verificando-se se os mapas de densidade eletrônica posteriormente corroboravam a presença. Em geral, exigia-se um contorno de pelo menos $0,5 \sigma$ nos mapas de densidade eletrônica para o posicionamento.

5.1.5 Análises estruturais

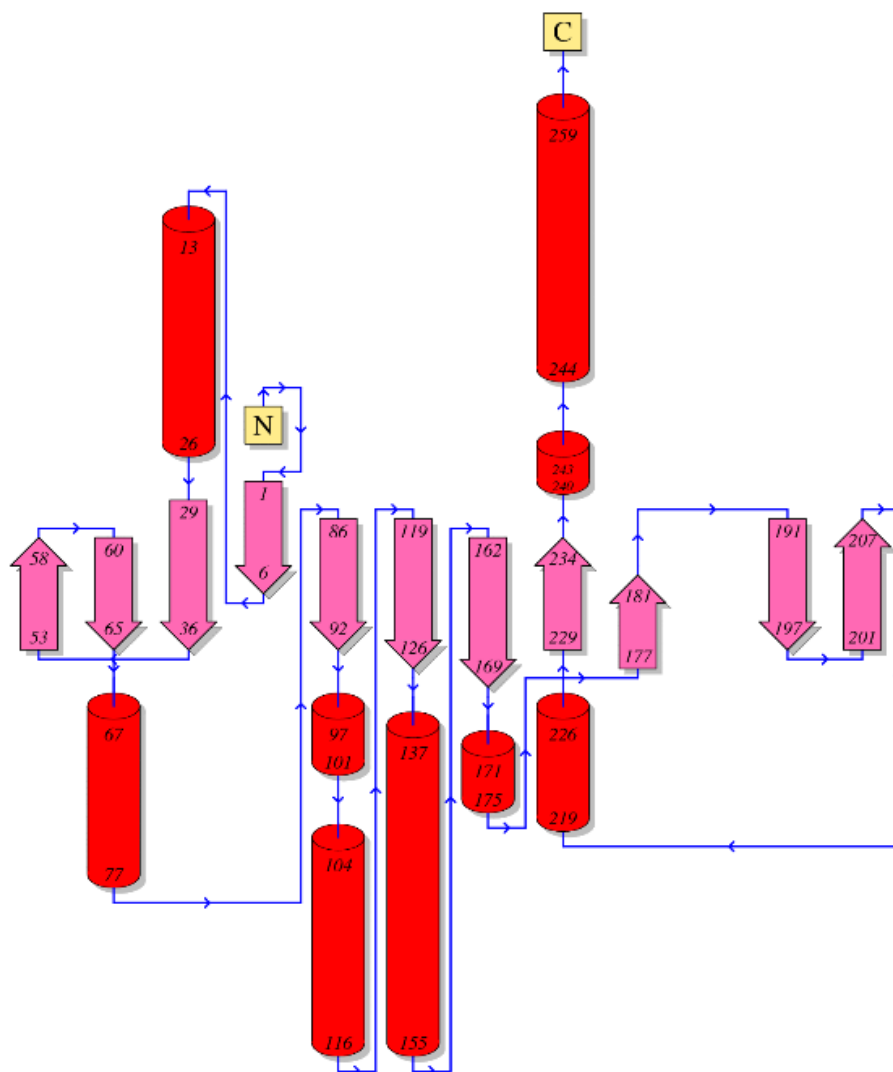
Em XFG3A3oriC4cen refinaram-se 6 íons I^- , 4 íons Mn^{2+} e 4 íons PO_4^{3-} , e em XFG3A3ampD5_I refinaram-se 3 íons I^- , 2 íons Mn^{2+} e 2 íons PO_4^{3-} . Ambas as estruturas iniciam com uma fita β pelo resíduo Met¹ e terminam no C-terminal em α -

hélice, cujo final é uma região relativamente flexível, sendo que seus resíduos apresentam-se desordenados.

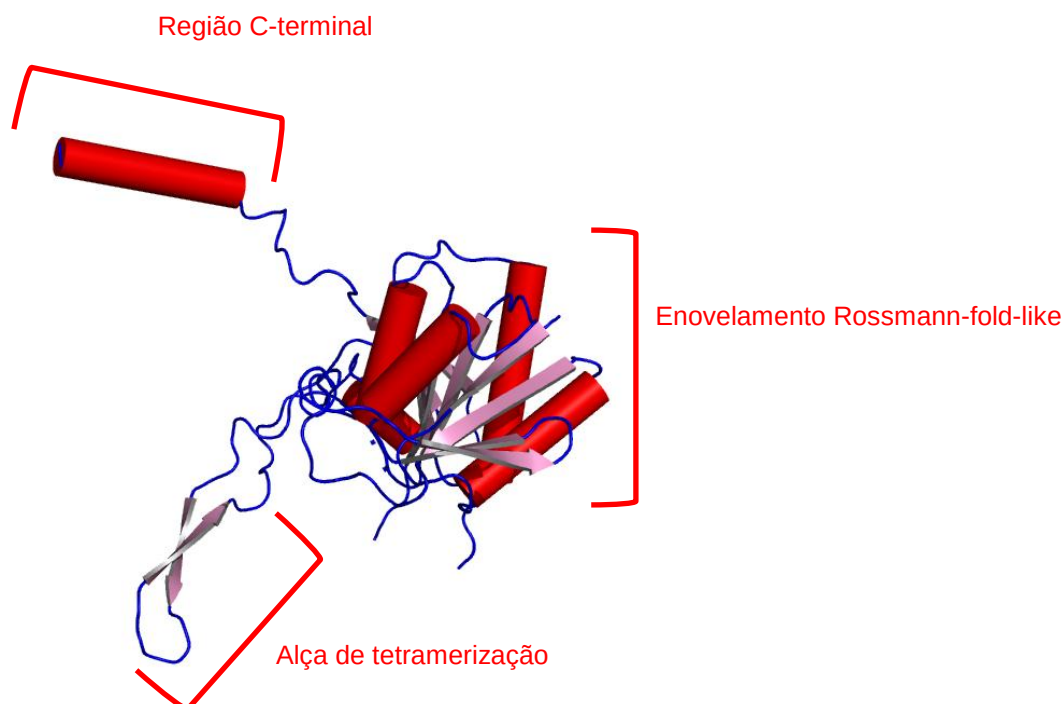
O uso da ferramenta Pdbsum (LASKOWSKI, 2009) contra a base de dados SCOP (CONTE *et al.*, 2002) indica que o enovelamento é do tipo *Rossmann-fold-like* $\alpha/\beta/\alpha$ e que a proteína pertence à família de SurE-Like, como seria de se esperar. A Figura 13 mostra um diagrama de topologia, explicitando os elementos de estrutura secundária, também produzido pela ferramenta Pdbsum para um monômero.

Figura 13: A) Diagrama da topologia da XfSurE; B) ilustração da estrutura para um monômero mostrando o enovelamento do tipo *Rossmann-fold-like* $\alpha/\beta/\alpha$. Em forma de bastão vermelho estão as α -hélices, em seta rosa folhas- β , em azul "loops".

A)



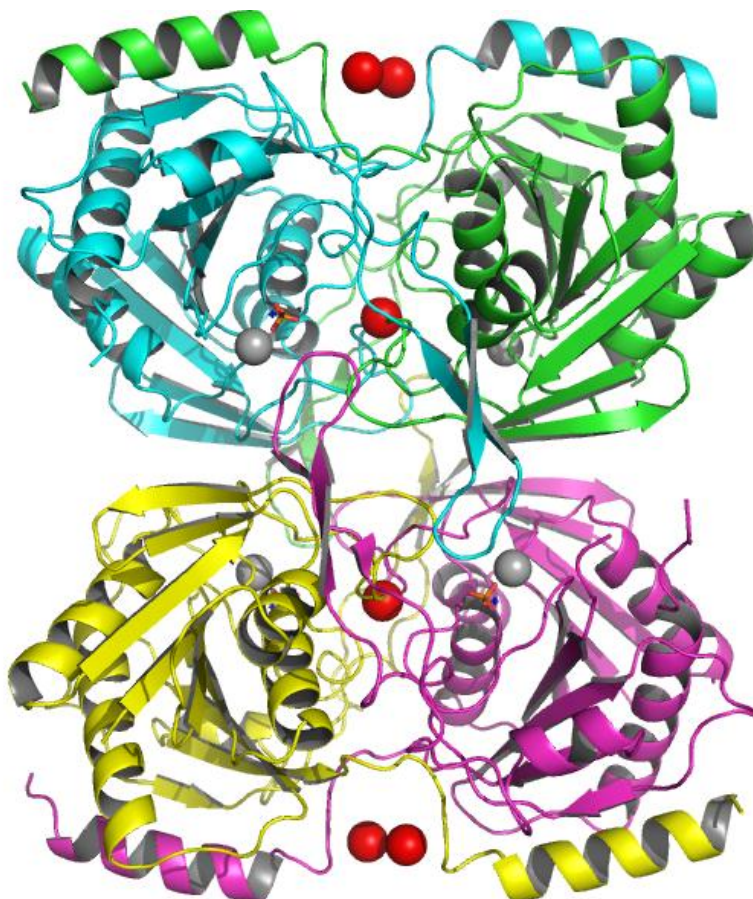
B)



A Figura 14 mostra o modelo estrutural do tetrâmero para XFG3A3oriC4cen e do dímero para XFG3A3ampD5_I como na unidade assimétrica, com cada monômero representado por uma cor distinta. Há um sítio ativo correspondente a cada monômero. O tetrâmero existente biologicamente deve corresponder ao da unidade assimétrica para XFG3A3oriC4cen e pode ser construído usando-se o eixo de ordem 2 cristalográfico para XFG3A3ampD5_I, em cujo caso chamar-se-ão os monômeros simetricamente equivalentes de C e D. Em qualquer caso, o monômero A participa de interações amplas com o monômero B, assim como o monômero C participa de interações amplas com o monômero D. O conjunto AB interage com o conjunto CD inclusive pelas alças de tetramerização. As áreas de superfície de contato entre os monômeros e o número de resíduos que estão em contato estão mostradas na Tabela 10, como obtido com programa PISA (KRISINEL *et al.*, 2007). A Figura 15, produzida com a ferramenta Pdbsum, esquematiza de forma proporcional a quantidade de superfície de contato entre os monômeros.

Figura 14: Estrutura Tridimensional da XfSurE; A) XFG3A3oriC4cen; B) XFG3A3ampD5_I. Em ambos os casos as esferas vermelhas representam os íons I^- e as esferas cinza os íons Mn^{2+} . Os íons PO_4^{3-} , representados como bastões, estão coordenando os íons Mn^{2+} .

