



UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE QUÍMICA DE SÃO CARLOS
LABORATÓRIO DE CROMATOGRAFIA

**“CLONAGEM, EXPRESSÃO E CARACTERIZAÇÃO DE
PROTEÍNAS RECOMBINANTES DE *XYLELLA FASTIDIOSA*”**

Célia Sulzbacher Caruso

Tese apresentada ao Instituto de
Química de São Carlos, da
Universidade de São Paulo, para
obtenção do título de Doutor em
Ciências, área Química Analítica.

Orientador: Prof. Dr. Emanuel Carrilho

São Carlos
2007

**Quando a saudade bater
No quarto escuro do peito
Você não pode negar
O que ficou bem marcado
O que ficou definido, neste lugar.
Lugar que nunca se esquece
Lugar que nunca envelhece
Lugar que tem seu lugar
Quando a saudade vier
Diga que fui pra lá lhe ver
Diga que fui buscar
Um pedaço do tempo que ficou
Como se fosse dois meninos
Se lambuzando de mel
Correndo pra festejar o amor, o amor, o AMOR....**

**As únicas coisas que importam são
as feitas de verdade e alegria”**

Célia Caruso

Agradecimentos

- A Deus, pela presença constante em minha vida;
- Ao Prof. Dr. Emanuel Carrilho, pela amizade, confiança, orientação e dedicação de sempre. Agradeço de todo o coração!
- À Prof. Dra. Ana Paula Ulian de Araújo, que muito fez por mim, meus sinceros agradecimentos! Que eu possa retribuir incondicionalmente a confiança, o respeito, a preocupação e ajuda de sempre.
- Á Profa. Dra. Eliana G. M. Lemos e a todo seu grupo de pesquisa que muito fizeram para o início deste trabalho, pelas horas de descontração, amizade, respeito, confiança e, acima de tudo, incentivo.
- Aos técnicos da Biofísica IFSC-USP: Bel, Andressa, e João, por toda a ajuda, paciência e principalmente suporte durante todo o período de realização deste trabalho.
- Aos professores do grupo de Biofísica Molecular “Sergio Mascarenhas”: Jabáh, Nelma, Leila e Otaciro por terem me recebido como membro do grupo.
- Aos amigos da Biofísica: Cris, Leandro, Fernando, Ana Paula, Priscila, Ceará, Fernanda, Wânius, Lia, Julio, Luis, Daniel e Militar, pela amizade!
- Aos meus grandes amigos Sheila, Júlio, Assuero, Joci, Débora, Fernando, Roberta e José Luiz.....meus amigos de todas as horas....tão especiais.... que só tenho a agradecer a DEUS por ter convivido com vcs....tudo foi muito bom!!! Eu amo vcs!!!

- Aos meus pais, José e Petronila, pelo amor de sempre, pela paciência e por serem aqueles que me trouxeram até aqui. Aos meus irmãos, Jeremias, Juliana e Cecília, pela mais sincera amizade. Eu amo vcs!

- A CAPES e a FAPESP pelo suporte financeiro para a realização deste trabalho.

SUMÁRIO

Dedicatória

Agradecimentos

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SÍMBOLOS

RESUMO

ABSTRACT

<u>I – INTRODUÇÃO</u>	1
I.1 - Reação em Cadeia da Polimerase – PCR	2
I.2 - Clonagem Molecular.....	4
I.2.1 - Conceito de Clonagem Molecular	4
I.2.2 - Enzimas de Restrição	5
I.2.3 - Vetores de Clonagem	10
I.2.4 - Construção do DNA Recombinante	12
I.2.4.1 - Enzimas DNA Ligases	13
I.2.5 - Transformação Bacteriana	15
I. 2.5.1 – Conceito de Transformação Induzida.....	15
I. 2.5.2 – Mecanismos de Captação do DNA.....	15
I. 2.6 – Detecção e análise dos Clones Recombinantes.....	17
I. 2.6.1 – Seqüenciamento de DNA	18
I. 2.6.1.1 – Seqüenciamento Manual.....	19
I. 2.6.1.2 – Seqüenciamento Automático	21
I. 3 – Expressão de Proteínas Heterólogas.....	22
I. 4 – Bactéria <i>Xylella fastidiosa</i>	25
I. 4.1 - Proteínas de <i>Xylella fastidiosa</i>	27
I. 4.1.1 - Hidroxinitrila Liase.....	27
I. 4.1.1.1 - Enzimas dependentes de FAD....	29
I. 4.1.1.2 - Enzimas independentes de FAD	30
I. 4.1.2 – Corismato Sintase.....	32
I. 4.1.2.1 – Biosíntese de Aminoácidos Aromáticos...35	
I. 4.1.3 – Proteína D do Sistema de Transporte e Secreção Tat	36
I. 5 – Objetivos e Perspectivas	40

II – MATERIAIS E MÉTODOS	43
II.1 - MATERIAIS	43
II.1.1 - Reagentes, Enzimas, Linhagens Bacterianas e Vetores	43
II.2 – MÉTODOS	44
II. 2.1 – Extração do DNA genômico	44
II. 2.2 – Construção dos Oligonucleotídeos e Amplificação dos Genes..	46
II. 2.3 – Purificação dos Produtos Amplificados por PCR.....	49
II. 2.4 – Preparação dos Produtos de PCR purificados – Reação de Adenilação.....	49
II. 2.5 – Clonagem dos Produtos de PCR adenilados nos Vetores de Clonagem pGEM-T easy - Reação de Ligação.....	50
II. 2.6 – Transformação da Bactéria <i>Escherichia coli</i> DH5 α com os Plasmídeos Recombinantes pGEM-T easy.....	51
II. 2.7 – Extração, Análise de Restrição e Sequenciamento dos Plasmídeos Recombinantes pGEM-T easy de <i>Escherichia coli</i> DH5 α	52
II. 2.8 – Subclonagem dos Genes no Plasmídeo de Expressão pET28a...	54
II. 2.9 – Transformação da Bactéria <i>Escherichia coli</i> DH5 α com os Plasmídeos Recombinantes pET28a	55
II. 2.10 – Extração e Análise de Restrição dos Plasmídeos Recombinantes pET28a de <i>Escherichia coli</i> DH5 α	55
II. 2.11 – Transformação da Bactéria <i>Escherichia coli</i> BL21(DE3) com os Plasmídeos Recombinantes pET28a.....	56
II. 2.12 – Expressão Heteróloga das Proteínas Recombinantes de <i>Xylella fastidiosa</i> (rXfHNL, rXftatD e rXfaroC) em <i>E.coli</i>	56
II. 2.13 – Purificação das Proteínas Recombinantes de <i>Xylella fastidiosa</i> (rXfHNL, rXftatD e rXfaroC)	58
II. 2.13.1 – Cromatografia de Afinidade em Resina de Níquel.....	59
II. 2.13.2 – Cromatografia de Exclusão por Tamanho em Colunas Superdex.....	62
II. 2.14 - Análises Físico-química e Estrutural das Proteínas Recombinantes de <i>Xylella fastidiosa</i> (rXfHNL, rXftatD e rXfaroC)	63
II. 2.14.1 – Determinação Quantitativa do Nível de Expressão Protéica.....	63
II. 2.14.1.1 - Eletroforese em Gel de Poliacrilamida (SDS-PAGE).....	64
II. 2.14.1.2 - Eletroforese em Gel através de Microchip (Método GelChip-CE).....	65
II. 2.14.2 - Determinação do Ponto Isoelétrico.....	66
II. 2.14.3 - Determinação da Sequência por Espectrometria de Massas.....	68
II. 2.14.4 – Espalhamento de Raios-X a Baixos Ângulos (SAXS) e Análises dos Dados.....	70
II. 2.14.5 – Espectroscopia de Dicroísmo Circular (CD) e Cálculo das Frações de Estrutura Secundária.....	73

II. 2.14.6 - Espectroscopia de Emissão de Fluorescência.....	76
II. 2.14.7 – Ensaios de Estabilidade Térmica e Química das Proteínas Recombinantes.....	78
II. 2.14.7.1 – Ensaio de Estabilidade Térmica Frente a Variações de pH.....	78
II. 2.14.7.2 - Ensaio de Estabilidade Química Frente a Variações na Concentração de Uréia.....	79
II. 2.14.8 – Estudos de Atividade Enzimática das Proteínas Recombinantes.....	80
II. 2.14.8.1 – Hidroxinitrila Liase de <i>Xylella fastiosa</i> (rXfHNL).....	80
II. 2.14.8.2 – Corismato Sintase de <i>Xylella fastiosa</i> (rXfaroC)	81
<u>III – RESULTADOS E DISCUSSÃO</u>	83
III. 1 – Hidroxinitrila Liase	83
III. 1.1 – Amplificação e Clonagem no Vetor pGEM-T <i>easy</i>	83
III. 1.1.1 – Análise de Restrição e Seqüenciamento dos Plasmídeos Recombinantes.....	87
III. 1.2 – Subclonagem no Vetor de Expressão e Análise de Restrição dos Plasmídeos Recombinantes.....	90
III. 1.3 – Expressão Heteróloga da Proteína rXfHNL a partir dos Vetores de Expressão.....	92
III. 1.4 – Purificação da Proteína rXfHNL por Cromatografia de Afinidade e Exclusão por Tamanho.....	93
III. 1.5 – Análises Físico-química e Estrutural da Proteína rXfHNL.....	98
III. 1.5.1 – Determinação Quantitativa do Nível de expressão Protéica	98
III. 1.5.2 – Determinação do Ponto Isoelétrico	102
III. 1.5.3 – Determinação da Seqüência por Espectrometria de Massas.....	103
III. 1.5.4 – Análise dos Dados de SAXS.....	105
III. 1.5.5 – Espectroscopia de Dicroísmo Circular.....	108
III. 1.5.5.1 – Estudos de Estabilidade Térmica.....	109
III. 1.5.5.2 – Estudos da Estabilidade Química.....	113
III. 1.5.6 – Espectroscopia de Emissão de Fluorescência.....	120
III. 1.5.6.1 – Estudos da Estabilidade Química.....	122
III. 1.5.7 – Ensaio de Atividade Enzimática	126
III. 1.5.8 – Conclusões Parciais	128
III. 2 – Corismato Sintase	130
III. 2.1 – Amplificação e Clonagem no Vetor pGEM-T <i>easy</i>	130
III. 2.1.1 – Análise de Restrição e Seqüenciamento dos Plasmídeos Recombinantes	132

III. 2.2 – Subclonagem no Vetor de Expressão, Análise de Restrição e Sequenciamento dos Plasmídeos Recombinantes	134
III. 2.3 – Expressão Heteróloga da Proteína rXfaroC a partir dos Vetores de Expressão	137
III. 2.4 – Purificação da Proteína rXfaroC por Cromatografia de Afinidade.....	139
III. 2.5 – Análises Estrutural da Proteína rXfaroC	141
III. 2.5.1 – Espectroscopia de Dicroísmo Circular	141
III. 2.5.2 – Espectroscopia de Emissão de Fluorescência	144
III. 2.5.3 – Ensaio de Atividade Enzimática.....	146
III. 2.6 – Conclusões Parciais	148
III. 3 – Proteína D do Sistema de Transporte e Secreção Tat	150
III. 3.1 – Amplificação e Clonagem no Vetor pGEM-T easy	150
III. 3.2 – Análise de Restrição e Sequenciamento dos Plasmídeos Recombinantes	152
III. 3.3 – Subclonagem no Vetor de Expressão e Análise de Restrição dos Plasmídeos Recombinantes	154
III. 3.4 – Expressão Heteróloga da Proteína rXftatD a partir dos Vetores de Expressão	156
III. 3.5 – Conclusões Parciais.....	158
IV – PERSPECTIVAS	159
V – REFERÊNCIAS BIBLIOGRÁFICAS	160

LISTA DE FIGURAS

Figura 1	O ciclo da reação em cadeia da polimerase	3
Figura 2	Tipos de clivagem gerados por enzimas de restrição	7
Figura 3	Moléculas de DNA de tamanhos diferentes podem ser separadas por eletroforese	9
Figura 4	Estrutura molecular de um plasmídeo típico usado em clonagem molecular	11
Figura 5	Construção de um DNA recombinante	13
Figura 6	Mecanismo molecular proposto para a transformação de <i>E. coli</i> com uma molécula de DNA exógeno.....	16
Figura 7	Seqüenciamento de DNA pelo método de Sanger	20
Figura 8	Expressão e purificação de uma proteína recombinante com <i>tag</i> de seis histidinas	23
Figura 9	Mecanismo enzimático de glicosídeos cianogênicos	28
Figura 10	Via do chiquimato	33
Figura 11	Destino do corismato a partir da via do ácido chiquímico e os produtos finais correspondentes	34
Figura 12	Previsões topológicas derivadas das seqüências previstas para os genes do sistema Tat.....	39
Figura 13	Seqüência codificadora da ORF XfHNL (<i>Xylella fastidiosa</i> Genome Project – número de acesso XFa0032)	47
Figura 14	Interação entre resíduos vizinhos da cauda <i>His-Tag</i> e a matriz Ni-NTA	60
Figura 15	Estrutura molecular do composto imidazol e do aminoácido histidina	60

Figura 16	Espectro de CD de vários tipos de estrutura secundária	74
Figura 17	Diagrama de Jablonski	77
Figura 18	Amplificação da ORF codificadora da proteína hidroxinitrila liase.....	85
Figura 19	Colônias Recombinantes e Não Recombinantes de Células <i>E. coli</i> DH5 α	86
Figura 20	Separação dos plasmídeos digeridos pGEM-XfHNL	88
Figura 21	Bases seqüenciadas do plasmídeo pGEM-XfHNL extraído do clone 1	89
Figura 22	Separação dos plasmídeos digeridos pET-XfHNL	91
Figura 23	Expressão da proteína rXfHNL a partir dos clones 1 e 2	92
Figura 24	Purificação da rXfHNL por cromatografia de afinidade	94
Figura 25	Perfil cromatográfico da rXfHNL por cromatografia de exclusão por tamanho e SDS-PAGE das frações eluídas	95
Figura 26	Determinação do grau de oligomerização aparente da rXfHNL por cromatografia de exclusão por tamanho	97
Figura 27	Géis representativos das análises por (A) SDS-PAGE e (B) GelChip-CE	99
Figura 28	Resultado da expressão da rXfHNL em função do tempo por GelChip-CE e SDS-PAGE	100
Figura 29	Ensaio de focalização isoeétrica da rXFHNL	102
Figura 30	Cobertura da seqüência de aminoácidos da proteína rXfHNL obtida pelas análises por LC-nanoESI-MS/MS	104
Figura 31	Curvas de espalhamento de raios-X a baixos ângulos e gráfico da função de variação de distância da rXfHNL	106
Figura 32	Modelo de SAXS para a rXfHNL	107