A)



B)

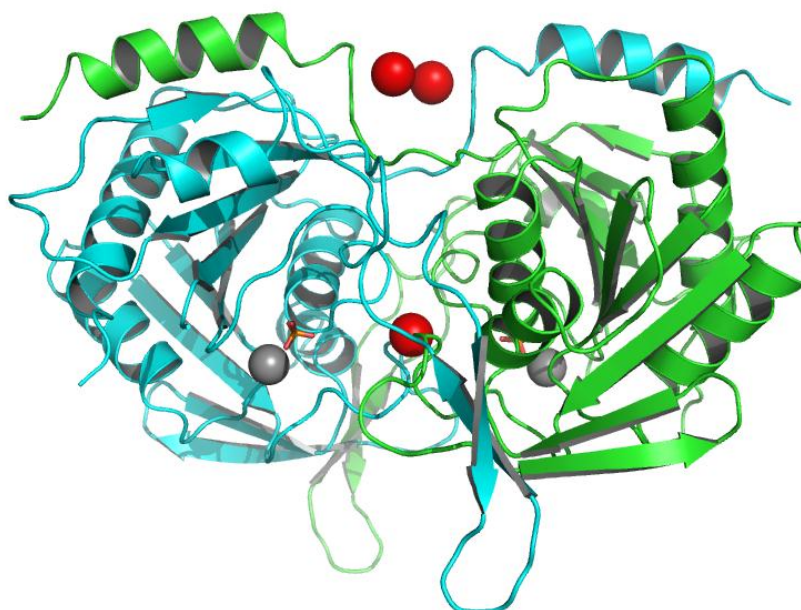
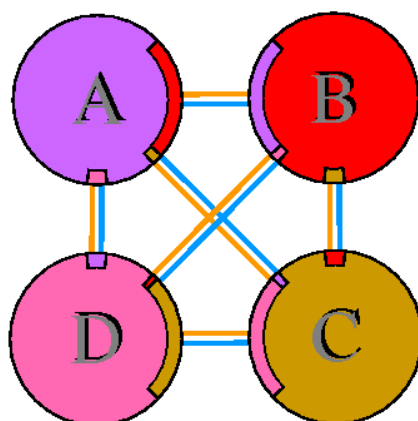


Tabela 10: Área de superfície e quantidades de resíduos em contatos intermonômeros de acordo com o programa PISA

Monômeros	XFG3A3oriC4cen		XFG3A3ampD5_I	
	Nº resíduos	Área de interface (Å ²)	Nº resíduos	Área de interface (Å ²)
C com D	97	3939,8	97	-
A com B	93	3799,7	97	3752,5
B com C	17	602,1	18	-
D com A	18	575,8	17	-
C com A	12	329,0	13	-
D com B	12	302,9	13	-

Figura 15: Regiões relativas de contato entre os monômeros no tetrâmero, produzida com a ferramenta Pdbsum (LASKOWSKI, 2009).



Podemos verificar que as áreas de contato entre os monômeros das duas estruturas são praticamente iguais, assim como praticamente são os mesmos resíduos que estão em contato (dados não mostrados).

5.1.5.1 Sítio ativo

No sítio ativo da XfSurE (Figura 16) há o íon Mn^{2+} , oriundo da presença de $MnCl_2$ durante a cristalização. Estudos realizados em (SARAIVA *et al.*, 2009) revelam que a XfSurE apresenta uma alta afinidade pelo íon Mn^{2+} , sendo este considerado um cofator. Ressalte-se que para cada monômero há um sítio ativo, localizado centralmente no tetrâmero, em região afetada pela mudança na interface intermonômeros (SARAIVA *et al.*, 2009). A Figura 17 mostra a interação desse cofator, Mn^{2+} , com os resíduos Asp⁸, Asp⁹, Asn⁹², Asn⁹³ do monômero correspondente e a coordenação dele pelo PO_4^{3-} . Os íons PO_4^{3-} provavelmente vieram da desfosforilação do 3-AMP co-cristalizado com cada estrutura e tiveram suas ocupações refinadas de forma alternada com uma molécula de água, como mostrado na Figura 17. A Tabela 11 lista os valores de B_{factor} médios para cada íon, para os resíduos e para as águas ao redor destes até 6 Å. No caso dos íons e da água alternada ao fosfato, mostra também as ocupações. Tais valores foram os considerados após vários testes.

Comparando as estruturas refinadas com estruturas da mesma família, analisamos que as interações do Mn^{2+} e PO_4^{3-} são relativamente conservadas e na maioria delas esses dois íons encontram-se coordenados pelos mesmos resíduos na mesma posição. Isso corrobora a alta afinidade da proteína pelo Mn^{2+} , como relatado na revisão bibliográfica.

Figura 16: Sítio ativo em apenas um monômero da estrutura da XfSurE.



Figura 17: A), B), C) e D) Interação dos íons para cada monômero (A, B, C e D, respectivamente) da estrutura XFG3A3oriC4cen; e) e f) Interação dos íons respectivamente para cada monômero (A e B, respectivamente) da estrutura XFG3A3ampD5_I. Em azul, nível de contorno 1 σ para o mapa de densidade eletrônica. Em verde, nível de contorno 3 σ para o mapa Fourier diferença. Em vermelho, nível de contorno -3 σ para o mapa Fourier diferença.

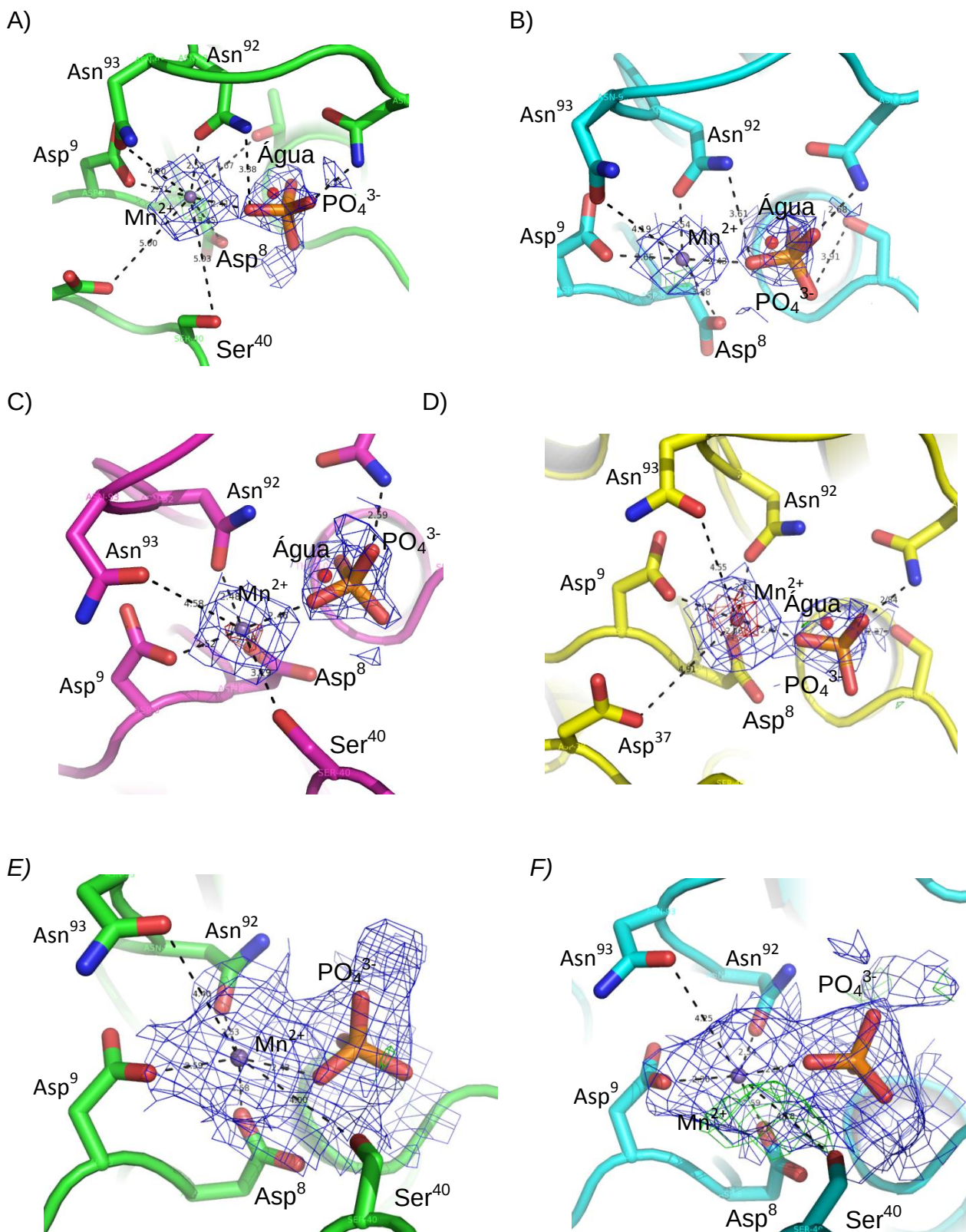


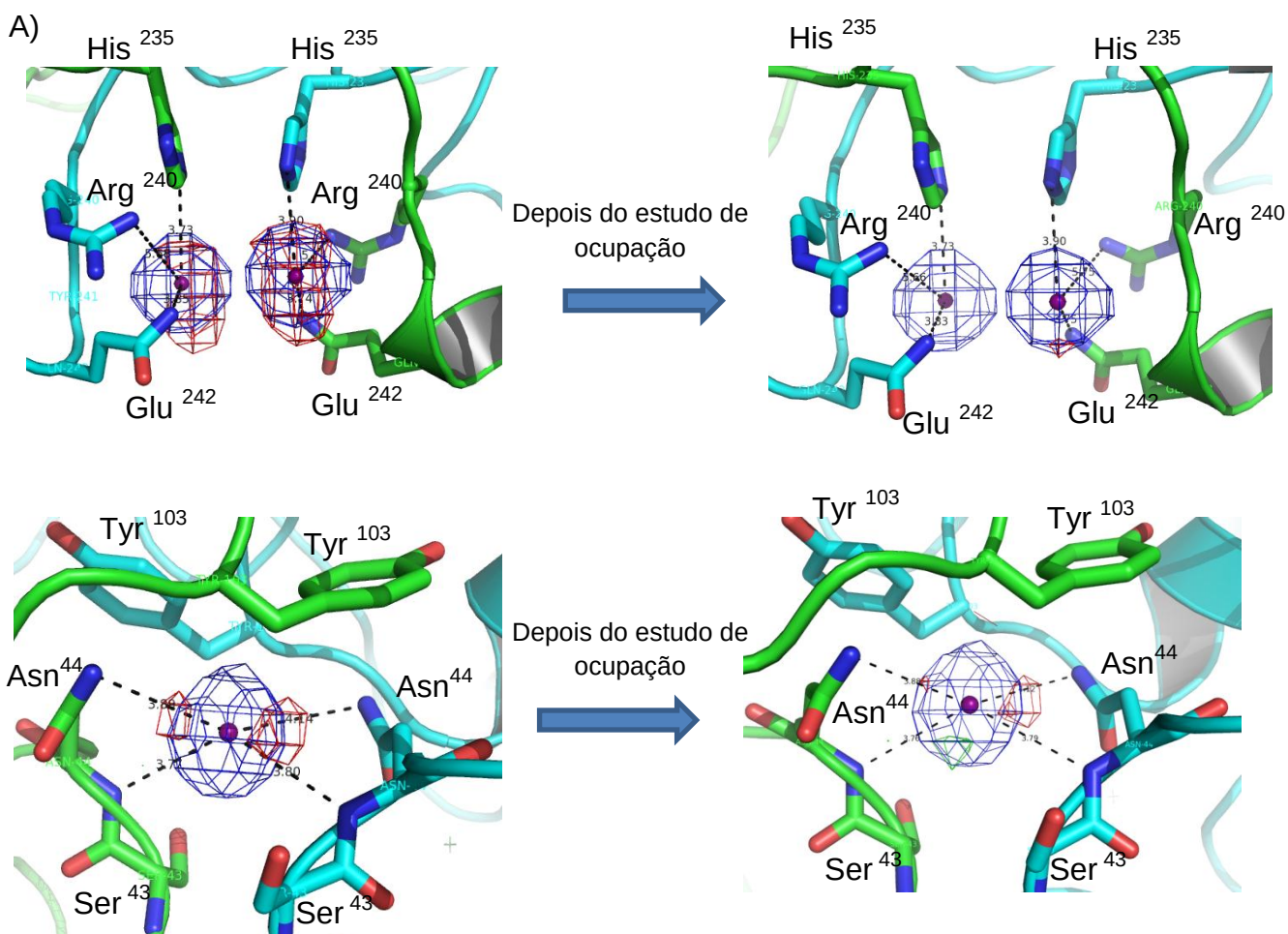
Tabela 11: Valores de ocupações e B_{factor} médios dos resíduos e águas ao redor dos íons Mn^{2+} e PO_4^{3-} , para cada estrutura. A estrutura XFG3A3oriC4cen apresenta os valores das águas com ocupações refinadas alternadas com o PO_4^{3-} .

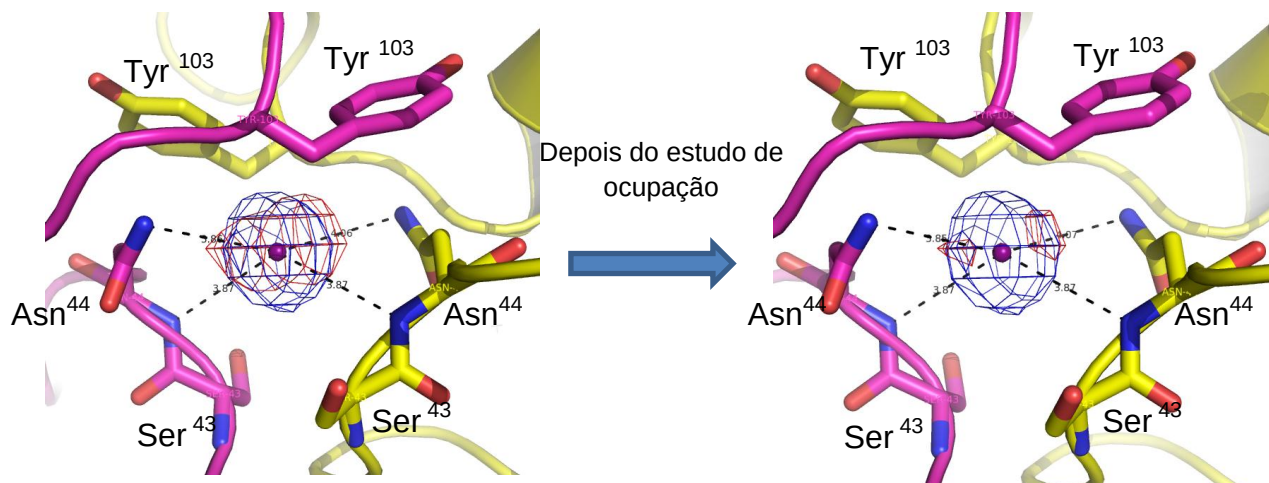
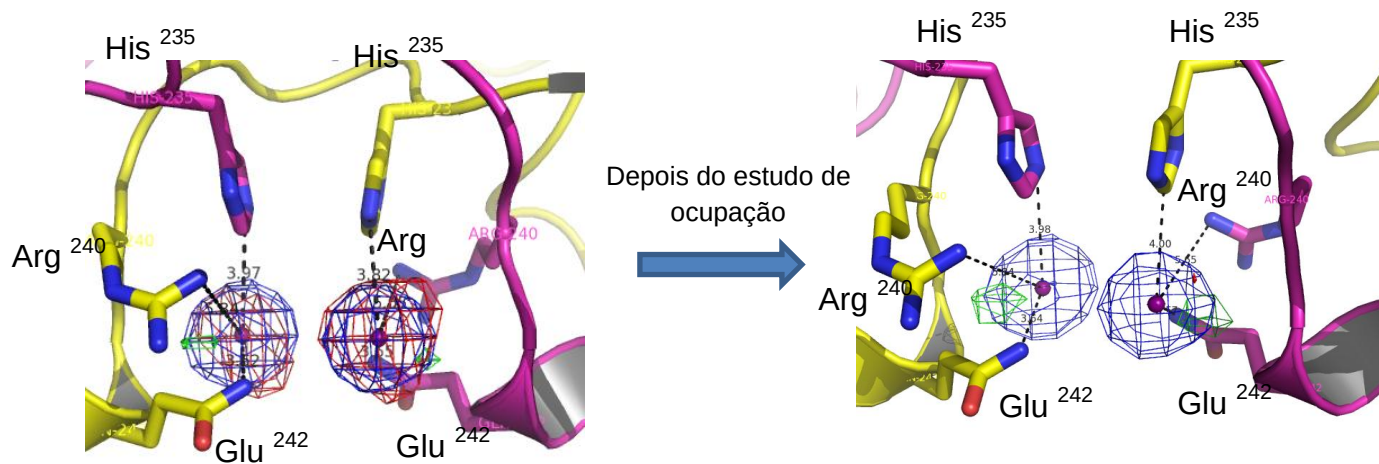
XFG3A3oriC4cen						XFG3A3ampD5_I					
Espécie	Ocupação (%)	B_{factor} médio íon/água	B_{factor} médio			Espécie	Ocupação (%)	B_{factor} médio íon	B_{factor} médio		
			Água	Proteína					Água	Proteína	
Mn^{2+}	1	100	44,07	33,24	33,04	Mn^{2+}	1	100	63,83	52,10	45,44
	2	100	44,29	30,60	33,49		2	100	54,45	47,97	49,71
	3	100	35,46	42,59	31,59						
	4	100	38,55	30,69	35,13						
PO_4^{3-}	1	39	16,80	32,99	23,90	PO_4^{3-}	1	100	47,64	46,85	37,27
	2	30	16,57	33,38	24,76		2	100	75,59	51,76	40,36
	3	43	17,64	52,59	26,16						
	4	31	17,47	46,78	26,83						
H_2O	1	61	9,80								
	2	70	8,59								
	3	57	13,41								
	4	69	11,31								

5.1.5.2 Íons fora do sítio ativo

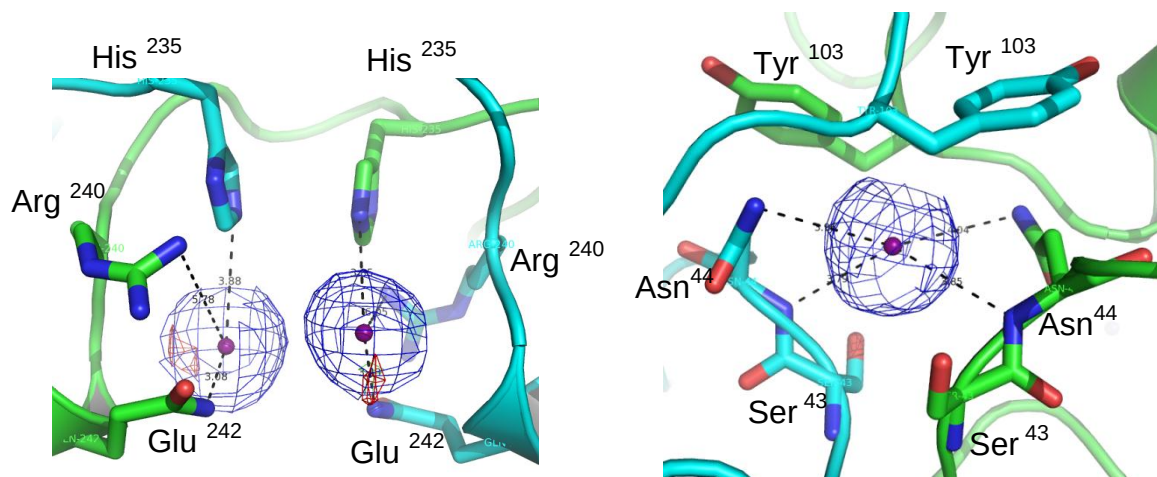
Durante o refinamento observaram-se outros seis sítios de interações para íons com carga negativa. Realizou-se teste de possíveis íons que durante a expressão, ou até mesmo na cristalização, pudessem ter se alocado na estrutura, a saber, Cl^- e I^- . Tentativas de refinar Cl^- sistematicamente levavam a densidade positiva em mapas de Fourier diferença, mesmo na ocupação de 100 %. Por isso, foi necessário realizar testes de ocupações para todos eles. A Figura 18 mostra a interação do I^- antes e após o estudo de ocupação.

Figura 18: A) Sítios de interações para I^- antes e após o estudo de ocupações para a estrutura XFG3A3oriC4cen. B) Sítios de interação para I^- com ocupação de 100 % para a estrutura XFG3A3ampD5_I. Em azul densidade eletrônica a 1 σ , em verde, nível de contorno 3 σ para o mapa Fourier diferença. Em vermelho, nível de contorno -3 σ para o mapa de Fourier diferença.





B)



A Tabela 12 mostra os B_{factor} para todos os I^- e as ocupações finais refinadas, estas orientadas baseando-se nos B_{factor} médios dos resíduos e águas ao redor de cada íon.

Tabela 12: B_{factor} dos I^- s, B_{factor} médios dos resíduos e das águas ao redor do I^- s.