Figura 33	Espectro de dicroísmo circular da rXfHNL.....	109
Figura 34	Espectros de CD da rXfHNL em diferentes pH	110
Figura 35	Transição da desnaturação térmica da rXfHNL	112
Figura 36	Dependência do T_m de rXfHNL com o pH, monitorado por CD	113
Figura 37	Desnaturação da rXfHNL induzida por uréia	114
Figura 38	Desnaturação e renaturação da rXfHNL induzida por uréia	118
Figura 39	Espectros de emissão de fluorescência da rXfHNL a vários pH	120
Figura 40	Efeitos da desnaturação induzida por uréia	122
Figura 41	Desnaturação induzida por uréia da rXfHNL por Fluorescência	124
Figura 42	Amplificação do gene codificador da proteína corismato sintase	131
Figura 43	Separação dos plasmídeos digeridos pGEM-aroC	133
Figura 44	Separação dos plasmídeos digeridos pET-XfaroC	135
Figura 45	Bases seqüenciadas do plasmídeo pET-XfaroC extraído do clone 3	136
Figura 46	Expressão piloto da proteína rXfaroC a partir do clone 1	138
Figura 47	Purificação da rXfaroC por cromatografia de afinidade	139
Figura 48	Purificação da rXfaroC por cromatografia de afinidade	140
Figura 49	Espectro de dicroísmo circular da proteína recombinante rXfaroC	142
Figura 50	Espectros de CD da rXfHNL em diferentes pHs	143
Figura 51	Espectros de emissão de fluorescência da rXfaroC a vários valores de pH	145

Figura 52	Análise de restrição dos plasmídeos pGEM-XftatD	152
Figura 53	Análise de restrição dos plasmídeos pET-XftatD	155
Figura 54	Expressão Piloto da proteína rXftatD a partir dos clones 1 e 2	156

LISTA DE TABELAS

Tabela 1	Endonucleases ou enzimas de restrição.....	8
Tabela 2	Oligonucleotídeos iniciadores utilizados na amplificação por PCR...	48
Tabela 3	Dados termodinâmicos para diferentes concentrações da rXfHNL por CD	119
Tabela 4	Dados termodinâmicos para diferentes concentrações da rXfHNL por Fluorescência.	125

LISTA DE ABREVIATURAS E SÍMBOLOS

Símbolo dos aminoácidos

Código de 3 letras	Código de 1 letra	Nome do aminoácido
Ala	A	Alanina
Arg	R	Arginina
Asn	N	Asparagina
Asp	D	Aspartato
Cys	C	Cisteína
Gln	Q	Glutamina
Glu	E	Glutamato
Gly	G	Glicina
His	H	Histidina
Iso	I	Isoleucina
Leu	L	Leucina
Lys	K	Lisina
Met	M	Metionina
Phe	F	Fenilalanina
Pro	P	Prolina
Ser	S	Serina
Thr	T	Treonina
Trp	W	Triptofano
Tyr	Y	Tirosina
Val	V	Valina

Abreviaturas

CA	Anidrase carbônica
BLAST	<i>Basic Local Alignment Search Tool</i>
BSA	Albumina sérica bovina
CD	<i>Circular Dichroism</i> , dicroísmo circular
D.O.	Densidade Óptica
EDTA	Ácido etileno diamino tetracético

LC-nanoESI-MS/MS	<i>Liquid chromatography nano Electros Spray Ionization – Mass Spectrometry</i> , Espectrometria de massa por ionização elétron-spray
IEF	<i>IsoElectric Focussing</i> , Focalização Isoelétrica
ITS	Inibidor de tripsina de soja
NCBI	<i>National Center for Biotechnology Information</i>
OVA	Ovalbumina
pI	Ponto isoelétrico
SDS-PAGE	<i>Sodium DodecylSulphate – Polyacrilamide Gel Electrophoresis</i> , eletroforese em gel de poliacrilamida-SDS
SELCON, Continll e CDSSTR	Programas de análise de espectros de CD que estimam o conteúdo de estrutura secundária de uma proteína
u.a.	Unidade arbitrária

RESUMO

A bactéria *Xylella fastidiosa* (*X. fastidiosa*) é um patógeno de planta que causa a clorose variegada dos citros, também conhecida como amarelinho. *Xylella fastidiosa* é uma bactéria limitada ao xilema, sendo transmitida de planta a planta através de insetos vetores. Através da determinação experimental da seqüência primária de algumas proteínas marcadamente expressas pela *X. fastidiosa* e a comparação destas com seqüências de ácidos nucleicos do genoma identificou entre elas três ORFs codificadoras de proteínas que podem estar relacionadas à patogenicidade desta bactéria no seu hospedeiro. No entanto, a função biológica destas proteínas ainda não foi funcionalmente caracterizada. Para investigar o papel biológico e realizar estudos de estrutura e função, as ORFs correspondentes as proteínas hidroxinitrila liase, corismato sintase e proteína D do sistema de transporte e secreção Tat foram clonadas, seqüenciadas e expressas em *Escherichia coli*. As proteínas recombinantes expressas foram purificadas por cromatografia de afinidade e suas identidades verificadas por seqüenciamento. As proteínas purificadas foram encontradas como uma única banda em SDS-PAGE. As proteínas recombinantes, pela primeira vez isoladas, foram caracterizadas em termos de estabilidade conformacional e comportamento estrutural frente a mudanças de pH e temperatura utilizando-se as espectroscopias de dicroísmo circular e fluorescência. O sucesso na construção do gene sintético e na clonagem em vetores de expressão de *E. coli* resultou em quantidade satisfatória de expressão das proteínas recombinantes. Duas delas apresentaram-se na formas solúveis, facilmente purificadas e ativas e permitindo suas caracterizações.

ABSTRACT

Xylella fastidiosa is the causal agent of citrus variegated chlorosis, also known as "amarelinho". *Xylella fastidiosa* inhabits exclusively the xylem vessels, and is transmitted by sharpshooter vectors. The experimental determination of the primary sequence of some markedly expressed proteins for *X. fastidiosa* and the comparison with the nucleic acids sequence of its genome aided the identification of three of them as being proteins involved in host pathogenicity. However, the biological role of these proteins should be assigned experimentally. In order to investigate the biological role, function and structure, ORFs corresponding to hydroxynitrile liase, chorismate synthase and type V secretory pathway proteins were cloned, sequenced and expressed in *Escherichia coli*. The expressed recombinant proteins were purified by immobilized metal affinity chromatography (Ni-NTA resin) and their identities were verified by protein sequencing. The purified proteins were found as a single band on SDS-PAGE. The successful cloning and heterologous expression of recombinant HNL in *E. coli* resulted in a satisfactory amount of expressed protein. Two of the expressed recombinant proteins were soluble, easily purified, and isolated in an active form. *X. fastidiosa* recombinant proteins, for the first time isolated, were characterized according to enzyme conformational stability and structural behavior as a function of pH and temperature using circular dichroism and fluorescence spectroscopy.

I - INTRODUÇÃO

Até meados da década de 70, o ácido desoxirribonucléico (DNA) foi o componente celular mais difícil de ser analisado por procedimentos científicos que levassem a sua caracterização químico-funcional dentro dos organismos vivos (STRYER, 1988). A partir de 1970, novas tecnologias provenientes da Microbiologia, Bioquímica, Imunologia e Genética Microbiana foram desenvolvidas, permitindo o isolamento e a purificação de genes específicos num processo chamado clonagem gênica (ZAHA *et al.*, 1996; STRYER, 1988). A clonagem gênica permitiu que a análise do DNA ganhasse um novo enfoque, fazendo com que a dificuldade encontrada na análise desta molécula diminuísse muito, tornando-o, assim, umas das moléculas mais fáceis de ser analisada. Muitas destas técnicas permitiram isolar regiões específicas do DNA, obtê-las em grande quantidade e determinar a sua seqüência numa velocidade de milhares de nucleotídeos por dia.

Como consequência do desenvolvimento destas tecnologias, a tecnologia do DNA recombinante, como se convencionou denominar este conjunto de técnicas, teve ampla aplicação tanto na área biológica como em diversas outras áreas, tais como: médica, forense, agrícola e química. Atualmente é possível realizar investigações de paternidade, diagnósticos de doenças genéticas e infecciosas, estudos dos mecanismos de replicação e expressão, estudos do desenvolvimento de culturas microbianas capazes de produzir substâncias úteis tais como a insulina humana, hormônio de crescimento e vacinas, e ainda, produzir enzimas de interesse industrial aplicáveis no processo de fabricação de pesticidas e fármacos, entre outros. Portanto, sua aplicação comercial ou biotecnológica parece ter um potencial inesgotável.

Diante da diversidade de aplicações, a tecnologia do DNA recombinante foi de extrema importância neste estudo tanto na obtenção dos genes de interesse como na confirmação de suas seqüências e, conseqüentemente, na obtenção das proteínas que eles codificam. Nos próximos tópicos pretende-se descrever os passos básicos envolvidos na construção de genes recombinantes, na expressão das respectivas proteínas codificadas por eles e, finalmente, explanar sobre o organismo e as funções das proteínas de interesse.

I. 1 – Reação em Cadeia da Polimerase - PCR

A tecnologia da reação em cadeia da polimerase (“Polymerase Chain Reaction”) foi concebida por Kary Müllis da década de 80 (ZAHA *et al.*, 1996; STRYER, 1988). A maior aplicação desta metodologia é a possibilidade de se amplificar uma determinada seqüência de DNA através da síntese enzimática *in vitro* de milhões de cópias de um segmento específico na presença da enzima DNA polimerase.

Uma simples cópia de um gene dentro de um genoma pode ser amplificada a alguns microgramas, partindo-se de quantidades mínimas como sub picogramas de DNA total. A condição básica para a aplicação da técnica, depende da construção de um par de oligonucleotídeos (“primers” ou iniciadores) os quais são complementares as duas fitas opostas do DNA e que delimitam a seqüência alvo da amplificação.

A PCR envolve basicamente 3 etapas: desnaturação, hibridização e extensão, os quais constituem um ciclo da reação de amplificação. A Figura 1 ilustra a representação de um ciclo básico da PCR.

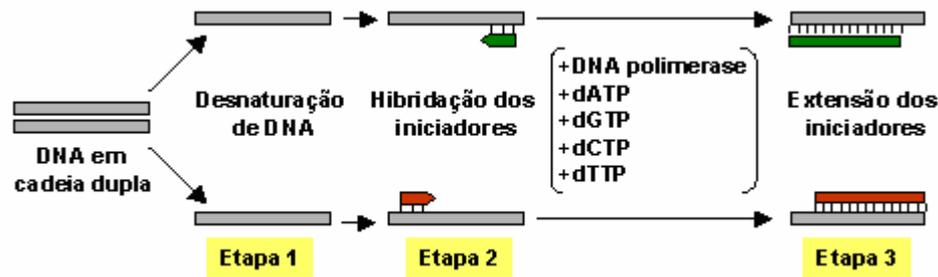


Figura 1. O ciclo da reação em cadeia da polimerase. O processo envolvendo estas três etapas pode ser repetido de 25 a 30 ciclos, sendo possível aumentar, em cada ciclo, duas vezes a concentração de DNA pré-existente. Em teoria, se for possível levar a cabo 25 ciclos de amplificação seguidos, a concentração de DNA aumentaria 2^{25} vezes embora, na prática, devido a ineficiência no processo de amplificação, esse aumento fique em alguns milhões de vezes.

Na primeira etapa, a fita dupla do DNA alvo sofre desnaturação a temperaturas em torno de 94 °C durante 5 min. Nesta temperatura, as duas fitas do DNA alvo se separam. Na etapa de anelamento, a temperatura é rapidamente reduzida para 35 a 60 °C permitindo que os oligonucleotídeos sejam hibridizados especificamente na região do DNA molde em que há complementariedade das bases. Finalmente, a temperatura é elevada para aproximadamente 72 °C para que a DNA polimerase realize a extensão a partir das extremidades 3' dos *primers*.

Inicialmente, a DNA polimerase da bactéria *Escherichia coli* (*E. coli*), cuja temperatura ótima é de 37 °C foi muito empregada e adicionada manualmente a cada ciclo na amplificação de DNA. Este processo manual era realizado a fim de preservá-la durante a PCR já que ela era inativa a 94 °C. Um importante avanço técnico surgiu com a descoberta da *Taq* DNA polimerase (ZAHA *et al.*, 1996; STRYER, 1988), oriunda da bactéria *Thermus aquaticus*, a qual vive em fontes de água a temperatura de 75 °C. Com esta nova enzima foi possível programar, de forma contínua e automatizada, vários ciclos de PCR em equipamentos denominados termocicladores. Por apresentar temperatura ótima de ação a 72 °C e ser razoavelmente estável mesmo

a 94 °C, esta descoberta permitiu que a adição de enzima fosse realizada uma única vez durante toda a reação.

Além do processo de amplificação de seqüências específicas, principal uso da PCR, rapidamente surgiram novas aplicações para esta poderosa metodologia, tais como: utilização no seqüenciamento de nucleotídeos, análise do polimorfismo e várias outras aplicações mais específicas atribuídas ao emprego da metodologia do DNA recombinante.

A amplificação via PCR é a etapa primordial para a obtenção da quantidade de DNA necessária a ser aplicada na etapa de clonagem molecular.

I. 2 - Clonagem Molecular

I. 2.1 – Conceito de Clonagem Molecular

A origem do termo clonagem vem da genética bacteriana que considera uma colônia de bactérias como um clone, uma vez que todos os indivíduos são geneticamente idênticos à bactéria inicial.

A técnica central da metodologia do DNA recombinante é a clonagem molecular, a qual consiste no isolamento e propagação de moléculas de DNA idênticas. Em geral, a clonagem molecular compreende dois estágios importantes:

i) a ligação de um fragmento de DNA de interesse, denominado inserto, a uma outra molécula de DNA, denominada vetor, a fim de formar uma terceira molécula, o DNA recombinante;

ii) a molécula do DNA recombinante é introduzida numa célula hospedeira compatível, num processo chamado de transformação. A célula hospedeira que

adquiriu a molécula do DNA recombinante é agora chamada de transformante ou célula transformada. Um único transformante, em condições ideais, sofre muitos ciclos de divisão celular, produzindo uma colônia que contém milhares de cópias do DNA recombinante.

I. 2.2 – Enzimas de Restrição

A descoberta das enzimas de restrição revolucionou a biologia molecular, pois possibilitou montar moléculas de DNA recombinante, tornando possível a manipulação do DNA com extrema precisão.

Em 1953, um fenômeno de deficiência da replicação de um vírus de bactérias (bacteriófagos ou simplesmente fagos) foi descrito dependente da célula hospedeira na qual este vírus estava inserido. Com o avanço das pesquisas em torno deste fenômeno, foi observado que a inabilidade de certos vírus crescerem em determinadas linhagens bacterianas era devido à presença de enzimas denominadas nucleases altamente específicas que clivavam o DNA viral, tornando-os inativos. Este processo foi encarado como um sistema de defesa bacteriano que degrada o DNA que lhe é estranho (DNA exógeno), e é denominado processo de restrição.

Neste processo, as bactérias protegem seu próprio DNA desta degradação através da sua modificação química após a síntese. Esta modificação consiste na metilação de nucleotídeos (bases) específicos ao longo da fita dupla de DNA. Sendo assim, as enzimas de restrição são capazes de distinguir o DNA exógeno de seu próprio DNA. A enzima responsável pelo processo de metilação, a metilase, também pode, teoricamente, adicionar radicais metil no DNA do vírus por engano. No entanto, a metilase é uma enzima de ação lenta, e antes que ela realize a metilação do DNA do

vírus, a enzima de restrição já o eliminou. Na bactéria o processo de metilação ocorre logo no início da formação de um novo indivíduo. Como consequência, este sistema é frequentemente descrito como fenômeno da restrição/modificação e existe em um grande número de bactérias.