Estruturas	I^-	Ocupação (%)	B_{factor}	B_{factor} médio	
				Água	Proteína
XFG3A3oriC4cen	1	74	39,5	40,3	34,38
	2	69	39,56	36,14	36,36
	3	82	26,15	39,99	27,02
	4	75	35,24	44,73	31,25
	5	90	36,36	39,89	32,59
	6	82	18,78	23,71	19,82
XFG3A3ampD5_I	1	100	48,19	45,3	38,46
	2	100	51,63	46,57	46,48
	3	100	72,28	46,50	46,12

Para reforçar a presença desse íon, cálculos de mapas anômalos foram realizados, cujas maiores densidades são nesses próprios sítios, como mostra a Tabela 13.

Tabela 13: Comparação entre os picos dos mapas anômalos com as posições refinadas dos íons I⁻ referentes as estrutura A) XFG3A3oriC4cen e B) XFG3A3ampD5_I.

A)

I ⁻	Posição dos picos, estrutura XFG3A3oriC4cen (x,y,z)	Altura relativa	Posição simetricamente equivalente dos I ⁻ refinados (x,y,z)	Ocupação (%)
1	64,080 / 61,863 / 69,813	22,80	-12,043 / 19,816 / -16,870	74
2	97,670 / 58,445 / 19,980	19,77	-11,705 / 16,342 / -19,967	69
3	77,979 / 19,789 / 57,729	29,93	-84,331 / 19,860 / -28,963	82
4	-1,251 / 58,663 / 25,492	24,07	-84,813 / 16,588 / -25,444	75
5	34,926 / 18,091 / 20,971	58,41	-34,860 / 19,151 / -20,951	90
6	100,599 / 17,999 / 62,621	37,83	-61,685 / 18,055 / -24,068	82

B)

I ⁻	Posição dos picos, estrutura XFG3A3ampD5_I (x, y, z)	Altura relativa	Posição simetricamente equivalente dos I ⁻ refinados (x, y, z)	Ocupação (%)
1	50,774 / 26,664 / 12,677	24,00	4,739 / -15,897 / 12,680	100
2	29,284 / 29,265 / 41,148	17,76	10,642 / -13,237 / 35,061	100
3	59,631 / 26,047 / 34,360	10,49	13,864 / -16,662 / 34,021	100

5.1.6 Comparações Estruturais

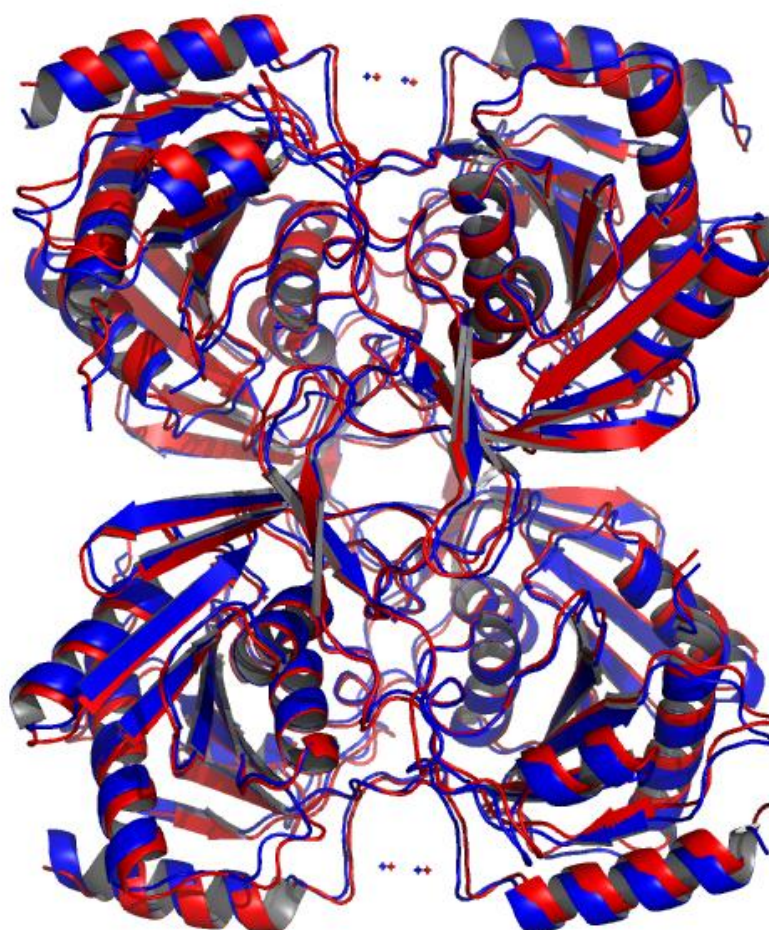
5.1.6.1 Comparações entre as estruturas refinadas

Realizou-se uma superposição entre os prováveis tetrâmeros biológicos das duas estruturas refinadas, como mostra a Figura 19. Observa-se, todavia, que quase não há diferenças significativas em relação à cadeia principal. Mesmo quanto às cadeias laterais, há poucas diferenças. Efetivamente, o rmsd (desvio médio

quadrático) global da superposição entre os tetrâmeros foi de 1,065 Å entre os carbonos α . O Gráfico 1 mostra o rmsd comparado para cada resíduo de ambas as estruturas. Em relação a cadeia lateral podemos listar resíduos que apresentam pelo menos um ângulo chi diferente de 120° ou mais, comparando-se entre as cadeias correspondentes para cada estrutura, que praticamente são os mesmos resíduos: Arg²; Asp⁸; Lys¹⁷; Met³¹; Arg³⁷; Arg³⁹; Ser⁴⁰; Ser⁴⁵; Arg⁵³; Lys⁵⁵; His⁶⁰; Asn⁸¹; Arg¹¹⁴; His¹⁴³; Asp¹⁶⁰; Thr¹⁶¹; Gln¹⁷⁴; Arg¹⁷⁶; Arg¹⁸⁶; His¹⁸⁷; Arg¹⁸⁸; Arg¹⁹⁹; Glu²¹²; Arg²²⁵, alguns desses resíduos estão em destaques na superposição de um monômero mostrada na Figura 19B.

Figura 19: A) Superposição entre a estrutura XFG3A3oriC4cen(em vermelho) e XFG3A3ampD5_I (em azul); B) superposição de apenas um monômero mostrando cadeia lateral de alguns resíduos listados com algum ângulo chi diferente de 120°.

A)



B)

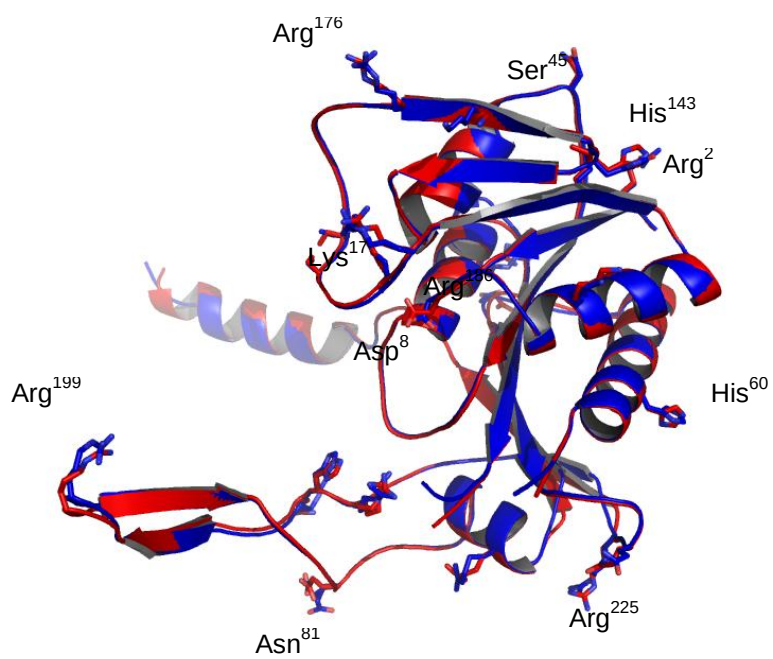
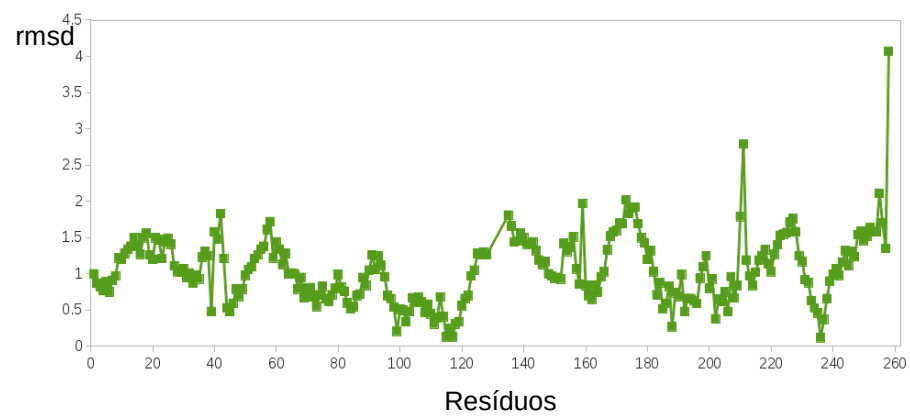
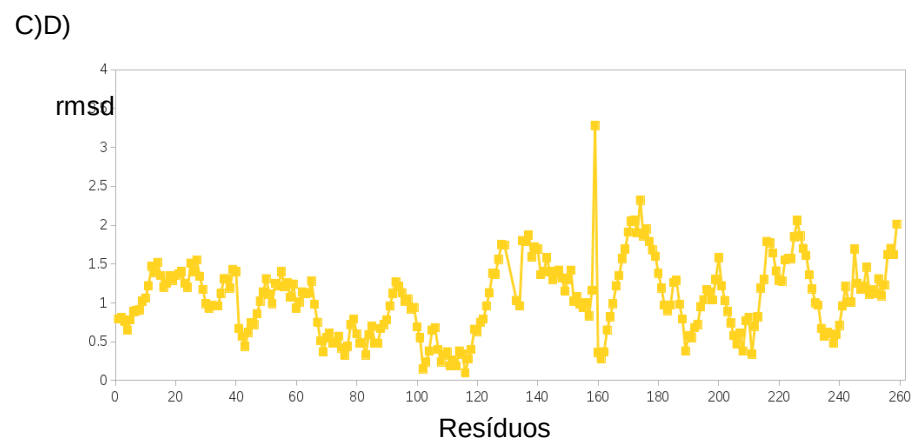
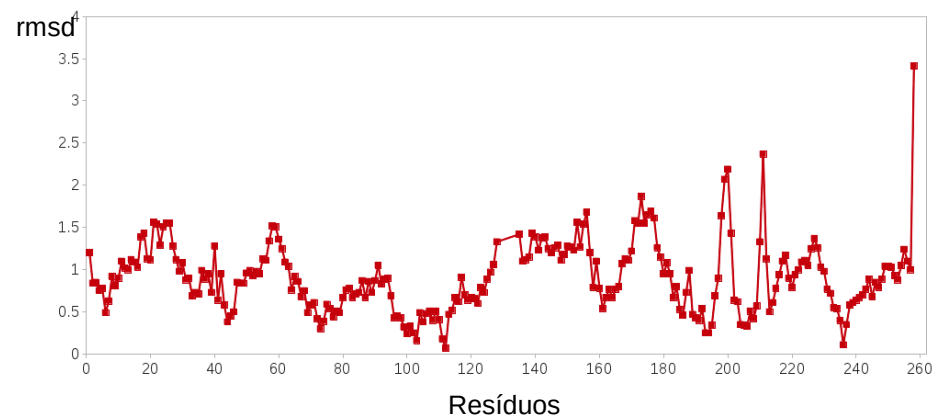
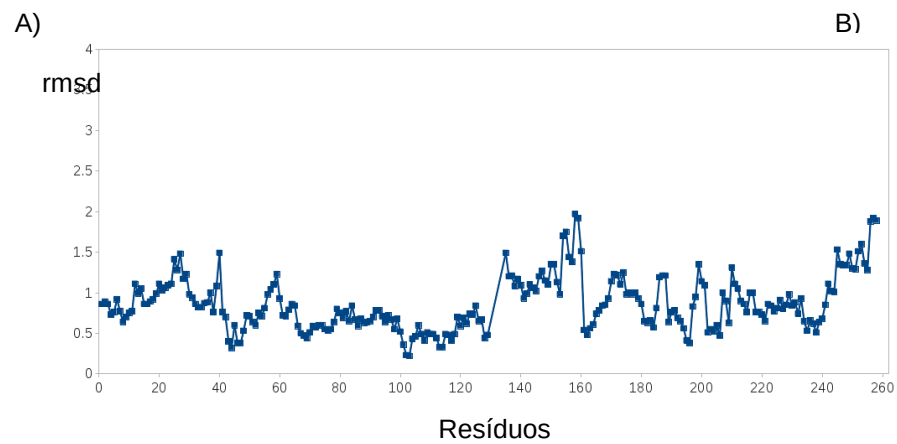


Gráfico1: Comparação entre as cadeias para cada resíduo de ambas as estruturas refinadas; A) cadeia A; B) cadeia B; C) cadeia C; D) cadeia D

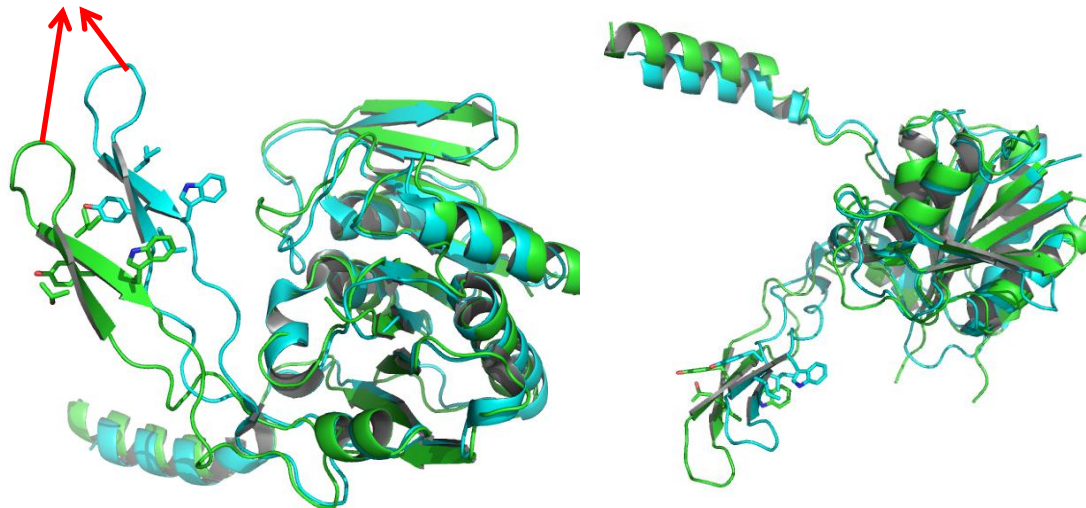


5.1.6.2 Comparação da estrutura XfSurE com a SurE de *Coxiella burnetti* (CbSurE)

A estrutura refinada foi comparada com a estrutura CbSurE (como descrito anteriormente, estrutura disponível com maior identidade, código PDB 3TY2). A Figura 20 mostra a superposição entre um monômero das estruturas. O rmsd global é 2,323 Å, superposição foi baseada na posição do carbono α e podemos realçar alguns resíduos mais distantes entre as estruturas XfSurE e CbSurE, respectivamente: Asn⁴⁴ e Ala⁵², Leu¹²⁸ e Glu¹³⁸, Ile¹⁹⁸ e Thr¹⁹⁴, Ile²⁰⁸ e Ile²⁰³, com destaque para os que fazem parte das alças de tetramerização, Tyr²⁰⁹ e Tyr²⁰⁴, Trp²¹⁰ e Trp²⁰⁵ entre outros.

Figura 20: Superposição de um monômero da XfSurE (em verde) com a CbSurE (em azul), evidenciando as diferenças na cadeia principal, um destaque maior nas alças de tetramerização que estão deslocadas uma em relação à outra.

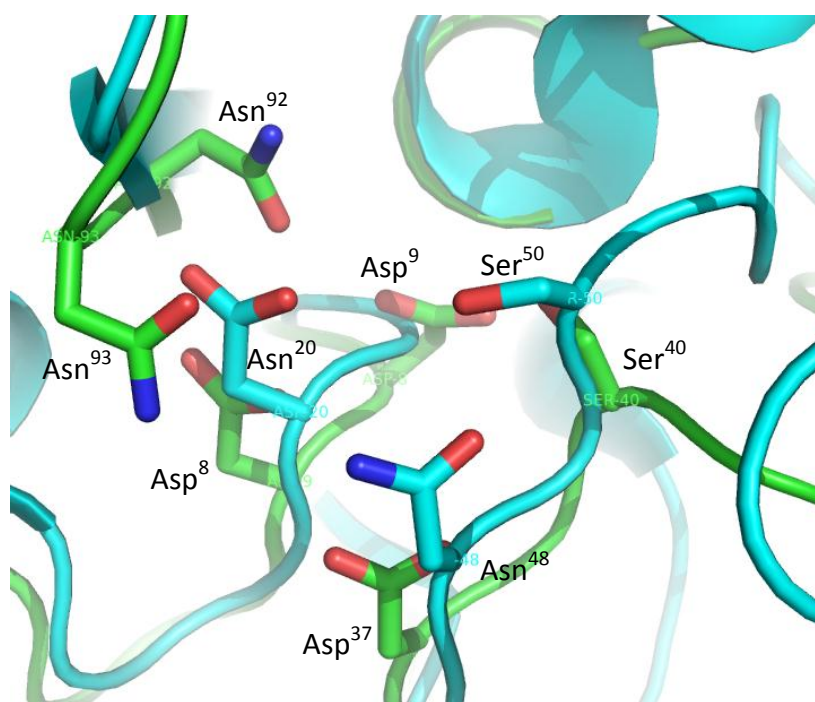
Alças de tetramerização



Destacamos as alças de tetramerização pelo fato de serem importantes para a oligomerização da proteína. Na CbSurE, a alça está mais fechada em relação à XfSurE. Particularmente, nos resíduos número 186 e 213 de XfSurE correspondentes aos resíduos 190 e 218 de CbSurE observa-se a significativa mudança de direção da cadeia.

Dadas as evidências de comportamento alostérico para a proteína XfSurE como descrito anteriormente, especula-se importância para esta propriedade o fato do sítio ativo do sítio ativo estar nas interfaces. Estudos mostram que as SurEs podem se apresentar tanto em conformação aberta como em conformação fechada, sendo essa última impeditiva para a interação com o substrato. Com o software CASTp (DUNDAS *et al.*, 2006), calcularam-se o volume e a área correspondentes ao sítio ativo de apenas um monômero para XfSurE e CbSurE, que resultaram, respectivamente, $V = 88,5 \text{ mm}^3$, $A = 106,4 \text{ \AA}^2$ e $V = 25,6 \text{ mm}^3$, $A = 35 \text{ \AA}^2$ (valores aproximados), o que indica que quando a conformação é aberta o sítio ativo está mais amplo. A Figura 21 mostra a comparação entre os dois sítios.

Figura 21: Superposição entre os sítios ativos, da XfSurE representado em verde e da CbSurE representado em azul.



6 CONCLUSÕES E PERSPECTIVAS FUTURAS

Concluimos que apesar das dificuldades nos sete conjuntos de imagens de difração de XfSurE, com cuidadosa escolha de parâmetros todos puderam ser processados. Todavia, os mapas de densidade eletrônica apresentam muito ruído; não obstante, duas estruturas foram cuidadosamente modeladas. Análise e comparação estruturais foram realizadas e aguarda-se no momento a adição de dados de outros estudos estruturais, como por SAXS, que somados aumentarão o entendimento do funcionamento desta proteína, inclusive nos tocante aos efeitos alostéricos.

7 REFERÊNCIAS

AFONINE, P.V.; Grosse-Kunstleve R.W.; Adams P.D. The Phenix refinement framework. **CCP4 Newsletter on Protein Crystallography**, v. 42, 2005.

ALBERTS, B. JOHNSON, Alexander, LEWIS, J. RAFF. M. ROBERTS, K. WALTER. P. *Biologia celular da célula*, 4ª ed. Artmed, 2006.

ANTONYUK, S. V.; ELLIS, M. J.; STRANGE, R. W.; BESSHO, Y.; KURAMITSU, S.; SHINKAI, A.; YOKOYAMA, S.; HASNAIN, S. S. Structure of SurE protein from *Aquifexaerobicus* VF5 at 1.5 Å resolution. **Acta Crystallographica Section F**. F.65p. 1204–1208. 2009.

Árvores contaminadas pela bactéria *Xylella fastidiosa* apresentam os sintomas da CVC em seus frutos e folhas. Disponível em: www.cnpmf.embrapa.br/publicacoes/produto_em_foco/citros_34.pdf. Acessado em 14/01/2013.

AYDIN, S. A comparison of ghrelin, glucose, alpha-amylase and protein levels in saliva from diabetics. **Journal of Biochemistry and Molecular Biology**, v.40, n. 1, p. 29-35, jan 2007.