Em 1973, o interesse por enzimas de restrição aumentou quando se percebeu que elas poderiam ser usadas para fragmentar o DNA deixando em suas extremidades fitas simples de DNA que permitiam a ligação dos fragmentos (ZAHA *et al.*, 1996; STRYER, 1988). Isto significava que a recombinação poderia ser efetuada *in vitro*. Além disto, DNA bacteriano poderia recombinar com DNA humano ou de qualquer outra espécie, abrindo a possibilidade de clonar genes ou isolar proteínas de culturas bacterianas.

Enzimas de restrição ou endonucleases de restrição são proteínas bacterianas que reconhecem seqüências nucleotídicas específicas de 4 a 8 pares de base (pb) em uma molécula de DNA, clivando a mesma em dois lugares da fita dupla que a constitui. A seqüência de reconhecimento é chamada de sítio de restrição. Estas enzimas são divididas em várias classes, dependendo da estrutura, da atividade e dos sítios de reconhecimento e clivagem.

Na tecnologia do DNA recombinante, as enzimas do tipo II, que podem ser proteínas monoméricas ou diméricas, são as mais importantes e cortam o DNA no mesmo sítio do seu reconhecimento. O sítio de reconhecimento deste tipo de enzima é normalmente uma seqüência denominada palindrômica, ou seja, ela tem um eixo de simetria e a seqüência de bases de uma fita é a mesma da fita complementar, quando lida na direção oposta. Há 2 tipos distintos de clivagens: a) os dois cortes ocorrem no eixo de simetria da seqüência específica, gerando extremidades abruptas, ou b) os cortes são feitos simetricamente, porém, fora do eixo de simetria, gerando

extremidades coesivas. Estas extremidades apresentam regiões constituídas de fita simples. Exemplos destes dois tipos de clivagens e seus produtos gerados foram representados na Figura 2.

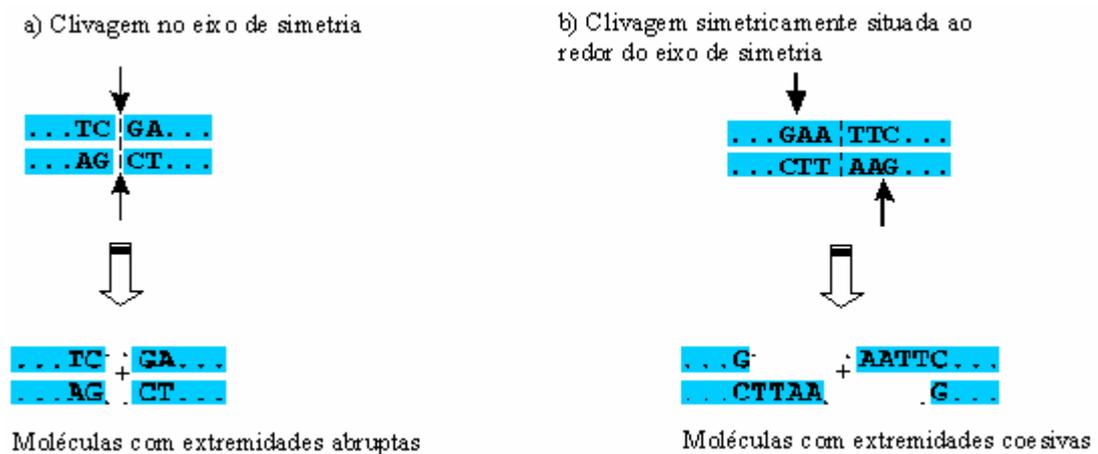


Figura 2. Tipos de clivagem gerados por enzimas de restrição. Os sítios de reconhecimento são compostos por seqüências palindrômicas. As setas indicam os sítios de clivagem. As linhas tracejadas representam o centro de simetria das seqüências.

As extremidades coesivas produzidas por várias enzimas de restrição permitem que dois fragmentos de DNA liguem-se mais facilmente do que fragmentos que contenham extremidades abruptas devido serem estruturalmente mais adequadas para a complementaridade das bases. Assim, fragmentos com extremidades abruptas são ligados com muito menos eficiência. Uma concentração muito maior de DNA ligase e dos DNAs envolvidos ou até uma modificação destas extremidades a fim de torná-las coesivas são necessários para que ocorra a reação de ligação.

Atualmente, mais de 1000 enzimas de restrição já foram identificadas (LEWIN, 1997; LEWIN, 1994). A nomenclatura desenvolvida para elas foi baseada na abreviação do nome do microrganismo do qual a enzima foi isolada. A primeira letra

representa o gênero e as outras duas à espécie, seguido de um algarismo romano ou outra letra que indica a ordem da descoberta ou a linhagem da qual ela foi isolada. Por exemplo, a enzima de restrição denominada de *EcoRI* é purificada da bactéria *E. coli* que carrega um fator de transferência de resistência RY, enquanto que a *HindIII* é isolada da *Haemophilus influenzae*, linhagem Rd III.

A Tabela 1 mostra o organismo de origem, a seqüência palindrômica e o sítio de clivagem de algumas enzimas de restrição. Note que há enzimas que originam terminações coesivas enquanto que uma gera cortes abruptos.

Tabela 1. Endonucleases ou enzimas de restrição. As setas verdes dentro das seqüências de reconhecimento indicam o local de clivagem em uma fita dupla de DNA.

Organismo	Denominação da Enzima	Seqüência de Reconhecimento
<i>Escherichia coli</i> RY 13	<i>EcoRI</i>	5' G↓AATTC 3' 3' CTTAA↑G 5'
<i>Haemophilus influenzae</i> Rd	<i>HindIII</i>	5' A↓AGCTT 3' 3' TTCGA↑A 5'
<i>Haemophilus aegyptius</i>	<i>HaeIII</i>	5' AG↓GCCT 3' 3' TC↑CGGA 5'
<i>Xantomonas hilicola</i>	<i>XhoI</i>	5' C↓TCGAG 3' 3' GAGCT↑C 5'
<i>Neisseria denitrificans</i>	<i>NdeI</i>	5' CA↓TATG 3' 3' GTAT↑AC 5'
<i>Termus aquaticus</i>	<i>TaqI</i>	5' A↓AGCTT 3' 3' TTCGA↑A 5'

A digestão de DNA por enzimas de restrição é um processo bastante simples. Basta colocar o DNA em contato com a enzima a uma temperatura ideal, geralmente 37 °C, e a enzima inicia imediatamente o processo de digestão, cortando o DNA em diversos fragmentos. O número de fragmentos produzido é estabelecido pelo número de sítios de restrição reconhecidos pela enzima utilizada. Uma importante consequência desta especificidade é que o número de clivagens feito por cada uma

delas no DNA de qualquer organismo é definido e permite o isolamento de fragmentos deste DNA. Portanto, cada enzima de restrição gera um conjunto único de fragmentos quando cliva uma molécula de DNA específica.

O conjunto de fragmentos gerados por digestão com enzimas de restrição é geralmente detectado pela separação destes fragmentos por eletroforese em gel de agarose. A Figura 3 mostra uma separação de DNAs digeridos e visualizados em gel de agarose.

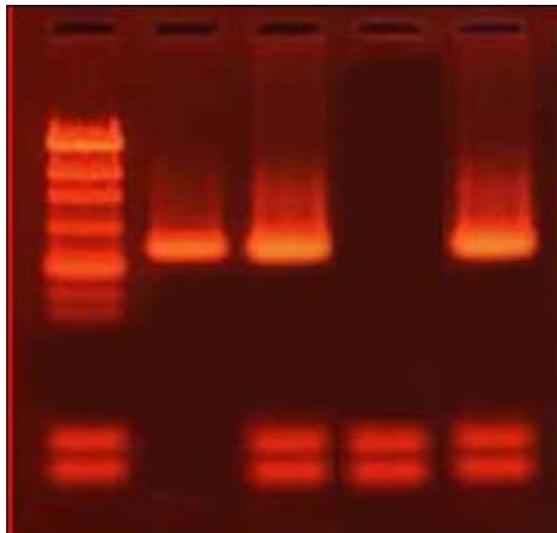


Figura 3. Moléculas de DNA de tamanhos diferentes podem ser separadas por eletroforese. Moléculas menores migram mais rapidamente que moléculas maiores, tornando-se, portanto separadas em bandas. O DNA é corado com brometo de etídeo, que fluoresce quando iluminado com luz ultravioleta.

A eletroforese é a técnica pela qual fragmentos de DNA de diferentes tamanhos são separados. O DNA é colocado em um poço do gel (que tem aspecto de gelatina) e este é submetido a um campo elétrico. O DNA irá se mover na direção do pólo positivo, uma vez que a molécula de DNA é negativa devido à presença de grupamentos fosfato. A separação dos fragmentos ocorre em função dos seus pesos

moleculares que determina o quão rápido o campo elétrico será capaz de mover o fragmento de DNA ao longo do gel. Quanto maior o fragmento, mais lenta será sua migração, deste modo, fragmentos menores chegam primeiro ao final do gel.

As enzimas de restrição constituem a principal ferramenta da tecnologia do DNA recombinante. Sua aplicação permite cortar genes específicos de interesse e colocá-los dentro de outros organismos, tornando possível recombinar ou criar organismos transgênicos, tão conhecidos popularmente.

I. 2.3 – Vetores de Clonagem

Após o isolamento de uma informação genética, os fragmentos de DNA obtidos pela clivagem com enzimas de restrição devem ser inseridos numa outra molécula de DNA, capaz de amplificar esta informação genética em centenas de cópias. Este processo de amplificação é obtido através do uso de moléculas de DNA que são os chamados de vetores de clonagem molecular.

Atualmente, os tipos básicos de vetores usados na metodologia do DNA recombinante apresentam características especiais que os tornam excelentes veículos de clonagem em diferentes situações. Os vetores de clonagem mais utilizados normalmente são os plasmídeos de bactérias - moléculas de DNA circular existentes nesses microorganismos, fora dos seus cromossomos e capazes de se replicar de modo independente. A técnica de DNA recombinante tem um interesse especial se uma das fontes de DNA clivado for um plasmídeo, justamente pelo seu potencial de replicação independente.

Vetores de clonagem são pequenas moléculas de DNA dupla fita, contendo os elementos necessários para a sua replicação e pelo menos um gene que confere

resistência a antibiótico. Estas moléculas variam de 5 a 400 quilobases e comumente estão presentes em duas ou mais cópias por célula. Os plasmídeos presentes num grande número de cópias são usados como veículos de clonagem desde que capacitem à amplificação do segmento do DNA neles clonado.

Um plasmídeo é considerado um bom vetor de clonagem se: possui uma origem de replicação (**O**), ou seja, uma seqüência de DNA que permita que o vetor seja replicado na célula hospedeira; apresenta dois ou mais sítios únicos de clivagem para endonucleases de restrição onde o inserto é incorporado e possui um gene que codifica um produto que distingue a célula transformada da célula não transformada. A Figura 4 ilustra as principais características estruturais de um vetor de clonagem.

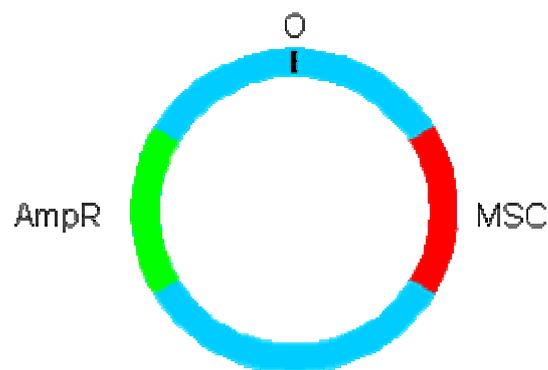


Figura 4. Estrutura molecular de um plasmídeo típico usado em clonagem molecular. O é a origem de replicação, Amp^R é o gene que confere resistência ao antibiótico ampicilina e MSC é o sítio de múltipla clonagem.

O conjunto dos sítios de clivagem é denominado de múltiplos sítios de clonagem (MSC) que permite sua abertura por várias enzimas, a escolha depende do DNA exógeno. Muitos vetores de clonagem carregam genes que conferem resistência a diversos antibióticos tais como ampicilina (Amp^R), tetraciclina e canamicina. As células transformadas com tais vetores são capazes de crescer num meio contendo o

antibiótico, enquanto que as células não transformadas acabam morrendo. Porém, existem ainda outras marcas de seleção úteis, as marcas auxotróficas. São representadas por genes que codificam proteínas cuja síntese é deficiente na célula hospedeira e que estão envolvidas na síntese de um composto essencial para a seleção dos clones positivos.

I. 2.4 – Construção do DNA Recombinante

O início do processo de clonagem envolve a construção de vetores contendo os genes de interesse, denominado construção do DNA recombinante. As construções gênicas, feitas através de reações enzimáticas que cortam e ligam o DNA em fragmentos específicos, são geralmente compostas de três partes principais dos genes: a região promotora, a região codificadora e a região terminal. Cada uma dessas regiões, no fragmento a ser transplantado, pode ter como origem um organismo diferente.

Na construção do DNA recombinante a primeira etapa envolve, como descrita anteriormente, a produção de fragmentos de DNA de origens diferentes com o mesmo conjunto de extremidades fitas simples, ou seja, a digestão das moléculas de DNA com a mesma enzima de restrição. Numa segunda etapa, estes dois fragmentos (por exemplo, provenientes de uma bactéria e de um vegetal) são ligados por renaturação das regiões de fita simples. Se a ligação for feita com uma enzima, a DNA ligase, após o pareamento das bases estes fragmentos serão ligados permanentemente.

A Figura 5 mostra uma representação da construção de um DNA Recombinante com somente um sítio de clivagem para uma determinada enzima de restrição.

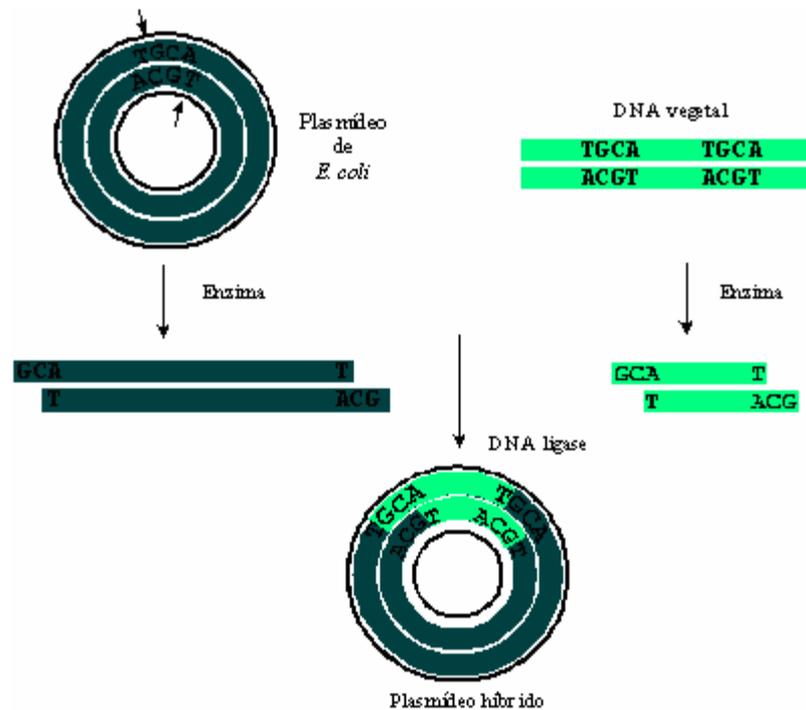


Figura 5. Construção de um DNA recombinante. Uma molécula de DNA híbrida é formada a partir de fragmentos de diferentes organismos obtidos com o uso de uma mesma enzima de restrição.

Se os fragmentos de DNA vegetal são misturados com os DNAs bacterianos linearizados previamente, permitindo a ligação entre eles, uma nova molécula de DNA plasmidial contendo DNA vegetal pode ser gerada. Este é denominado plasmídeo híbrido e pode ser inserido numa bactéria por transformação. Sendo assim, o inserto será replicado como se fosse parte integrante do plasmídeo.

I. 2.4.1 – Enzimas DNA Ligases

Algumas bactérias e vírus, como a *E. coli* e o fago T4, codificam enzimas DNA ligase capazes de unir fragmentos de DNA contendo extremidades abruptas ou coesivas. A maior diferença entre elas é que as DNA ligases isoladas de bactérias requerem moléculas de nicotinamida adenina dinucleotídeo (NAD^+) como cofator

para exercerem suas atividades, enquanto que a isolada do bacteriófago T4 requer adenosina tri-fosfato (ATP).

Após preparação das duas moléculas de DNA que se pretendem recombinar *in vitro*, segue-se o processo de ligação. Sendo o vetor de clonagem e o fragmento de DNA a clonar tratados com a mesma enzima de restrição, suas extremidades serão compatíveis e a reação de ligação poderá ocorrer na presença de uma DNA ligase. Consequentemente, o emparelhamento entre bases complementares por ligações de hidrogênio também ocorrerá na seqüência.

A ligação é promovida pela ligação covalente (uma ligação fosfodiéster) entre os dois nucleotídeos das extremidades das duas moléculas. A enzima DNA ligase catalisa a ligação entre o grupo fosfato ligado ao carbono 5' e o grupo hidroxila, OH livre, ligado ao carbono 3' das moléculas de desoxirribose presentes nos nucleotídeos que constituem estas moléculas.

Neste passo, pode acontecer que os vetores de clonagem recircularizem, promovendo uma ligação intramolecular, em vez de haver integração do outro fragmento de DNA a clonar, resultando no retorno ao processo inicial. Além disso, pode ainda haver formação de dímeros do vetor sem ligação do fragmento de DNA. Embora seja relativamente fácil selecionar as colônias que abrigam os vetores de clonagem utilizando os marcadores de resistência a antibióticos, é muito mais difícil selecionar o clone que contenha o gene de interesse. Mas existe, no entanto, estratégias complementares de seleção a fim de minimizar estes problemas de recircularização do vetor.

A etapa seguinte consiste na transformação das células hospedeiras adequadas com as moléculas de DNA recombinante (rDNA ou vetor com o DNA exógeno clonado).

I. 2.5 – Transformação Bacteriana

I. 2.5.1 – Conceito de Transformação Induzida

O processo de transformação constitui um evento de alta importância na técnica de manipulação gênica. A transformação natural, descrita por Griffith e Avery, é um evento raro verificado apenas em determinadas espécies e em condições fisiológicas específicas. No entanto, em 1970, Mandel e Higa encontraram que a *E. coli* tornou-se marcadamente competente para transformação com DNA exógeno, quando a bactéria foi suspensa em cloreto de cálcio gelado e submetida a um curto choque térmico a 42 °C (LEWIN, 1997; ZAHA *et al.*, 1996, LEWIN, 1994). Estes mesmos autores também verificaram que as bactérias crescidas até a fase de crescimento exponencial (fase log) eram mais competentes do que aquelas isoladas de outros estágios do crescimento.

A eficiência de transformação é geralmente expressa como o número de células transformadas (geralmente de 10^5 a 10^7) obtido a partir de um micrograma de DNA plasmidial intacto (LEWIN, 1997; LEWIN, 1994). O tamanho e a conformação da molécula de DNA afetam o processo de transformação. Plasmídeos pequenos são mais facilmente incorporados pela célula bacteriana competente. DNA linear é pobremente incorporado pelo fato de talvez sofrer degradação pelas exonucleases presentes no espaço periplasmático.

I. 2.5.2 – Mecanismos de Captação do DNA

O mecanismo de captação da molécula de DNA pela bactéria competente ainda é desconhecido. Uma hipótese é que as moléculas do DNA passam através de canais

situados nas chamadas zonas de adesão, que são locais onde as membranas interna e externa da célula bacteriana se unem formando poros. Estes poros só estão presentes durante o crescimento bacteriano na fase log. Em condições naturais, a captação do DNA torna-se difícil, devido à repulsão eletrostática existente entre as cargas negativas da camada de fosfolípidos da membrana bacteriana com a carga negativa dos grupos fosfato da molécula de DNA. A Figura 6 ilustra esta hipótese para o mecanismo de captação do DNA por íons Ca^{2+} .

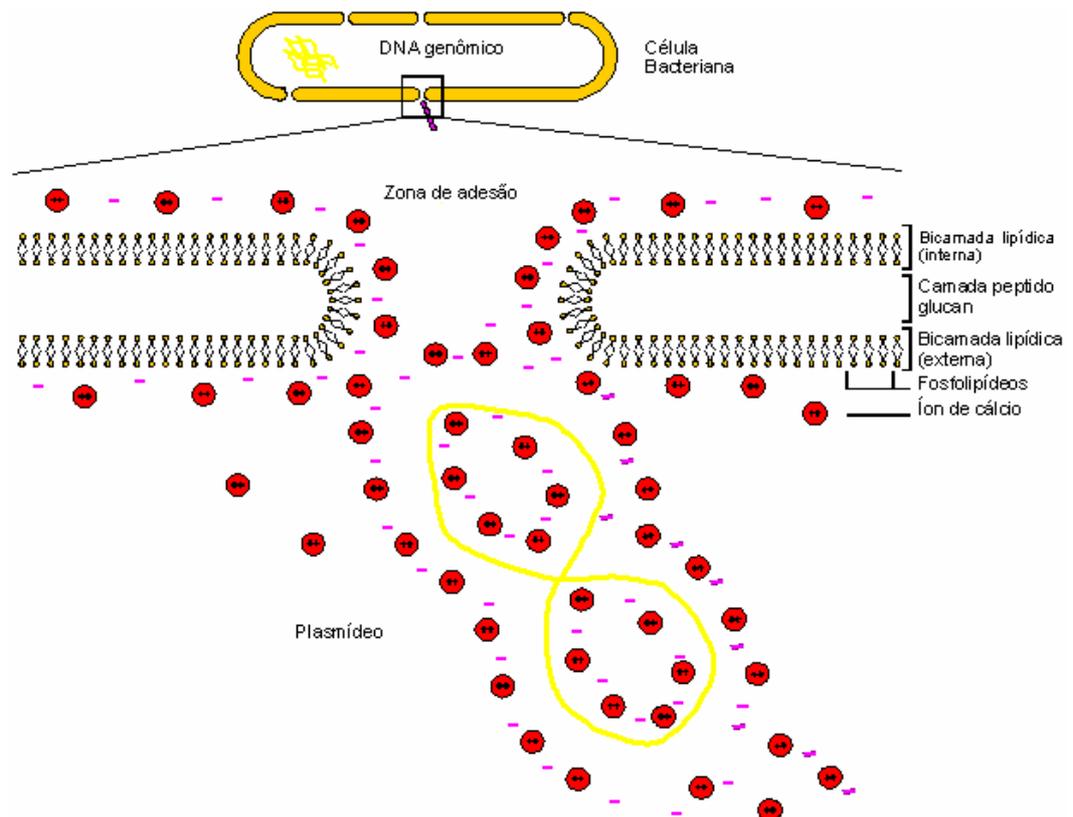


Figura 6. Mecanismo molecular proposto para a transformação de *E. coli* com uma molécula de DNA exógeno. Os círculos em vermelho representam os íons Ca^{2+} . Os sinais negativos em rosa são representações da carga elétrica negativa proveniente de grupos fosfatos carregados presentes nas membranas celulares das bactérias.

Quando soluções geladas de cloreto de cálcio são empregadas para incubar as células bacterianas, a fluidez da membrana celular é cristalizada, estabilizando a distribuição dos fosfatos carregados. Os íons Ca^{+2} formam um complexo com os fosfatos carregados, cobrindo as cargas negativas, facilitando agora a atração eletrostática com as moléculas do DNA na zona de adesão. O choque térmico complementa este processo de captação, provavelmente criando um desequilíbrio térmico entre o interior e o exterior da célula bacteriana, auxiliando o bombeamento do DNA através da zona de adesão.

I. 2.6 – Detecção e análise dos Clones Recombinantes

Um passo crucial na tecnologia do DNA recombinante é encontrar o clone correto na mistura de clones que foi criado durante os procedimentos de construção. Note que, no caso do vetor ser um plasmídeo, milhões de colônias de bactérias crescidas em placas de Petri contendo ágar deverão ser examinados para a detecção daquelas colônias que possuam o gene de interesse.

Dentre as marcas de seleção, a presença de genes repórteres é um dos artifícios mais importantes e frequentemente empregado na escolha de um vetor de clonagem. Genes repórteres funcionam como marcadores porque fornecem informações quanto ao sucesso ou não dos experimentos com DNA recombinante.

O gene *lacZ* é muito utilizado como gene repórter no qual produz uma enzima chamada β -galactosidase. A presença desta enzima faz com que a colônia da bactéria que tem em seu interior o plasmídeo contendo este gene seja azul quando colocada na presença dos compostos químicos isopropil-3-D-tiogalactopiranosídeo (IPTG) e 5-bromo-4-cloro-3-indoil-b-D-galactosidase (X-Gal). O IPTG induz a produção da

enzima β -galactosidase que degrada o substrato X-Gal formando um produto colorido, de cor azul.

Os sítios de restrição das enzimas utilizadas para cortar este tipo de plasmídeo se localizam exatamente na região onde está o gene que codifica a enzima β -galactosidase. Quando adicionamos um fragmento exógeno de DNA dentro deste plasmídeo, introduzindo-o nesta região, o gene que produz a enzima β -galactosidase é destruído. Sendo assim, as colônias de bactérias que possuem o plasmídeo com a essa enzima intacta em seu interior irão produzir colônias azuis, enquanto a colônia de bactéria que possuir o plasmídeo que teve a β -galactosidase destruída irá produzir colônias brancas.

Assim, os genes repórteres em adição aos antibióticos também permitem selecionar as colônias, porém selecionam somente aquelas constituídas por células que possuam o vetor ligado com o fragmento de interesse inserido, em detrimento das colônias que possuam o vetor de clonagem tal como ele se apresentava antes do início da manipulação.

I. 2.6.1 – Seqüenciamento de DNA

O seqüenciamento de DNA é um método de análise dos clones positivos que se faz necessário para se obter a seqüência completa, ou pelo menos, uma porção do gene alvo. Sua estratégia envolve a PCR consistindo na capacidade da DNA polimerase em adicionar um nucleotídeo na extremidade de alongamento de um oligonucleotídeo copiando uma fita de DNA molde.

Desde a década de 70, o método descrito por Sanger é o mais utilizado para análise de seqüências de DNA e é conhecido também como método da terminação da

cadeia (LEWIN, 1997; ZAHA *et al.*, 1996, LEWIN, 1994; STRYER, 1988). No início dos anos 70, o processo era manual e eram necessários dois anos para seqüenciar 20 pares de bases de DNA. Atualmente, com o desenvolvimento das técnicas de seqüenciamento automático, um aparelho de seqüenciamento é capaz de seqüenciar meio milhão de pares de bases por dia (LEWIN, 1997; ZAHA *et al.*, 1996, LEWIN, 1994; STRYER, 1988).

Estas duas versões da técnica diferem entre si basicamente pelos tipos de dideoxynucleotídeos (ddNTPs) empregados. No seqüenciamento manual, os ddNTPs são marcados com radiação; já no seqüenciamento automático são marcados com compostos fluorescentes. Na técnica manual se faz necessário uma autoradiografia para a leitura da seqüência do DNA e na técnica automática um equipamento provido de um laser é necessário para detectar os ddNTPs fluorescentes e permitir a leitura direta da seqüência.

I. 2.6.1.1 – Seqüenciamento Manual

O seqüenciamento de DNA descreve a clonagem dos fragmentos de DNA de interesse em vetores apropriados como os da série M13mp os quais têm sido largamente utilizados para esta finalidade. A reação de seqüenciamento pelo método de Sanger consiste na hibridização de um oligonucleotídeo sintético com uma região conhecida do vetor M13mp, imediatamente adjacente ao inserto a ser seqüenciado. A seguir, a síntese da fita complementar, pela enzima DNA polimerase, é promovida na presença de uma mistura de dideoxynucleotídeos (ddNTPs) marcados com ^{32}P ou ^{35}S e desoxirribonucleotídeos (dNTPs).

A Figura 7 ilustra uma representação do método manual de seqüenciamento de DNA.

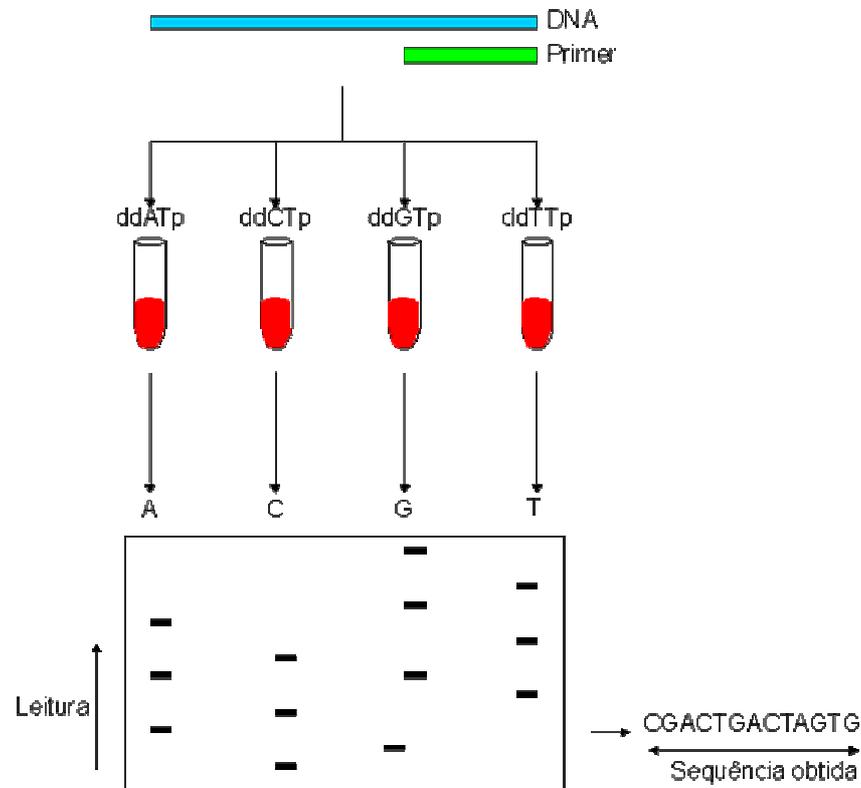


Figura 7. Seqüenciamento de DNA pelo método de Sanger. Os dideoxinucleotídeos representados como ddATP, ddCTP, ddGTP e ddTTP são os compostos 5'-desoxiadenosina trifosfato, 5'-desoxicitosina trifosfato, 5'-desoxiguanina trifosfato e 5'-desoxitimina trifosfato, respectivamente.