BOND, C. S.; WOLFGANG. A.; SCHUTTELKOPF.; ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. **Acta Crystallographica Section D**. V. 65. P. 510-512. 2009.

BLOW, D. **Outline of Crystallography for Biologists**. 3 ed. Oxford University Press, 2005.

CLUSTALW2- Multiple Sequencia Alignment. Disponível em: <http://www.ebi.ac.uk/tools/clustaw2/>.

CHAMPMAN, S. M. SOMASUNDARAM, T. De-ice: recovery of diffraction intensities in the presence of ice rings. **Acta Crystallographica Section D**. D66, p- 741-744, 2010.

CHEN, V. B. BRYAN, W. A. HEADD, D. A. KEEDY, M. A. IMMORMINO, M. R. KAPRAL, G. J. MURRAY, J. S. RICHARDSON, D. C. Molprobity: all atom structure validation for macromolecular crystallography. **Acta Crystallographica Section D**. D. 66. P-12-21. 2010

CHYZASZCZ, T.; JANICKI, J. Inactivating substances in saliva by plant paralyzers and the presence of inactivating substances in solution of animal amylases. **Biochem.J.**, v.28, p.296-304, 1934.

CHOKSHI, D. Subchronic oral toxicity of standardized white kidney bean (*Phaseolus vulgaris*) extract in rats. **Food Chem. Toxicol.**, v. 45, n. 1, p. 32-40, jan. 2007.

CCP4 (Collaborative Computacional Project N. 4): Software for macromolecular X-ray crystallography. Disponível em: <http://www.ccp4.ac.uk/>.

CONTE, L. L.; BRENNER, E. S.; HUBBARD, P. J. T.; CHOTHIA, C.; MURZIN, G. A.; Scop database in 2002: refinements accommodate structural genomics. *Nucleic Acids Research*. V. 30., p. 264-267, 2002.

CULLITY, B. D. **Elements of Ray- X Diffraction**. 2 ed. Addison-wesley Publishing Company Inc. 1978.

DELANO, W. L. The Pymol Molecular Graphics System. San Carlos, CA: **Delano Scientific**, 2002.

Diagrama de fases para a água. Disponível em: www.lsbu.ac.uk/water/phase.html, acessado em 21/04/2013.

DRENTH, J. **Principles of Proteins X- Ray Crystallography**. Springer-Verlang. New York. USA. 1994.

DOWELL, L. G. RINFRET, A. P. Low- temperature forms of ice as studied by x-ray diffraction. **Nature**, v. 188, p- 1144-1148. 1960.

DUNDAS, J.; OUVANG, Z.; TSENG, J.; BINKOWSKI, A.; TURPAZ, Y.; LIANG, J.; CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. **Nucl.Acids Res**. V. 34, p. 116-118, 2006.

ELSPETH, F. G. SCHNEIDER, T. R. Macromolecular Crystallography, **Journal of Applied crystallography**, v. 30.p-211-237. 1997.

EMSLEY, P.; COWTAN, K. Coot: Model-Building Tools for Molecular Graphics. **Acta Crystallographica Section D**, p. 2126-2132, 2004.

FILHO, O. A. S.; ALENCASTRO, R. B. Modelagem de proteínas por homologia. **Química Nova**. V. 26, n. 2, p. 253-259, 2003.

GALPERIN. M. Y. MOROZ. I. V. WILSON. K. S. MURZIN. A. G. House cleaning a part of good housekeeping. **Molecular Microbiology**, v. 59, 5-19, 2006.

GONÇALVES, A. M. D.; REAGO, A. T. THOMAZ. M. ENQUITA. F. J. CARRONDO. M. A. Expression, purification, crystallization and preliminar X-ray characterization of

two Crystal forms of stationary-phase survival E protein from *Campylobacter jejuni*. **Acta Crystallographica Section F-Structural Biology and Crystallization Communications**, v. 64, 213-216, 2008.

GERRARD, J. A.; Prince, M. J.; ABELL, A. D. Kinetic characterization of enediol-based inhibitors of α -amylase. **Bioorganic e Medicinal Chemistry Letters**, v. 10, p.1575-1576, 2000.

HOOFF, R. W. W.; SANDER, C.; VRIEND, G. What_check: Errors in protein structures. **Nature**, v.381, p.272, 1996.

IWASAKI, W. MIKI, K. Crystal structure of the stationary phase survival protein SurE with metal ion and AMP. **Journal of Molecular Biology**. V 371, 123-136, 2007.

KABSCH, W. XDS: Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. **Journal of Applied Crystallography**, v. 26, p.795-800, 1994.

KANDRA, L.; ZAJACZ, J.; REMENYIK, J.; GYEMANT, G. Kinetic investigation of a new inhibitor for human salivary α -amylase. **Biochemical and Biophysical Research Communications**, v. 334, p. 824-828, 2005.

KIM, U, C. CHEN, Y-F. TATE, W, M. GRUNER, M, S. Pressure-induced high-density amorphous ice in protein crystals. **Journal of Applied Crystallography**. V-41.p-1-7. 2008.

KLEYWEGT, G. J.; JONES, T. A. xdi MAPMAM and xdiDATAMAN – programs for reformatting, analysis and manipulation of biomacromolecular electron-density maps and reflection data sets. **Acta Crystallographica Section D**, V.52.p-826-828, 1996.

KOHL, I. MAYER, E. HALLBRUCKER, A. The glassy water –cubic ice system: a comparative study by x-ray diffraction and differential scanning calorimetry, **Phys. Chem.** V. 2. p-1579-1586, 2000.

KRISSINEL, E.; HENRICK, K.; Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions, 2004.

KRISSINEL, E. HENRICK, K. Inference of macromolecular assemblies from crystalline state. **J. Mol. Biol.** V. **372**, 774—797. 2007

KUNZ. B. A. KOHALMI. S. E. KUNKEL. T. A. MATHEWS. C. K.; MCINTOSH. E. M. REIDV. J. A. Deoxyribonucleoside triphosphate levels – a critical factor in the maintenance of genetic stability. **Mutation Research- Reviews in Genetic Toxicology**. V 318, 1-64, 1994.

LAEMMLI, U. K., Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature**. 227(259): 680-5, 1970.

LAMBAIS. M. R.; GOLDMAN. M. H. S.; CAMARGO. L. E.; GOLDMAN. G. H. A genomic approach to the understanding of *Xylella fastidiosa* pathogenicity. **Current Opinion in Microbiology**. V 3, 459-462, 2000.

LASKOWSKI, R A. PDBsum new things. **Nucleic Acids Res.**, V. 37, 2002

LAYER, P.; RIZZA, R. A.; ZINSMEISTER, A. R.; CARLSON, G. L.; DIMAGNO, E. P. Effect of a purified amylase inhibitor on carbohydrate tolerance in normal subjects and patients with diabetes mellitus. **Mayo Clin.Proc.**, v. 61, p.447, 1986.

LEE. J. Y.; KWAK. J. E.; MOON. J. EOM. S, H. LIONA. E. C. PEDELACA. J. D. BERENDZEN. J. SUH. S. W. Crystal structure and functional analysis of the SurE protein identify a novel phosphatase family. **Nature Structural Biology**. V8. 789-794, 2001.

LESLIE, A. G. W. Mosflm. **ActaCryst. D** 62, 48–57. 2006

LI, C.; ICHIKAWA. J. K; RAVETTO. J. J. KUO. H. C.; FU. J. C; CLARKE. A new gene involved in stationary-phase survival located at 59 minutes on the *Escherichia coli* chromosome. **Journal of Bacteriology**. V 176, 6015- 6022, 1994

MARSHALL, J. J.; LAUDA, C. M. Purification and properties of phaseolamin, an inhibitor of α amylase, from the kidney bean, *Phaseolus vulgaris*. **J. Biol. Chem.**, v. 250, p. 8030-8037, 1975

MEIDANIS, J.; BRAGA, M. D. V.; ALMEIDA, S. V. Whole-Genome Analysis of Transporters in *Xylella fastidiosa* the Plant Pathogen. **Microbiol.Mol. Biol.** v. 66, p. 272-299, 2002.

MCCOY, A. J.; GROSSE-KUNSTLEVE, R. W.; ADAMS, P. D.; WINN, M. D.; STORONI, L. C.; READ, R. J. Phaser crystallographic software. **Journal of Applied Crystallography**, v. 40, p. 658-674, 2007.

Microscopia eletrônica da *Xylella fastidiosa* nos vasos de xilema em laranjeiras. Disponível em: <http://aeg.lbi.ic.unicamp.br/xf/home/mmachado>. Acesso em 23/11/2012.

MURA. C.; KATZ. J.E. CLARCKE. S. G. EISENBERG. D. Structure and function of an archaeal homolog of survival protein E (SurE α): and Acid phosphatase which purine nucleotide specificity. **Journal of Molecular Biology**, v. 326, 1559, 2003

MURAI, A.; IWAMURA, K.; TAKADA, M.; OGAWA, K.; USUI, T.; OKUMURA, J. Control of postprandial hyperglycemia by galactosylmaltobionolactone and its novel anti-amylase effect in mice. **Life Sciences**, v. 71, p. 1405-1415, 2002.

National Center for Biotechnology Information (NCBI), disponível em: www.ncbi.nlm.nih.gov.

ODA, Y. Tertiary and quaternary structures of 0,19 alpha-amylase inhibitor from wheat kernel determined by X-ray analysis at 2,06 Å resolution. **Biochemistry**, v. 36, n, 44, p. 13503-13511, 1997.

PAINTER, J. MERRIT, E. A. TLMSD web server for the generation of multi-group TLS models. **Journal of Applied Crystallography**. V. 39, p-109-111. 2006.

PAPPACHAN.A.; SAVITHRI. H, S. MURTHY. M. R. N. Structural and functional studies on a mesophilic stationary phase survival protein (SurE) from *Salmonella typhimurium*. **Febs Journal**, v. 275.5855-5864, 2008.

PARK, Y. SHEETLIN, S. MA, N. MADDEN, L, T. SPOUGE, L, J. New finite-size correction for local alignment score distributions, **BMC Research notes**. V.5 p-286. 2012.

PAYAN, F. Structural basis for the inhibition of mammalian and insect α -amylases by plant protein inhibitors. **Biochemical and Biophysical Acta**, v. 1696, p. 171-180, 2004

POOLER, M, R. HARTUNG. Specific PCR detection and identification of *Xylella fastidiosa* strains causing citrus variegated chlorosis. **Cur Microbiol**. V. 31. p-377-381. 1995.

PRODFOOT.M. Kuznetsova. E. Brown. G. Rao. N. N Kitagawa. M .Mori. H. Savchenko. A, Yakunin. A. F. General enzymatic screens identify three new nucleotidase in *Escherichia coli* – Biochemical characterization of SurE. **Journal of Biological Chemistry** 279(52): 54687 – 54694.