Neste processo são realizadas quatro reações independentes, sendo que em cada uma delas é adicionado um tipo de ddNTP. Em cada reação, um número muito grande de moléculas de fita complementar é produzido simultaneamente. A hidroxila 3' (3' OH) do oligonucleotídeo reage com um dNTP formando uma nova ligação fosfosdiéster. No entanto, em um dado momento da extensão, uma pequena porcentagem de moléculas irá incorporar um ddNTP e neste caso a reação de alongamento da cadeia será automaticamente interrompida naquele sítio, devido à ausência do grupamento 3' OH do ddNTP. Por outro lado, em outras moléculas a

reação de extensão continua a ocorrer até que um ddNTP seja incorporado. Sendo assim, em cada uma das quatro reações, haverá fragmentos de todos os tamanhos sendo que todos eles têm início comum, ou seja, adjacente ao iniciador.

Uma vez terminadas as reações de extensão, os fragmentos recém sintetizados são separados em gel desnaturante de poliacrilamida e em seguida autoradiografados. Como este tipo de gel apresenta alto poder de resolução é possível através da sua auto-radiografia, visualizar cada um dos fragmentos sintetizados como uma banda específica.

I. 2.6.1.2 – Seqüenciamento Automático

O desenvolvimento e a utilização de seqüenciadores automáticos tornaram mais rápido e eficiente o seqüenciamento de DNA na medida em que as etapas de leitura do gel e o processamento de seqüências são realizados através de computadores. Assim, o surgimento e a utilização de seqüenciadores automáticos foi o maior benefício alcançado com o desenvolvimento da técnica.

Em concordância com as técnicas manuais, a reação de seqüenciamento é também realizada através do método de Sanger. Entretanto, devido aos ddNTPs serem marcados com fluoróforos ao invés de isótopos radioativos, os fragmentos de DNA são detectados através de fluorescência induzida a laser. Quatro diferentes fluoróforos são empregados e uma vez excitados por um feixe de laser emitem luz em diferentes comprimentos de onda. É possível marcar com estes fluoróforos o *primer* universal M13mp ou então, cada um dos ddNTPs. Assim, uma vez que em cada uma das reações foi empregado um fluoróforo diferente é possível juntar estes produtos numa só reação e realizar a corrida em uma única raia do gel de seqüenciamento. Os

produtos da reação de seqüenciamento, marcados com os fluorocromos, ao serem submetidos à eletroforese passam pelo feixe de laser, que promove a excitação dos fluorocromos. A luz emitida pelos fluoróforos é detectada por um fotomultiplicador e a informação é processada através de um computador.

I. 3 – Expressão de Proteínas Heterólogas

Grande parte dos esforços despendidos na expressão de proteínas heterólogas, visando obtê-las em quantidade adequada para estudos funcionais e estruturais posteriores, se concentra na experimentação de sistemas de expressão alternativos, incluindo bactérias, leveduras, células de insetos e mamíferos; e no uso de proteínas de fusão como facilitadoras do enovelamento e purificação de proteínas recombinantes. A expressão destas proteínas permitiu uma abordagem poderosíssima que vem revolucionando os estudos de estrutura, função, purificação e identificação de novas proteínas.

O sistema mais comumente utilizado para expressão de proteínas heterólogas é o que utiliza *E. coli* como célula hospedeira (LEWIN, 1997; ZAHA *et al.*, 1996, LEWIN, 1994; STRYER, 1988). Este sistema é amplamente difundido devido à facilidade e baixo custo de se cultivar esta bactéria e pela reprodutibilidade e abundância de proteínas que a mesma produz. Além disso, modificações nos vetores e linhagens de *E. coli* são freqüentemente feitas no sentido de aumentar a eficiência e versatilidade do sistema original.

De maneira ideal, quando se pensa em expressão heteróloga, espera-se que a proteína de interesse seja estável, não tóxica para a bactéria, solúvel, expressa em grande quantidade e possa ser facilmente purificada. Um procedimento muito

utilizado é o de expressar a proteína de interesse em fusão com uma pequena sequência de aminoácidos específicos (“tag”) que permita sua fácil purificação através de cromatografia por afinidade em resinas às quais se encontram acoplada moléculas ligantes, permitindo que a *tag* possa se ligar especificamente. A coluna reterá a proteína de fusão até que os sítios ligantes sejam liberados pela mesma frente à passagem de compostos cujos grupamentos apresentem maior afinidade por estes sítios. Um exemplo do procedimento para a obtenção da proteína híbrida através da clonagem em vetores é dado na Figura 8.

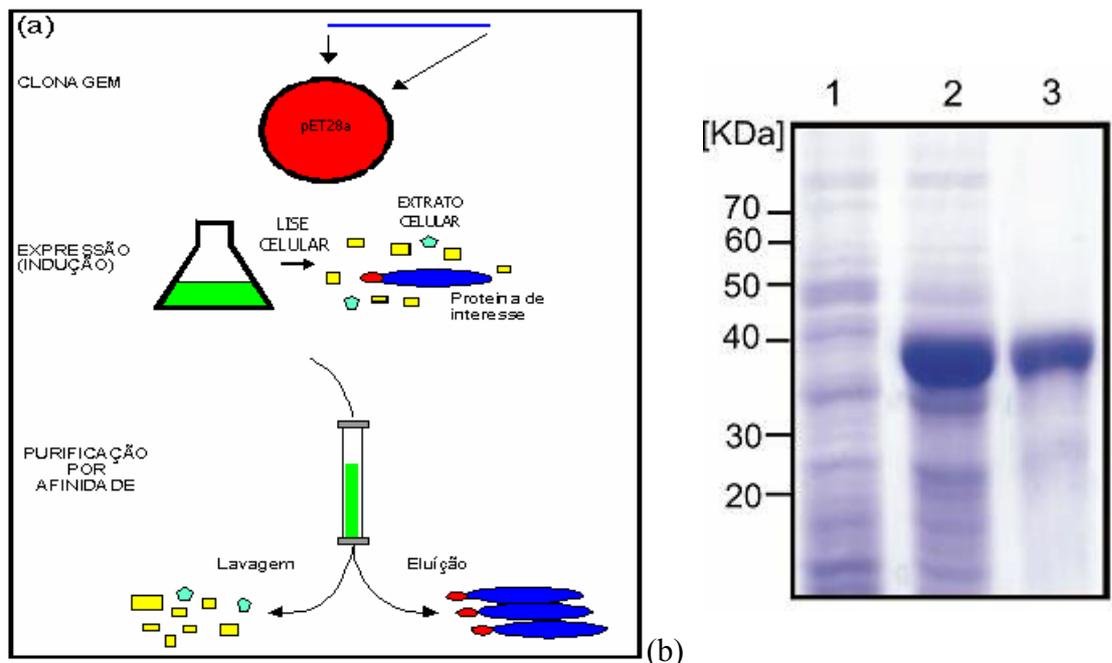


Figura 8. Expressão e purificação de uma proteína recombinante com *tag* de seis histidinas. a) Esquema ilustrando os passos da expressão em vetor pET28a e purificação da proteína recombinante; b) Gel de poliacrilamida de frações da expressão e purificação. (1) extratos de células não induzidas, (2) extratos de células induzidas e (3) fração eluída contendo a proteína recombinante purificada.

A produção de proteínas híbridas é geralmente muito simples, bem como, eficiente e de baixo custo financeiro podendo suprir imediatas necessidades da

pesquisa básica, tais como, produção de anticorpos, sondas em experimentos variados e para estudos funcionais e estruturais da proteína expressa. Permite, também, a expressão em grande escala, para fins industriais ou clínicos de enzimas, hormônios e anticorpos quando a atividade da proteína é preservada.

É importante mencionar que a situação ideal exposta acima nem sempre é atingida. Algumas proteínas produzidas em bactérias têm a sua atividade biológica plenamente recuperada, porém outras são inativas. Assim como, podem ser insolúveis, tóxicas para a célula hospedeira, expressas em baixos níveis, formar agregados extremamente insolúveis mesmo na presença de agentes desnaturantes, ou até mesmo interagir com a sua fusão inibindo-as de se ligar à resina de afinidade e tornando a purificação menos eficiente. Nestes casos, procedimentos para atenuar estes problemas devem ser bem estudados a fim de empregá-los convenientemente e não comprometer a pesquisa científica das proteínas de interesse.

Uma vez que a expressão de proteínas heterólogas provenientes da manipulação de genes específicos representa a última etapa da relevante tecnologia do DNA recombinante, o subsequente estudo bioquímico destas proteínas pode elucidar as diversas interações entre organismos vivos e patógenos. Visando focar isto e sabendo das possíveis similaridades entre enzimas de diferentes organismos em termos de função biológica, a literatura descrita a seguir servirá como base para o conhecimento do fitopatógeno *Xylella fastidiosa* e de algumas de suas proteínas metabólicas depositadas em seu respectivo banco de dados. A análise de enzimas de uma bactéria de crescimento lento como a *Xylella fastidiosa* representa não só um desafio, mas também uma contribuição dos estudos monitorados por conceituados pesquisadores envolvidos no Projeto Genoma/Onsa da FAPESP de 1998.

I. 4 – Bactéria *Xylella fastidiosa*

A bactéria *Xylella fastidiosa* (*X. fastidiosa*) (PURCELL, *et al.*; 1996) pertence ao grupo das bactérias gram-negativas, restrita aos vasos xilemáticos das plantas. Sua colonização provoca o bloqueio dos vasos obstruindo a passagem de água e nutrientes das raízes para a copa das plantas, gerando estresse híbrido e impedindo o desenvolvimento vegetal. É transmitida e disseminada para as plantas hospedeiras através de insetos vetores como a cigarrinha das famílias Cicadellidae e Cercopidae, que se alimentam da seiva bruta das plantas (MEIDANIS, 2002; HARTUNG, 1994). São estritamente aeróbicas e não fermentativas, não halofíticas, apigmentadas, com temperatura ótima de desenvolvimento de 26 a 28 °C e pH ótimo de 6,5 a 7,0. Possuem dimensões de 0,25 a 0,35 µm de largura por 0,9 a 5,0 µm de comprimento e apresentam ondulações na parede celular que podem estar envolvidas no reconhecimento do hospedeiro pelo patógeno (PURCELL, *et al.*; 1996)

A *X. fastidiosa* foi primeiramente relatada em 1987 nos Estados de São Paulo e Minas Gerais, afetando pomares com variedades de laranja doce. A clorose variegada dos citros (CVC), conhecida como amarelinho, é uma doença causada por esta bactéria e atinge todas as variedades de citros comerciais. A doença é amplamente distribuída em todas as regiões de produção de citros no Brasil, principalmente no estado de São Paulo, onde sua difusão ocorreu de forma extremamente rápida. As plantas cítricas afetadas pela CVC mostram inicialmente, na parte mediana e superior da copa, folhas amareladas com sintomas de deficiências nutricionais e frutos reduzidos, endurecidos, persistentes, com amarelecimento precoce e lesões pardas na casca, imprestáveis para o comércio. Além disso, danosas conseqüências aos

equipamentos de extração do suco são provocadas por estes frutos (PURCELL, *et al.*; 1996).

Numerosas espécies de plantas são colonizadas pela *X. fastidiosa* sem provocar sintomas. Freitag (PURCELL *et al.*, 1996) demonstrou que mais de 100 espécies de plantas arbóreas apresentam esta bactéria. Estas plantas não apenas suportam vastas populações de vetores, mas também foram excelentes fontes da bactéria, o que explica porque a remoção de videiras afetadas foi de limitado valor na prevenção da propagação da doença nos vinhedos da Califórnia e Flórida (PURCELL, *et al.*; 1996). Além dos citros, *X. fastidiosa* causa doenças em diversas plantas de importância econômica como cafeeiro (*Coffea arabica* L., PURCELL *et al.*, 1996; HARTUNG *et al.*, 1994), ameixeira (*Prunus salicina* Lindl. PURCELL *et al.*, 1996; HARTUNG *et al.*, 1994), videira (*Vitis vinifera* L.), hibisco (*Hibiscus schizopetalus*, PURCELL *et al.*, 1996; HARTUNG *et al.*, 1994), seringueira (*Hevea brasiliensis*, PURCELL *et al.*, 1996; HARTUNG *et al.*, 1994) e em várias espécies florestais (AZEVEDO *et al.*, 2000). Fica evidente que a gama de hospedeiros deste fitopatógeno é grande e diversa.

A interação patógeno-hospedeiro pode ser analisada sob aspectos morfológicos, genéticos, bioquímicos ou moleculares, isoladamente ou em conjunto. Independente da abordagem é clara a importância dos fatores que delimitam a colonização do tecido hospedeiro por estas bactérias e como eles vivem neste ambiente. Uma grande parcela destes fatores é formada pelas proteínas produzidas por elas onde, as enzimas exercem funções cruciais, dependente de cada interação no desenvolvimento do patógeno e, conseqüentemente, no surgimento dos sintomas no hospedeiro. Desta forma, a descrição de alguns aspectos interessantes sobre função, classificação e uso destes tipos de enzimas se faz necessário para caracterizá-las adequadamente.

I. 4.1 - Proteínas de *Xylella fastidiosa*

A determinação experimental das seqüências primárias seguidas de suas subseqüentes comparações com seqüências protéicas depositadas no banco de dados público SWISSPROT (Lausanne, Suíça), permitiram durante o projeto genoma caracterizar diversas proteínas como sendo de prováveis funções conhecidas. Dentre elas, três proteínas apresentando homologias com a de outros organismos foram denominadas prováveis enzimas hidroxinitrila liase (HNL), corismato sintase (CS) e proteínas D da via de transporte e secreção Tat (TatD), potencialmente envolvidas na interação planta-patógeno.

I. 4.1.1 - Hidroxinitrila Liase

A habilidade de organismos vivos em produzir ácido cianídrico (HCN) é denominada cianogênese. Apesar de muito estudada em plantas, pode ser encontrada também em bactérias, fungos e insetos.

Fontes de HCN foram identificadas em aproximadamente 300 espécies de plantas como sendo glicosídeos cianogênicos e cianolipídeos denominados precursores cianogênicos e que, em sua maioria, são derivados de cinco aminoácidos: a L-fenilalanina, L-tirosina, L-valina, L-isoleucina e L-leucina. Os precursores mais comuns são as cianohidrininas (definidas quimicamente como glicosídeos de α - hidroxinitrilas). O HCN não é produzido livremente na maioria das plantas, porém é liberado por precursores cianogênicos como resultado de uma ação enzimática. A Figura 9 ilustra o mecanismo enzimático de precursores cianogênicos até a liberação de ácido cianídrico.

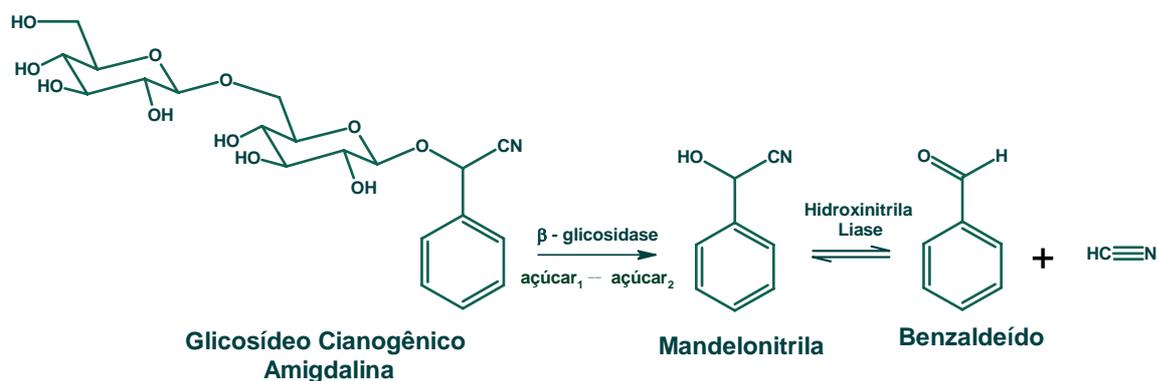


Figura 9. Mecanismo enzimático de glicosídeos cianogênicos. O grupo açúcar em um glicosídeo cianogênico é normalmente uma D-glucose.