Protein Data Bank (PDB). Disponível em: www.rcsb.org.

Quando os raios X atingem diferentes planos, são refletidos em diferentes ângulos sendo detectado por um filme; Imagem de um padrão de difração. Disponível em: www.azevedolab.dominiotemporario.com/doc./diffraction1.pdf. Acesso em 17 fev. 2012

REIS, M. A.; SARAIVA, A. M.; SANTOS, M. L.; SOUZA, A. P.; APARICIO, R.; Crystallization and preliminary X-ray analysis of stationary phase survival protein E (SurE) from *Xylella fastidiosa* in two crystal forms, 04/2012, **Acta Crystallographic. Section F**, Vol. 68, pp.464-467, Copenhagen, Alemanha, 2012 *

ROMÃO, M. J. Cristalografia de proteínas: metodologias e aplicações de proteínas. **Biologia estrutural**. N. 53, p. 18-36, 1996.

SILANO, Vitório. Alpha-amylaseinhibitors. In: KRUGER, James: STAUFFER, Clyde E. **Enzymes and their role in cereal technology**. St. Paul: American Association of Cereal Chemists, cap. 6 p. 141-199 1987.

SAMBROOK, J.;FRITSCH, E.F.; MANIATIS, T. Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press, **Molecular cloning: a laboratory manual** 1989.

SARAIVA, A. M.; REIS. A. M.; TADA. F. S. LUCIANA, K.; MURAI, R.; SCHNEIDER. S. R.; PELLOSO. C. A.; TOLEDO. S. A. M.; GILES. C.; APARICIO.R.; SOUZA.P. A.; Functional and small-angle X ray scattering studies of a new stationary phase survival protein E (SurE) from *Xylella fastidiosa* – evidence of allosteric behavior. **FEBS Journal**. V. 276, p. 6751-6762, 2009.

SIMPSON, A. J. G., et al., The genome sequence of the plant pathogen *Xylella Fastidiosa*. **Nature**, v. 406, p. 151-157, 2000.

SIVAKUMAR, S.; MOHAN, M.; FRANCO, O.L.; THAYUMANAVAN, B. Inhibition of insect pest α -amylases by little and Finger millet inhibitors.**Pesticide Biochemistry and Physiology**, v. 85, n.3, p. 155-160, 2006.

SPECTOR, S.; FLYNN, J. M.; TIDOR, B.; BAKER, T. A.; SAUER, R. T. Expression of N-formylated proteins in Escherichia coli. *Protein ExprPurif* 32, p. 317-322, 2003.

STEIN, N. CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. *J. Appl. Cryst.*, 41, 641-643, 2008.

STRYER, L.; TYMOCZKO, J. L.; BERG, J. M. **Bioquímica**, Ed: 5º, Edit: Guanabara, 2004.

TEIXEIRA, A. P. et al. O efeito da adição de yacon no suco de laranja industrializado sobre a curva glicêmica de estudantes universitários. **Alim. Nutr.**, Araraquara, v. 20, n. 2, p. 313-319, abr./jun. 2009

WANG, J. R.; PU, Z.; LAN, X.; BAUM, B. R.; YAN, Z.; ZHENG, Y.; WEI, Y. Phylogenetic analysis of dimeric alpha-amylase inhibitor sequences from an orthologous region in 21 different genomes of the tribe Triiceae (Poaceae).**Biochemical Systematics and Ecology**. V. 38, p. 708-714, 2010.

WANG, J. R.; YAN, Z. R.; WEI, Y. M; NEVO, E.; BAUM, B. R.; ZHENG, Y. L. Molecular characterization of dimeric alpha-amylase inhibitor genes in wheat and development of genome allele-specific primers for the genes located on chromosome 3BS and 3DS. **Journal of Cereal Science**, p. 1-9, 2006.

WESTBROOK, J. ValidationPDB- Validation of protein structures for protein data bank.**Methods in Enzymology**, 2003.

WHITFORD, D. **Proteins structure and function**. Copyright, 2005.

WOLEVER T. M.; RADMARD, R.; CHIASSON, J. L.; HUNT, J. A.; JOSSE, R. G.; PALMASON, C.; RODGER, N. W.; ROSS, S. A.; RYAN, E. A.; TAN, M. H. One year acarbose treatment raises fasting serum acetate in diabetic patients. **Diabetic Medicine**, v. 12, p. 164-172, 1995.

8 APÊNDICE

8.1 Teste de expressão dos inibidores de α -amilase presente no gênero *Triticum*

8.1.1 Inibidores de α -amilase do gênero *Triticum*

O gênero *Triticum* abrange uma variedade de espécies do trigo, sendo que a família das gramíneas da tribo *Triticeae* é a quarta maior família de plantas floríferas, com mais de 8.700 espécies classificadas em cerca de 650 gêneros. O trigo é das culturas alimentares mais importantes para a nutrição humana e animal, muitos estudos são realizados na busca e identificação de inibidores de α -amilase de natureza proteica. As primeiras investigações desses inibidores de α -amilase em cereais foram descritas para o trigo (CHERZASZCZ *et al.* 1934). Desde então, vários inibidores já tiveram sua estrutura tridimensional determinada, por exemplo, o inibidor 0.19 (ODA, Y *et al.* 1997) presente na espécie *Triticum aestivum*.

Os inibidores de α -amilase pertencem à classe de inibidores de enzimas digestivas e sua atividade consiste em inibir a atividade das amilases (MARSHAL, 1975). Esses inibidores estão amplamente distribuídos em diversos cereais, leguminosas, tubérculos, oleaginosas, enfim, em diversas plantas como cevada, centeio, milho, aveia, sorgo, milhetos e tritcale (SILANO, 1987), entre outros (SIVACUMAR *et al.* 2006; WANG *et al.* 2006). Nas albuminas, essa proteína compõe cerca de dois terços da sua composição (WANG *et al.* 2010).

8.1.1.1 Importância dos inibidores proteicos de α -amilase

O estudo de inibidores de α -amilases é importante devido às suas diversas potenciais aplicações: tratamento de *Diabetes mellitus*, diagnóstico de distúrbios no pâncreas, controle de obesidade, tratamento de cáries dentárias e como possíveis ferramentas para engenharia de resistência de plantas cultivadas, entre outros (PAYAN, 2004).

De acordo com a Organização Mundial da Saúde (OMS), mais de 180 milhões de pessoas em todo o mundo são portadoras de *Diabetes mellitus* com provável aumento para mais que o dobro em 2030. Segundo estimativa da OMS,

aproximadamente 1,6 bilhões de adultos no mundo estão acima do peso e pelo menos 400 milhões são obesos, números que poderão chegar a 2,3 bilhões e mais de 700 milhões, respectivamente, em 2015 (TEIXEIRA *etal.* 2009).

Dietas suplementadas por “starchblockers”, como são popularmente denominados os inibidores de α -amilase, têm sido adotadas de forma crescente. Acredita-se que o inibidor extraído do feijão branco diminua o conteúdo calórico fornecido pela ingestão de carboidratos por inibir a enzima α -amilase, responsável pela conversão dos carboidratos complexos (amido) em açúcares simples, e portanto, levando a menor absorção de açúcar no intestino delgado (CHOKSHI, 2007).

Diabetes mellitus é caracterizada por hiperglicemia e é uma doença complexa; ela diminui substancialmente a qualidade e a expectativa de vida, visto que está relacionada a complicações sérias. Estudos mostram que inibidores retardam a digestão e a absorção geral do amido e reduzem também os picos de glucose pós-prandial, induzindo a tolerância a carboidratos (AYDIN, 2007).

Outros estudos revelam contribuição desses inibidores no tratamento para obesidade. Eles diminuem os níveis pós-prandiais de insulina e aumentam a sensação de saciedade, por meio do retardamento do esvaziamento gástrico, além de reduzirem importantes fatores adipogênicos (LAYER *etal.* 1986; GERRARD *et al.* 2000; MURAI *etal.* 2002). Uma diferença importante em relação aos inibidores não proteicos seria é que forte inibição exercida por este provoca um acúmulo excessivo de polímeros de carboidratos de cadeia curta. Isso causa diarreia e outros efeitos colaterais (WOLEVER *etal.* 1995). Já os inibidores proteicos não provocam acúmulo excessivo de dissacarídeos, pois a microbiota do intestino humano é capaz de metabolizar grandes quantidades de amido não absorvidos presentes no lúmen intestinal, com isso a má-absorção de carboidrato induzida por esse inibidor não aumenta a carga osmótica, não causando nenhum efeito colateral. (LAYER *etal.* 1986).

Os inibidores podem auxiliar no tratamento de cáries dentárias. A α -amilase humana é capaz de se ligar ao esmalte dentário (hidroxiapatita), podendo levar à formação de placas dentárias, e também se liga com afinidade às bactérias (*Streptococcus*); ainda assim, ela conserva a sua ação perante o amido e forma

produtos que podem ser usados como fonte de alimento pelas bactérias, que são metabolizados a ácido láctico (KANDRA *etal.*2005).

Uma outra contribuição importante seria no desenvolvimento de transgênicos geneticamente resistentes, como por exemplo, a brocas, já que são capazes de expressar inibidores específicos para enzimas de predadores (SIVAKUMAR *et al.* 2006). Vários tipos de inibidores de α -amilase e proteinase em sementes e órgãos vegetativos agem para regular o número de insetos fitófagos (WANG *etal.*2010).

8.1.2 Materiais e reagentes:

Na eletroporação foram utilizados: cubetas de plástico com distância entre eletrodos de 0.1 cm (Bio-Rad), placas de Petri (Prolab), alça de Drigalsk, microtubos, pipeta de Pasteur e o eletroporador (Bio-Rad, modelo MicroPulser). Na expressão das proteínas foram utilizados: tubos de ensaios com tampa, Shaker (Nova Ética modelo 430-RDBP), cuba de eletroforese (Bio-Rad, modelo Mini Protean), espectrofotômetro (Varian, modelo Cary50).

Reagentes utilizados na preparação de meios de cultura segundo SAMBROOK *et al.* (1989): a) meio de cultura Luria Bertani LB, contendo Triptona 10 g L⁻¹, Extrato de levedura 5 g L⁻¹ (Biogras), NaCl (Cloreto de sódio) 10 g L⁻¹ (Merck P.A.), b) LA (meio LB contendo ágar 15 g L⁻¹) c) meio SOB (Triptona 2 g L⁻¹, Extrato de levedura 0,5 g L⁻¹, NaCl 0,05 g L⁻¹) e d) meio SOC (meio SOB contendo MgCl₂ 1 mL, C₆H₁₂O₆ (Glucose) 1 mol L⁻¹). Na expressão utilizaram-se IPTG (isotropil- β -D-tiogalactopiranosídeo) 1mmol L⁻¹ e o antibiótico Canamicina 50 μ g mL⁻¹ (Km⁵⁰). Na preparação do tampão de ressuspensão do *pellet* utilizaram-se H₃PO₄ (Ácido fosfórico) (Nuclear P.A.) e NaOH (Hidróxido de sódio) (Nuclear P.A.). Na eletroforese utilizaram-se C₄₅H₄₄N₃NaO₇S₂ (ComassieBriliante Blue R-250) (Isopar), HCl (Ácido clorídrico) (Merck P.A.), C₂H₄O₂ (Ácido acético glacial) (Synth P.A.), C₇H₁₀N₂O₂ (bis-acrilamida) (Sigma), NaC₁₂H₂₅SO₄ (Sulfato dodecilde sódio) (SDS) (Reagen), C₁₉H₁₀Br₄O₅S (Azul de bromotimol) (Vetec), C₃H₈O₃(Glicerol (Reatec P.A.) C₃H₅NO (Acrilamida) (Vetec), N₂H₈S₂O₈ (Persulfato de sódio) (Fisher Biotech), C₆H₁₆N₂ (Temed 1-2 bis-etano (dimetilamino)) (Sigma) e C₂H₅NO₂ (Glicina) (Vetec), C₆H₁₃NO₅ (Tricina) (Sigma).

8.1.3 Parte experimental

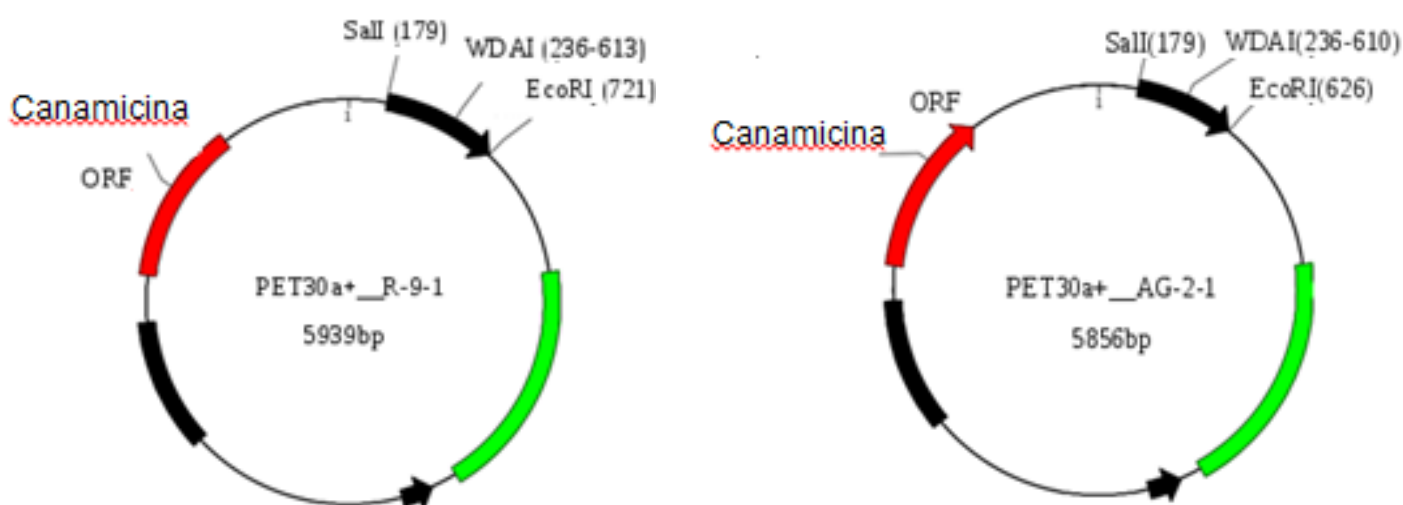
8.1.3.1 Introdução dos vetores de expressão em *E.coli* por eletroporação

Plasmídeos Pet30+ (Figura 22) com os genes R9 e Ag2 em microtubos foram enviados pelo Prof. Dr. Ji-Rui Wang da Universidade Agrícola de Sichuan, China. A esses microtubos foram adicionados 10 µl de água ultrapura autoclavada, desse volume foi retirado 1 µl que foi adicionado a outro microtubo contendo 40 µl de bactérias *Escherichia coli* (*E.coli*). As bactérias com os plasmídeos foram transferidas para uma cubeta e submetidas a um choque elétrico, após esse procedimento, imediatamente foi adicionado na cubeta 1mL de meio SOC em que se deixou por 55 min sob agitação de 250 RPM em Shaker a 37 °C, para a recuperação da membrana das bactérias.

Após o tempo de recuperação, realizou-se o plaqueamento em meio LA contendo Canamicina 50 µg/mL, utilizando-se nas placas de Petri diluições de 1:1, 1:10, 1:30, 1:300. As placas foram incubadas a 37 °C por aproximadamente 12 h.

O procedimento relatado foi realizado igualmente para os dois inibidores, em separado.

Figura 22: Plasmídeos Pet30+ com os genes dos inibidores de α -amilase R9 e Ag2



8.1.3.2 Teste de Expressão dos inibidores de α -amilase R9 e Ag2

Para o pré-inóculo foram utilizados 10 mL de meio de cultura LB contendo canamicina a 50 $\mu\text{g/mL}$ e escolheram-se as melhores colônias crescidas nas placas de Petri. Incubou-se por 16 h a 37 °C.

Em seguida, diluiu-se 100 μL do pré-inóculo em um volume de 10 mL de meio LB com canamicina 50 $\mu\text{g/mL}$. Incubou-se novamente até que a densidade ótica (D.O) atingisse entre 0,4 e 0,6 no comprimento de onda de 600 nm, para posteriormente se induzir a expressão adicionando-se IPTG com concentração de 1mmol L^{-1} por 4 h a 37 °C, sob agitação de 250 RPM. Após, a cultura foi centrifugada a 10.000 g por 5 min então o sobrenadante foi descartado. As bactérias foram ressuspensas em tampão fosfato de sódio 5mmol L^{-1} pH 6,9 e lisadas por choque térmico com nitrogênio líquido e banho Maria a 40 °C (10 vezes por aproximadamente 20 segundos). Retirou-se uma alíquota da fração total e realizou-se então uma nova centrifugação a 10.000 g por 5 min, coletou-se uma fração do sobrenadante (porção solúvel) e o ppt foi ressuspenso novamente em tampão fosfato de sódio 5mmol L^{-1} PH 6,9, do que coletou-se uma fração.

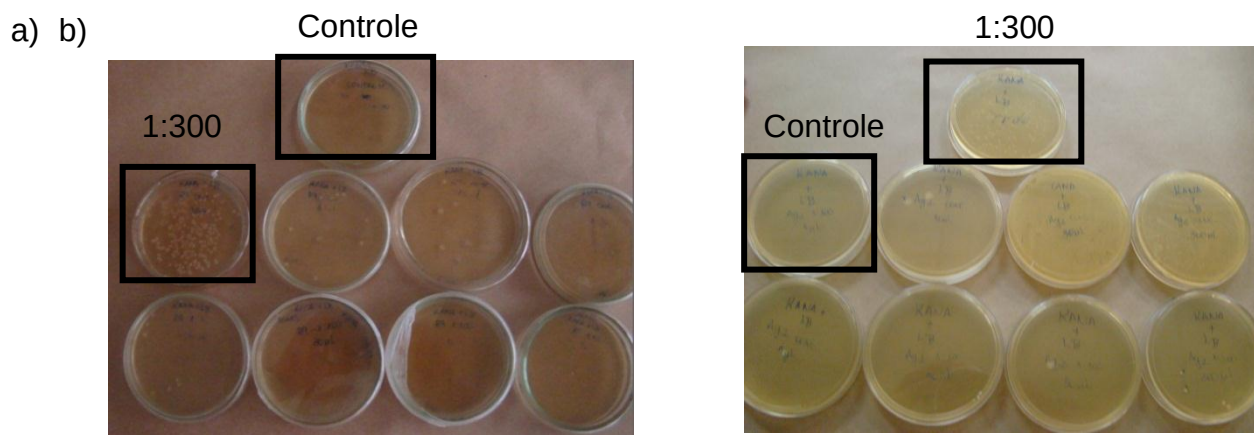
Para a análise da expressão dos inibidores, usou-se eletroforese pelo método de LAEMMLI (1970) com gel a 15% a uma tensão de 150 V durante aproximadamente 1 h.

8.1.4 Resultados e discussões

8.1.4.1 Introdução dos vetores de expressão em *E.coli* por eletroporação

A Figura 23 mostra o crescimento das colônias de bactérias transformadas para ambas as proteínas, observa-se que o melhor crescimento ocorreu na placa com diluição 1:300 μL , enquanto que na placa com bactérias não transformadas não houve crescimento.

Figura 23: Placas com colônias de bactérias transformadas: a) colônias de bactérias com o gene da proteína R9; b) colônias de bactérias com o gene da proteína Ag2 (onde a posição do controle está trocada com da diluição 1:300)



8.1.4.2 Teste de Expressão dos inibidores de α -amilase R9 e Ag2

No eletroforegrama (Figura 24), observa-se que não há presença de bandas de superexpressão, por algum motivo não foram obtidas a produção das proteínas, que devem ter massa molecular de aproximadamente 13 kDa, ou seja, revelar uma banda forte na parte inferior do gel. Algumas hipóteses levantadas para tal fato são: a transformação não tenha sido com uma alta eficiência, o rompimento da membrana da bactéria não ter ocorrido no momento da lise por choque térmico e as condições de expressão não terem sido favoráveis para a expressão da proteína.

Mediante o fato da oportunidade de trabalhar em colaboração com o Prof. Dr. Igor Polikarpov e o Prof. Dr. Ricardo Aparicio na parte computacional de estrutura de proteínas a partir de um número grande de conjuntos de dados, o que levou ao aprendizado desta parte (computacional) em soma ao já realizado em termos experimentais de bancada, preferiu-se enfatizá-lo no trabalho.

Figura 24: Eletroforegrama SDS-PAGE para teste de expressão dos inibidores R9 e Ag2: linha 1: Extrato bruto sem indução; linha 2 extrato bruto com indução, linha 3 sobrenadante corresponde à proteína solúvel. Linha 4 precipitado corresponde às proteínas insolúveis; linha 5: marcadores de massa molar (MW): 116kDa β -galactosidase, 66kDa albumina soro bovina, 45kDa ovoalbumina 35kDa lactato desidrogenase 25kDa REaseBsp 98, I 18kDa β -lactoglobulina, 14kDa lisozima, respectivamente para cada gel. a) proteína Ag2; b) proteína R9.