Quando a integridade celular dos tecidos de plantas é afetada por ataque de fungos, herbívoros ou ainda mecânico, os chamados precursores cianogênicos (glicosídeos e cianolípídeos) entram em contato com enzimas β - glicosidases e são hidrolisados e dissociados a cianohidrinas. As cianohidrinas, por sua vez, são clivadas a HCN e ao correspondente aldeído ou cetona, sob ação de outra enzima catabólica chamada α - hidroxinitrila liase (HNL).

As HNLs estão naturalmente envolvidas no passo final da biodegradação de glicosídeos cianogênicos nas células de diversas espécies de plantas de diferentes origens filogenéticas. Como todas as reações enzimáticas envolvidas na cianogênese são reversíveis, as HNLs podem ser utilizadas também na formação de cianohidrinas.

As HNLs de diferentes famílias de plantas representam uma nova classe de enzimas que originam da evolução independente de diferentes precursores de proteínas. De fato, a falta de identidade seqüencial associada à seqüência de aminoácidos de recentes HNLs clonadas enfatiza esta idéia, embora algumas

apresentem nenhuma ou limitada identidade seqüencial a outras proteínas conhecidas. Um exemplo disto é a HNL de *Sorghum bicolor* que possui identidade seqüencial surpreendente com uma enzima serina carboxipeptidase, especialmente com aquelas de trigo. Alguns sítios de importância funcional nas carboxipeptidases são conservados, como exemplo, motivos de N-glicosilação, resíduos da tríade catalítica (sítio ativo) e de cisteína.

As HNLs podem ser classificadas de acordo com diferentes propriedades enzimáticas tais como especificidade a substratos (R) ou (S) e independência ou não de coenzimas (FAD - flavina-adenina dinucleotídeo), representando a classificação bioquímica dessas enzimas.

I. 4.1.1.1 - Enzimas dependentes de FAD

Enzimas HNLs que possuem grupos prostéticos FAD (moléculas não protéicas) ligados próximo ao seu sítio catalítico são denominados hidroxinitrila liases dependentes de FAD. Estas enzimas foram isoladas de monocotiledôneas da espécie *Rosaceae* nas subfamílias *Prunoideae* e *Maloideae*.

Elas são glicoproteínas de cadeia longa com massa molecular variando de 50 a 80 kDa, contendo mais de 30% de carboidratos ligados à mesma por ligação N-glicosídica. Hidroxinitrila liases dependentes de FAD encontradas no gênero *Prunus amygdalus* são formadas por uma única subunidade e existem em pequenas isoformas diferentes, originadas de variações na sua estrutura primária e/ou do seu conteúdo carboidrato heterogêneo.

Apresentam pontos isoelétricos variando de 4,2 a 4,8 e possuem como substrato natural a (R)-(+)-mandelonitrila que é derivada dos glicosídeos cianogênicos (R)-

prunasina ou amigdalina. Sua forma ativa contém FAD ligado não covalentemente próximo ao seu sítio ativo, apresentando um espectro de absorção típico com máximos na região 380-390 e 460-480 nm. Estas enzimas estão presentes em concentração relativamente alta nos tecidos de grãos e sementes, necessitando de 5 a 10 purificações para alcançar alto grau de homogeneidade. A região codificadora de proteínas (ORF – “open reading frame”) em *Prunus amygdalus* codifica polipeptídeos de 61 kDa (563 aa) que é preservada pela remoção de uma seqüência sinal N-terminal de 27 aminoácidos (aa), levando a uma hidroxinitrila liase madura de 58 kDa (sem carboidrato) (ASANO *et al.*, 2005).

I. 4.1.1.2 - Enzimas independentes de FAD

As hidroxinitrila liases independente de FAD são mais heterogêneas e foram isoladas de diferentes famílias de grupos taxonômicos de monocotiledôneas e dicotiledôneas de Poaceae, Euphorbiaceae, Linaceae, Filitaceae e Olacaceae. Elas diferem entre si no conteúdo de carboidratos, em massa molecular, na estrutura primária, bem como em suas subunidades de polimerização. Além disso, elas não estão relacionadas às HNLs dependentes de FAD a nível de seqüência de aminoácidos e de sua superfície estrutural antigênica. Em geral, as HNLs deste grupo possuem caráter mais ácido com pontos isoelétricos variando de 3,9 a 4,6 e necessitam de 100 a 1500 purificações para se tornar homogênea (ASANO *et al.*, 2005; FECHTER *et al.*, 2004; HICKEL *et al.*, 1996; WAJANT *et al.*, 1996).

Dentre as HNLs pertencentes a este grupo as mais estudadas são de *Sorghum bicolor*, *Manihot esculenta* e *Hevea brasiliensis*. HNLs da família Poaceae, isolada da espécie *Sorghum bicolor* apresentam massa molecular de 105 kDa, do qual, 7% é

carboidrato N-ligado. São dímeros que apresentam conformação heterotetramérica essencial para sua atividade enzimática. In vivo, apresentam conformações homo ou heterodíméricas. Possuem como substrato natural a p-hidroxi-(S)-mandelonitrila que é derivada do glicosídeo cianogênico (S)-dhurrin. Já as HNLs da família Euphorbiaceae, isoladas das espécies *Manihot esculenta* e *Hevea brasiliensis* são proteínas não glicosiladas com uma subunidade de massa molecular 30 kDa e conformação natural homoligomérica; apresentando identidade de 75% da sua estrutura primária com outras proteínas já identificadas da mesma família. Seus substratos naturais são acetato de cianohidrina e cianohidrina 2-butanona (descrita somente para a HNLs isolada de *Manihot esculenta*) derivadas dos glicosídeos cianogênicos linamarina e (R)-lotaustralina, respectivamente.

Baseado nas características bioquímicas das diferentes hidroxinitrila liases já isoladas por diversos pesquisadores, observou-se ainda que, para garantir o mecanismo enzimático ou a atividade enzimática das HNLs dependentes de FAD, se faz necessário a presença de uma serina, uma cisteína e um cofator FAD no meio reacional; já para HNLs independentes de FAD, estruturas α,β -hidrolases com tríades catalíticas (formada por três aminoácidos Ser/Asp/His sendo a serina nucleofílica) são enzimaticamente essenciais. A função dos grupos prostéticos FAD não foi ainda bem esclarecida, porém, alguns estudos sugerem que eles são essenciais para manter a estrutura conformacional total da molécula ativa.

A mais relevante aplicação das enzimas HNL está na síntese orgânica enantioseletiva de compostos ativos utilizados extensivamente nas indústrias farmacêutica, química e de cosmético. As HNLs são utilizadas como biocatalisadores industriais para a síntese de cianohidrinhas quirais através da reação enzimática reversa que elas catalizam a partir de HCN e aldeídos ou cetonas. As cianohidrinhas quirais são

compostos essenciais necessários na obtenção dos ativos empregados na fabricação de antibióticos, broncodilatadores, inseticidas, germicidas e cosméticos em geral (SHARMA *et al.*, 2005; ASANO *et al.*, 2005; FECHTER *et al.*, 2004; HICKEL *et al.*, 1996; WAJANT *et al.*, 1996).

I. 4.1.2 – Corismato Sintase

Em plantas e microrganismos, todos os compostos aromáticos chave envolvidos no mecanismo primário, incluindo os três aminoácidos aromáticos encontrados em proteínas, são produzidos pela via do chiquimato ou do ácido chiquímico. A via do chiquimato é essencial em algas, plantas superiores, bactérias e fungos, porém ausente em mamíferos, porque estes últimos dependem destes compostos para sua dieta.

A bactéria *X. fastidiosa* é capaz de sintetizar uma ampla variedade de cofatores enzimáticos e grupos prostéticos, alguns deles necessários e essenciais à via do chiquimato, incluindo biotina, ácido fólico, pantotenato e coenzima A, ubiquinona, glutationa, tireodoxina, glutaredoxina, riboflavina, mononucleotídeo flavina (FMN), dinucleotídeo flavina-adenina (FAD), nucleotídeos pirimidínicos, porfirina, tiamina, 5-fosfato piridoxal (PLP) e lipoato.

A via metabólica do chiquimato consiste de sete enzimas que catalisam a conversão seqüencial de eritrose-4-fosfato e fosfoenolpiruvato à corismato (KITZING *et al.*, 2005; HERRMANN *et al.*, 1999; MACHEROUX *et al.*; SPRINSON *et al.*, 1964). A Figura 10 ilustra a via do chiquimato completa.

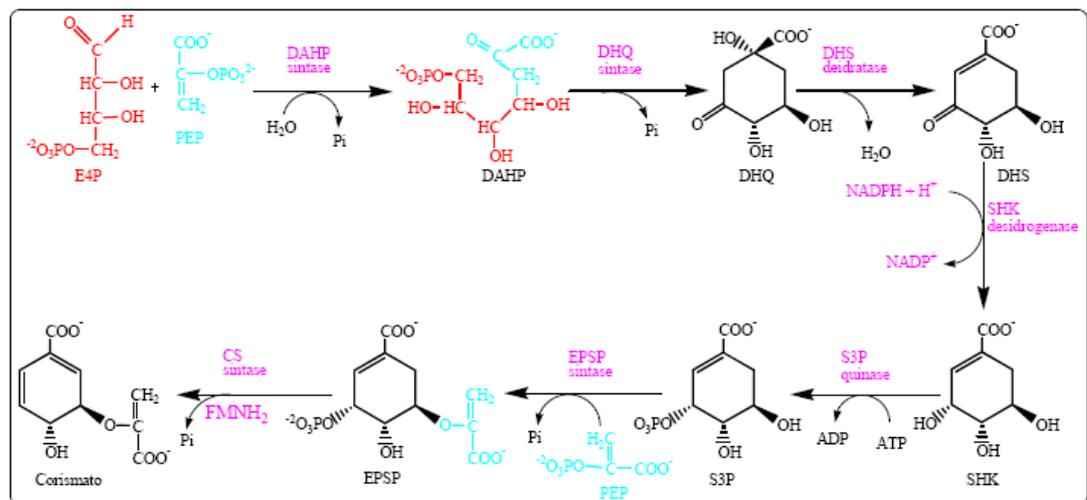


Figura 10. Via do chiquimato. As enzimas descritas tratam-se das enzimas 2-ceto-3-desoxi D-arabinoeptulose-7-fosfato sintase (DAHP sintase), desidroquinato sintase (DHQ sintase), 3-desidroquinato desidratase (DHS desidratase), chiquimato desidrogenase (SHK desidrogenase), chiquimato sintase (S3P quinase), 5-enolpiruvil chiquimato-3-fosfato sintase (EPSP sintase) e corismato sintase (CS sintase).

Todos os intermediários da via podem ser considerados compostos importantes porque podem servir como substratos para outros caminhos metabólicos. A organização molecular e estrutural das enzimas da via do chiquimato varia consideravelmente entre grupos de microorganismos. As bactérias possuem sete polipeptídeos individuais, cada um possuindo uma única atividade enzimática, nos quais são codificados por genes separados. Plantas têm um arranjo molecular similar ao de bactéria, isto é, as enzimas são codificadas por genes separados, com exceção da DHQ e a DHS, que são domínios separados de um polipeptídeo bifuncional. Grande parte das enzimas de plantas, embora codificadas no núcleo, é ativa no cloroplasto e possuem uma seqüência N-terminal transitória. Em contraste, todos os fungos possuem corismato sintases monofuncionais e um polipeptídeo pentafuncional chamado AROM. O polipeptídeo AROM tem domínios análogos às enzimas bacteriais: DHQ, EPSP, S3P, DHQ e SHK.

O corismato é convertido, por cinco enzimas distintas, aos metabólitos intermediários prefenato, antranilato, aminodeoxicorismato, isocorismato e p-hidroxibenzoato (KITZING *et al.*, 2005; HERRMANN *et al.*, 1999; MACHEROUX *et al.*; SPRINSON *et al.*, 1964). A Figura 11 apresenta a rota para obtenção destes metabólitos intermediários e consequentemente os aminoácidos, ácidos e vitaminas.

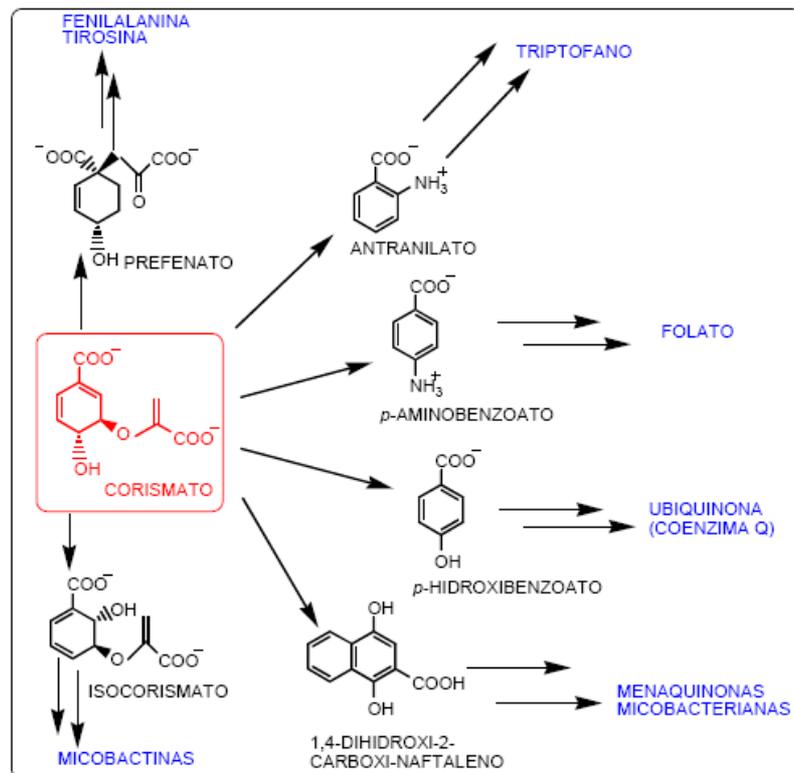


Figura 11. Destino do corismato a partir da via do ácido chiquímico e os produtos finais correspondentes. Os metabólitos intermediários são essenciais para a viabilidade do organismo.

Estes metabólitos compreendem os primeiros intermediários na biossíntese da fenilalanina (Phe), Tirosina (Tyr), triptofano (Trp), folato, menaquinona e os sideróforos enterobactina e ubiquinona, respectivamente. A síntese destes precursores é em muitos casos altamente regulada. Entre os muito metabólitos aromáticos

secundários estão os flavonóides, muitas fitoalexinas, indol acetato, alcalóides (como morfina), protetores de luz UV e o mais importante, lignina.

I. 4.1.2.1 – Biosíntese de Aminoácidos Aromáticos

Na biossíntese dos aminoácidos aromáticos, Phe, Tyr e Trp, os precursores essenciais iniciais são o fosfoenolpiruvato e a eritrose-4-fosfato. Sua condensação forma o 2-ceto-3-desoxi-D-arabinoheptulose-7-fosfato, um composto de sete carbonos que se torna cíclico e é convertido a chiquimato que dará origem ao corismato, o precursor comum dos aminoácidos aromáticos. O corismato é transformado em prefenato por meio de uma mutase. O prefenato sofre uma desidratação e depois é descarboxilado originando o fenilpiruvato que é transaminado formando a fenilalanina. O prefenato pode, ainda, sofrer uma descarboxilação oxidativa formando p-hidroxi fenilpiruvato. O p-hidroxi fenilpiruvato é então transaminado formando a tirosina.

Os mamíferos sintetizam a tirosina através da hidroxilação da fenilalanina. A glutamina cede uma amina de sua cadeia lateral para o corismato, esta reação libera piruvato e forma o antranilato. O antranilato se condensa com o PRPP (fosforribosil pirofosfato) formando o N-5-fosforribosil antranilato. Este composto sofre então um rearranjo, uma desidratação e uma descarboxilação gerando o indol 3-glicerol fosfato. O indol 3-glicerol fosfato reage com a serina formando o triptofano e liberando gliceraldeído 3-fosfato. A triptofano sintase é a enzima que catalisa as duas reações finais da biossíntese do triptofano. Sua estrutura é um tetrâmero $\alpha_2\beta_2$. A subunidade α cliva o indol-3-glicerol fosfato em indol e gliceraldeído-3-fosfato. A subunidade β une o indol com a serina para formar o triptofano. O indol é uma molécula hidrófoba

que pode atravessar a membrana e ser perdido da célula. A triptofano sintase possui um canal que liga suas subunidades catalíticas permitindo que o indol seja difundido através da enzima, ou seja, o indol formado não sai da enzima quando a serina está presente e as duas reações parciais são coordenadas. Este processo é chamado de canalização do substrato e é extremamente vantajoso, pois assim, além de aumentar mais de 10 vezes a taxa catalítica global, ele previne a perda de um intermediário da síntese do triptofano.

Enzimas da via do chiquimato são alvos atrativos para o desenvolvimento de drogas como as contra a tuberculose, no entanto, alguns novos inibidores para as enzimas da via do chiquimato têm alta especificidade para enzimas de *X. fastidiosa*, uma vez que elas também estão presentes em plantas.

I. 4.1.3 – Proteína D do Sistema de Transporte e Secreção Tat

Os sistemas de transporte e secreção são encontrados em várias bactérias, fitopatógenos, patógenos humanos e microrganismos simbiotes. Como o próprio nome diz, sistemas de transporte e secreção exercem a função de transportar para fora da célula (meio extracelular) através das membranas, diversas proteínas sintetizadas em seu interior cujas funções são extremamente específicas.

Fisiologicamente, algumas bactérias classificadas como gram-positivas possuem apenas um espaço intracelular denominado citoplasma sendo limitado por uma única membrana. Já outras, as gram-negativas, possuem o periplasma além do citoplasma, localizado entre duas membranas, uma interna e outra externa. Aqui residem as diferenças na exportação de proteínas através de membranas. Algumas proteínas exportadas das bactérias gram-negativas permanecem no periplasma, mas outras

atravessam as duas membranas, sendo secretadas para fora das células. A bactéria gram-negativa *E. coli*, em particular, não faz bem a secreção para fora da célula, e muitas proteínas que são exportadas em outras espécies gram-negativas, quando tem o gene inserido em *E. coli*, são transportadas no máximo até o espaço periplasmático, mas não além dessa região.

O mecanismo de transporte e secreção de proteínas em bactérias gram-negativas é nas últimas duas décadas, um campo de grande pesquisa com o intuito de verificar a virulência destes organismos e a produção de enzimas de interesse (FILLOUX *et al.*, 1998; SARGENT *et al.*, 1998). Estas bactérias são capazes de secretar enzimas degradativas e toxinas diretamente envolvidas nos processos infecciosos para o meio ambiente ou diretamente para o citoplasma da célula hospedeira/alvo. A produção destas enzimas está envolvida na aquisição de substratos, degradação da parede celular das plantas ou tecido animal e no efeito tóxico de alguns compostos ocasionando a morte de células alvos.

A maioria das bactérias gram-negativas apresenta um mecanismo comum de secreção via periplasma (FILLOUX *et al.*; 1998; SARGENT *et al.*, 1998). Esse mecanismo envolve um processo com dois passos: primeiro, as proteínas são transferidas através da membrana citoplasmática, via um mecanismo dependente de uma seqüência sinal. O segundo passo é a transferência através da membrana externa, um processo que envolve funções que podem ser específicas para proteínas individuais ou em grupos (FILLOUX *et al.*; 1998; SARGENT *et al.*, 1998).

Um total de 140 genes que codificam proteínas relacionadas com o transporte foi identificado no genoma de *X. fastidiosa*, representando 4,8 % de todos os genes codificando para proteínas e estão incluídos em dois sistemas de transporte: sistemas gerais de secreção ou Sec dependente e sistema Tat (FILLOUX *et al.*; 1998;

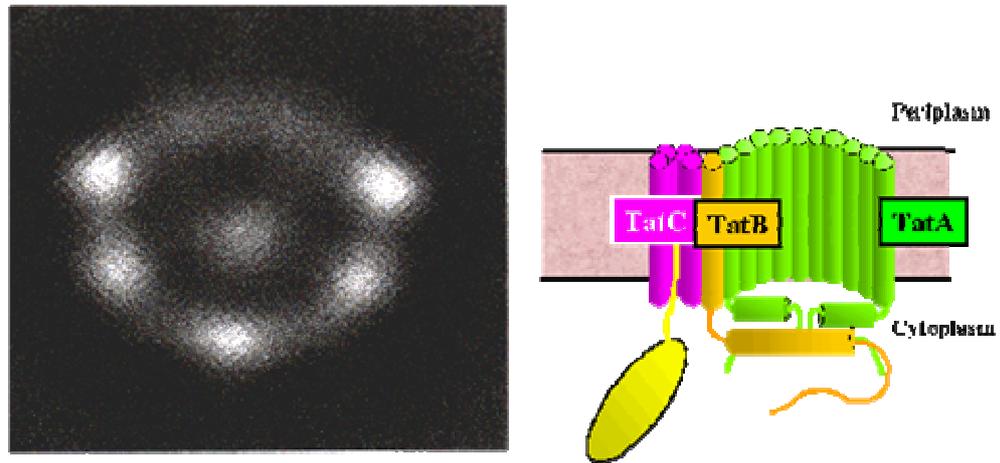
SARGENT *et al.*, 1998). Estes genes referem-se a transportes de ânions, cátions, carboidratos, peptídeos, proteínas e de substâncias relacionadas com as vias de secreção.

A maquinaria de sistemas gerais de secreção envolve a translocação de proteínas em conformação estendida nas quais são freqüentemente ligadas a outras proteínas que previnem seu enovelamento antes de serem exportadas. Proteínas como fotossintéticas e de cadeias respiratórias alternativas (anaeróbicas, por ex.) produzidas dentro das células e que estão envolvidas no metabolismo energético, necessitam de ser translocadas associadas à cofatores. Essas proteínas já adquirem uma configuração terciária ao se ligarem aos cofatores no citoplasma, e conseqüentemente necessitam de um canal especial na célula, que é designado por sistema Tat (“Twin Arginine translocation”). Dentre os cofatores possíveis, incluem-se íons cobre, enxofre, ferro, níquel e nicotinamida adenina dinucleotídeo-fosfato (NADP). A maioria das bactérias possui esta via de secreção e é totalmente distinta das outras vias de secreção (WEINER *et al.*, 1998; SARGENT *et al.*, 1998).

Proteínas da via Tat recebem este nome devido seus precursores serem lançados para a via através de um peptídeo sinal localizado na zona N-terminal da proteína reconhecível por uma seqüência conservada (S/T)RRxFLK que inclui dois resíduos consecutivos de arginina (R). Os resíduos de arginina são invariantes em proteínas desta via sendo que os outros resíduos ocorrem com freqüência de 50% nas proteínas já descritas. O que mais distingue estes peptídeos sinal dos convencionais é que estes possuem muitas vezes cargas positivas nas regiões hidrofóbicas e regiões N-terminais mais compridas do que o normal. Mas, a característica mais pronunciada da via Tat é que ela funciona para transportar proteínas enoveladas de dimensões variadas através da membrana citoplasmática.

A Figura 12 mostra dois arranjos propostos para proteínas pertencentes ao sistema Tat.

A)



B)

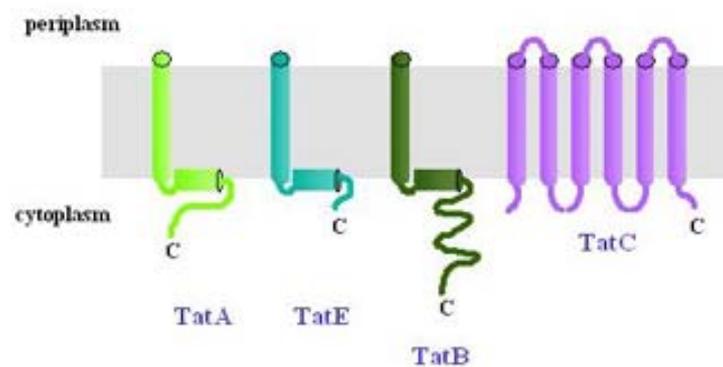


Figura 12. Previsões topológicas derivadas das seqüências previstas para os genes do sistema Tat.

A) Imagem da microscopia eletrônica e modelo do complexo TatABC purificados da bactéria *Streptomyces coelicolor* (Sargent *et. al.*, 1998). B) Modelo para as proteínas de translocação Tat de *E. coli*. Em ambos os organismos, somente as proteínas Tat C apresentam suas extremidades N- e C-terminais localizados no citoplasma.

As proteínas destes sistemas são classificadas como proteínas transportadoras do tipo ABCDE. São comumente denominadas TatA, TatB, TatC, TatD e TatE. O canal de passagem das proteínas a serem excretadas é majoritariamente composto por TatA. Transportadores do tipo ABE (“ATP-binding cassette”), em sua maioria, estão geralmente reunidos na formação de um único canal e envolvidos no transporte transmembranar de substratos com consumo de ATP. Os transportadores do tipo C estão envolvidos no transporte de proteínas dependentes do pH. Já transportadores do tipo D existem, mas não se sabe qual ou quais são suas funções até o momento. Somente sabe-se que eles são proteínas citoplasmáticas e não estão fazendo o papel de interação direta com a membrana.

Conceitualmente, o transporte dependente de sistemas Sec é mais rápido ou de menor custo energeticamente para a célula do que sistemas Tat. Porém, a preferência da bactéria por um sistema ou outro depende das velocidades relativas e da conformação das proteínas que necessitam ser secretadas.

I. 5 – Objetivos e Perspectivas

Através da análise por microarranjos de DNA foram observados 23 genes diferencialmente expressos da *X. fastidiosa* quando presente em plantas de citros com sintomas de CVC comparado com plantas de citros sadias. Desses, três genes estão envolvidos na patogenicidade e distribuídos segundo as diferentes categorias funcionais especificadas durante a anotação do Projeto Genoma da *X. fastidiosa* (<http://unicamp.lbi.ic.unicamp.br/xf>). É importante mencionar que as funções atribuídas a estes genes foram baseadas apenas por identidade seqüencial com proteínas de funções conhecidas de outros organismos.

Com base no projeto genoma e na literatura (SIMPSON *et al.*, 2000), os genes AroC e HNL depositados no banco de dados desta bactéria apresentam fortes indícios de que são importantes sustentadores da sua colonização, adaptação e patogenicidade no hospedeiro. Além disso, por terem sido expressos em grande quantidade pela bactéria em diversas condições são também possíveis grandes alvos biotecnológicos podendo ser aplicados na produção de pesticidas, cosméticos e fármacos.

Entre as proteínas hipotéticas conservadas, a HNL especificamente identificada como potencialmente envolvida num processo de detoxicação contra o HCN liberado pela planta será especialmente caracterizada e priorizada em termos estruturais a fim de permitir sua manipulação futura no desenvolvimento de possíveis moléculas inibidoras ou silenciadores do gene que a codifica.

Com as HNLs envolvidas em processos de liberação de toxinas e sendo possíveis alvos diretos de enzimas das vias de secreção e transporte, a similaridade de uma terceira proteína, também expressa diferencialmente com enzimas do tipo D da via Tat, fomentou a investigação desta, principalmente por não haver na literatura trabalhos que estudem a função deste tipo de proteína.

Acreditamos que a presença de genes com alta similaridade às enzimas liases, corismato e de transporte e secreção conhecidas de outros organismos, fornecem material enzimático adequado que contribui muito para seu metabolismo, permanência e conseqüente degradação dos componentes celulares na planta hospedeira. O conjunto das potencialidades enzimáticas previstas para estas proteínas estimulou, em nosso grupo, a pesquisa com a *X. fastidiosa*, principalmente porque poderá contribuir na aplicabilidade individual de cada uma na área biotecnológica. Assim, os objetivos gerais do presente trabalho foram: clonar, expressar, purificar e

elucidar as propriedades físico-químicas e estruturais das enzimas de *X. fastidiosa* obtidas através da propagação em *E. coli*.

Neste intuito, as seguintes etapas foram desenvolvidas no referente trabalho:

1- Expressão e purificação das proteínas de interesse. Obtenção das proteínas de interesse a partir das principais ferramentas da biologia molecular (extração, quantificação e reações de amplificação do DNA genômico, purificação dos fragmentos, clonagem em vetores de expressão, seleção dos clones, expressão do produto de PCR, verificação da expressão), bem como da bioinformática (construção de primers, alinhamento e caracterização da seqüência protéica obtida frente a proteínas já descritas na literatura/banco de dados), que servirão como ferramenta essencial no esclarecimento das etapas envolvidas na identificação das proteínas e na elucidação do funcionamento e da interligação das diversas vias metabólicas.

2- Caracterização das proteínas de interesse. Esta etapa envolve: separação, purificação, e determinação do peso molecular das proteínas por cromatografia de exclusão por tamanho, cromatografia de afinidade e SDS-PAGE; determinação do pI por focalização isoeétrica; medida de concentração através da espectrofotometria por absorção UV-VIS; caracterização das diferenças estruturais e identificação de alterações conformacionais por fluorescência, dicroísmo circular e SAXS; análise das seqüências através espectrometria de massas e, finalmente, ensaios enzimáticos a fim de determinar a atividade enzimática das proteínas em estudo.

A biocaracterização, ou caracterização bioanalítica destas proteínas recombinantes, nunca antes isolada em *Xylella fastidiosa*, poderá ser útil para a compreensão dos mecanismos de ação destas proteínas quando a bactéria está presente em plantas hospedeiras.

II – MATERIAIS E MÉTODOS

II. 1 – MATERIAIS

II. 1.1 – Reagentes, Enzimas, Linhagens Bacterianas e Vetores

Os reagentes químicos utilizados tais como Tris-HCl, β -mercaptoetanol, SDS, NaCl, MgCl₂, acrilamida, Na₂HPO₄, NaOH, HCl, EDTA, brometo de N-cetil-N,N,N-trimetil amônio (CTAB), ditioneitol (DTT), iodoacetamida, acetonitrila, ácido fórmico, ácido trifluoracético (TFA), glicerol, uréia, ácido acético glacial, clorofórmio, álcool isoamílico, isopropanol, fenol, ácido bórico, acetato de sódio, borato de sódio, etanol, bicarbonato de amônio, *Comassie Brilliant Blue G-250 e R-250*, albumina sérica bovina (BSA – 66 kDa), ovalbumina (OVA – 45 kDa), anidrase carbônica (CA – 30 kDa), inibidor de tripsina de soja (ITS - 20,1 kDa), citocromo C (cit C - 12,4 kDa), α -lactoalbumina (a-Lacto), extrato de levedura, bacto-triptona, ágar e mandelonitrila foram de grau P.A. ou equivalentes e comprados da Sigma–Aldrich Chemical (St. Louis, MO, EUA). O reagente químico 5-bromo-4-chloro-3-indolyl-beta-D-galactopiranosideo (X-Gal) foi obtido da GIBCO BRL (Rockville, MD, EUA).

Os reagentes para PCR e a polimerase *Taq Platinum High Fidelity* foram obtidas da Invitrogen Life Technologies (Carlsbad, CA, EUA). As enzimas *Thermophilic DNA polymerase*, T4 DNA ligase, *XhoI* e *NdeI*, seus respectivos tampões e tripsina modificada grau seqüenciamento foram obtidas da Promega (Madison, WI, EUA). Os oligonucleotídeos específicos (iniciadores ou *primers*) foram obtidos da Integrated DNA Technologies (Coralville, IA, EUA).

As linhagens de *Escherichia coli* (*E. coli*) utilizadas nos experimentos de clonagem e expressão foram DH5 α e BL21(DE3) obtidos da Promega e Novagen (San Diego, CA, EUA), respectivamente. Os vetores de propagação e expressão utilizados foram pGEM[®]-T *easy* e pET28a obtidos da Promega e Novagen, respectivamente. Os kits *Wizard[®] Plus SV Miniprep DNA Purification Systems* e *Wizard[®] SV Gel and PCR Clean-Up System* foram obtidos da Promega. A Resina *Ni-NTA superflow* foi obtida da Qiagen (Hilden, Alemanha). Os antibióticos ampicilina, canamicina e o indutor isopropil- β -D-tiogalactosidase (IPTG) foram comprados da Calbiochem (La Jolla, CA, EUA).

A coluna cromatográfica de exclusão por tamanho, Superdex 75-HR 10/30 foi comprada da GE Healthcare (Uppsala, Suécia).

II. 2 – MÉTODOS

Na descrição deste trabalho, as seqüências nucleotídicas provenientes de *X. fastidiosa* serão referidas como ORFs (*open reading frame*). As proteínas resultantes da tradução das ORFs de *X. fastidiosa* e expressas de forma heteróloga são tratadas como proteínas recombinantes e são designadas com a sigla r antes do nome.

II. 2.1 – Extração do DNA genômico

O método utilizado para a extração do DNA genômico foi o método descrito por Ausubel e colaboradores em 1987 (AUSUBEL *et al.*, 1995), com algumas modificações.

As suspensões bacterianas de *X. fastidiosa* foram transferidas para tubos de microcentrifuga estéreis (*eppendorfs*) de 2 mL contendo 100 μ L de água purificada e estéril (Milli-Q, Millipore, Bedford, MA, EUA) e centrifugadas em uma microcentrifuga *Eppendorf* (modelo 5415 D, Cologne, Alemanha) por 20 min a 14000 rpm. Os sobrenadantes foram descartados e os precipitados (*pellets*) ressuspensos em uma solução contendo 400 μ L tampão TE 10 mM (10 mM de Tris e 1 mM EDTA, ajustado com HCl concentrado a pH 8,0), 70 μ L de SDS a 10% e 10 μ L de proteinase K a 20 mg mL⁻¹. As soluções foram incubadas a 37 °C por uma hora. Em seguida, 10 μ L de RNase a 200 μ g mL⁻¹ foram adicionados aos *eppendorfs*, incubando-se novamente à 37 °C por mais uma hora e meia. Volumes de 100 μ L de uma solução de NaCl a 5 mM e CTAB a 2% foram adicionados a mistura reacional, que em seguida foi submetida à incubação final a 65 °C sob agitação por 10 min.

Após o período de incubação, volumes de 750 μ L de uma mistura composta de fenol, clorofórmio e álcool isoamílico (25:24:1) foram adicionados, homogeneizados cuidadosamente e em seguida foram submetidos a centrifugação por 10 min a 14000 rpm em microcentrifuga. Os sobrenadantes foram transferidos para novos microtubos estéreis.

Aos sobrenadantes foram adicionados 450 μ L de isopropanol e estes foram agitados algumas vezes por inversão. Após centrifugação por 10 min a 14000 rpm e descarte dos sobrenadantes, os DNAs foram lavados com 1 mL de etanol 70%, centrifugados e secos a temperatura ambiente por aproximadamente 30 min.

Os DNAs foram ressuspensos em 30 μ L de tampão TE 10 mM pH 7,5 e posteriormente armazenados em *freezer* a -20 °C. Para estimar a quantidade do DNA

extraído foi medida a absorvância a 260 e 280 nm contra o tampão TE, utilizando-se o espectrofotômetro UV-Vis Hitachi, modelo U-2800 (Hitachi High-Technologies, Tóquio, Japão). A relação padrão de que uma unidade de densidade óptica (DO) equivale a aproximadamente 50 µg de DNA dupla fita por mL de solução foi considerada (SAMBROOK *et al.*, 1989).

II. 2.2 – Construção dos Oligonucleotídeos e Amplificação dos Genes

As seqüências nucleotídicas das ORFs que codificam as proteínas putativas hidroxinitrila liase (XfHNL), corismato sintase (XfaroC) e proteína D do sistema de transporte e secreção Tat (XftatD) de *X. fastidiosa* foram amplificadas através da reação em cadeia da polimerase (PCR). A Figura 13 demonstra a seqüência codificadora da proteína XfHNL como um exemplo da estratégia de amplificação utilizada para obtenção dos genes de interesse deste estudo.

AATTCATATGAACGCATACCGTCCTATCGCTTTTGCCTCATCCGCAGCATCTGATAAGCCAA
 CCGTTTTACTGAAACCGACAATTGTGCTCGTACACGGTGCTTTTGCTGATGGTTCTACATGGA
 ACAAAGTAATCCATCAGCTCCAGGCAAAGGGCTGAATGCTGTGTCCGTTCAAATCCGCTGA
 CGTCTTTCGAAGATGATGTAGCAGCAACACGCCGTGTGCTTGCTCTGCAAACGGGACCTGTAG
 TGCTCGTTGGCCACTCTTATGGGGGAGCAGTGATTACTGAAGCTGGACAAGATGAACGAGTTA
 AAGCCCTAGTTTATATTGCCGCATTTGCTCCATCCGAGGGAGAGTCGGTTGCTGATCTGAATA
 AAAAATATCCATTACCTTCTGGATATAATCATCTCAGCAGTGACAAAGAAGGGTTCCTGATGC
 TAACGCCAGAAGGTGTGGAAAAATACTTGGCTCAGGATATCCCCCTTGAGCAAACCTCGCCTGA
 TCATCGCCACTCAGCACCCCTATACGCGGCGCCAATTTTGAAAAAAGGTTTCAGCTGCCGCTT
 GGGAAACTAAACCATCATGGTATTTGGTGAGTGAGAATGATCTTATGCTTCAGCCAGCGCTGC
 AAAAAGAGATGGCTCAAAAAATAGGTGCTCATATCGTCAATGTGGCAGCCAGCCATGTGCCAC
 ATCTTTCACATGCTGCGGAAGTAGAAAAATCATCATGAGTGCAGTCAAC**GATGTTGAAGTAT**
TAGAATAACTCGAGCGG

Figura 13. Sequência codificadora da ORF XfHNL (*Xylella fastidiosa* Genome Project – número de acesso XFa0032). A sequência possui 773 nucleotídeos sendo que 756 nucleotídeos codificam para a proteína putativa hidroxinitrila liase XfHNL. Os nucleotídeos em negrito vermelho e azul correspondem aos códons de início e de parada tradução, respectivamente. As seqüências em negrito representam os oligonucleotídeos *forward* e *reverse* (5' e 3') utilizados como iniciadores (primers). Os nucleotídeos sublinhados representam os sítios de clivagem para as enzimas de restrição *NdeI* e *XhoI*.

A estratégia resumiu-se em utilizar dois *primers* (oligonucleotídeos iniciadores) complementares às extremidades 3' e 5' do gene de interesse as quais hibridizavam com esta região e serviram como ponto de partida para a polimerização pela *Taq* DNA polimerase.

O DNA genômico, isolado de acordo com Ausubel (AUSUBEL *et al.*, 1995) e tratado com RNase, foi utilizado como molde para a amplificação das ORFs (*full-length*) anotadas no banco de dados genômico da *X. fastidiosa* (códigos de acesso XFa0032, XF1369 e XF 1913). Os oligonucleotídeos *forward* e *reverse* com sítios de restrição para as enzimas *NdeI* e *XhoI*, respectivamente, foram sintetizados e

utilizados na amplificação a fim de facilitar a clonagem dentro dos vetores de expressão escolhidos. Estes oligonucleotídeos iniciadores estão listados na Tabela 2.

Tabela 2. Oligonucleotídeos iniciadores utilizados na amplificação por PCR. Os nucleotídeos em azul representam as trincas dos códons de iniciação (*start codon*) e terminação (*stop codon*) presentes na seqüência das ORFs de interesse nas quais codificam para a metionina inicial e o término do processo de tradução, respectivamente.

PROTEÍNA	CLONES	OLIGOS	SEQÜÊNCIA
Hidroxinitrila Liase	XfHNL	XFa0032F	5'-AATTCCATATGAACGCATACCGTCC-3'
		XFa0032R	5'-CCGCTCGAGTTATTCTAATACTTCAACATC-3'
Corismato Sintase	XfaroC	XF1369F	5'-GGAATTCCATATGGGGGCGAACACGTTTGG-3'
		XF1369R	5'-CCGCTCGAGTCAATCATCGGCATTCTG-3'
Proteína D do Sistema Tat	XftatD	XF1913F	5'-GGAATTCCATATGCAATTGATCGACATCGG-3'
		XF1913R	5'-CCGCTCGAGTCATGAACCGGAACCGG-3'

A amplificação via PCR foi realizada em um termociclador PTC-100™ *Programable Thermal Controller* (MJ-Research Inc., Watertown, MA, EUA) empregando-se um volume final de 50 µL constituído de 100 pmol de cada *primer*, 133 ng de DNA genômico, 10 mM de dNTPs, 1 U de polimerase *Taq Platinum High Fidelity* e 2 mM de MgSO₄.

As condições de amplificação compreenderam uma desnaturação inicial a 94 °C por 120 s, seguida de 39 ciclos constituídos de desnaturação a 94 °C por 60 s, hibridização (*annealing*) variando de 52 e 58 °C (de acordo com o gene empregado) por 60 s e extensão dos *primers* a 72 °C por 90 s; com extensão final a 72 °C por 10 min. Ao final da reação a mesma foi interrompida por resfriamento a 4 °C.

Os produtos das amplificações foram analisados aplicando-se 2 μL das misturas de reação em gel de agarose preparativo a 0,8% contendo 0,5 $\mu\text{g mL}^{-1}$ de brometo de etídeo, em tampão 1X TAE (80 mM Tris; 40 mM ácido acético glacial; 2,5 mM EDTA, pH 8,3) e em seguida foi submetido a uma tensão de 100 V por 30 min. Em seguida, foram visualizados sob luz UV após migração.

II. 2.3 – Purificação dos Produtos Amplificados por PCR.

Uma vez confirmado o tamanho correto dos fragmentos de DNA obtidos, os produtos de PCR restantes (cada 48 μL) foram separados novamente por eletroforese em gel de agarose preparativo, conforme descrito anteriormente.

Os produtos de PCR foram excisados do gel e purificados utilizando-se o *Kit Wizard[®] SV Gel and PCR Clean-Up System* (Promega), de acordo com as instruções do fabricante. O resultado da extração foi analisado por eletroforese em gel de agarose 1% e quantificado por espectrofotometria no UV-visível empregando-se o método de Sambrook e colaboradores (SAMBROOK *et al.*, 1989).

II. 2.4 – Preparação dos Produtos de PCR purificados – Reação de Adenilação

Para uma correta e eficiente ligação dos produtos de PCR no vetor de clonagem pGEM-T *easy* fez-se necessária a modificação prévia destes através da introdução *in vitro* de uma adenina (A) em suas extremidades 3' terminais mediante uma reação denominada *A-tailing procedure*. Nesta reação, os substratos são o produto da PCR e

uma deoxiadenosina tri-fosfato (dATP) e a enzima é uma *Taq* DNA polimerase sem atividade corretiva (exonucleásica).

As reações de adenilação foram executadas em um termociclador PTC-100™ empregando-se volumes finais de 10 µL constituídos de 6 µL do produto de PCR purificado, 1 µL de tampão *Thermophilic DNA poly 10x Buffer* contendo MgCl₂, 2 µL de dATP a 1 mM e 1 µL de enzima *Thermophilic DNA polymerase* (5U). Uma vez preparadas, as reações foram incubadas por 30 minutos a 72 °C no termociclador.

II. 2.5 – Clonagem dos Produtos de PCR adenilados nos Vetores de Clonagem pGEM-T *easy* - Reação de Ligação

O sistema pGEM-T *easy* (Promega) é um kit que contém um plasmídeo linear que possui uma deoxitimidina (T) em cada uma de suas extremidades 3' e que permite a inserção direta de produtos de PCR adenilados, pela enzima T4 DNA ligase. A circularização do DNA pela inserção do produto de PCR adenilado nada mais é que a formação do plasmídeo recombinante a ser introduzido posteriormente dentro da bactéria *E. coli* visando à propagação plasmidial.

As seqüências codificantes dos genes de interesse, amplificadas e adeniladas, foram inseridas no vetor pGEM-T *easy* através de uma reação de ligação. Cada mistura de ligação foi preparada empregando-se um volume final de 10 µL constituídos de 2 µL produto de PCR adenilado a 100 ng µL⁻¹, 1 µL da enzima T4 DNA ligase (6U), 5 µL do tampão de ligação para a enzima T4 DNA ligase (2X) e 1 µL da solução contendo o vetor pGEM-T *easy* a 50 ng µL⁻¹. Água Milli-Q foi utilizada para completar o volume total. Esta reação foi mantida a 20 °C por 24

horas. As misturas de ligação foram utilizadas na transformação por choque térmico de células competentes de *E. coli* DH5 α (Promega), competentes por tratamento com CaCl₂.

Os plasmídeos recombinantes obtidos foram denominados de pGEM-XfHNL, pGEM-XfaroC e pGEM-XftatD correspondentes aos insertos codificando as proteínas putativas hidroxinitrila liase, corismato sintase e proteína D do sistema de transporte e secreção Tat, respectivamente.

II. 2.6 – Transformação da Bactéria *Escherichia coli* DH5 α com os Plasmídeos Recombinantes pGEM-T *easy*

As suspensões de células competentes DH5 α , produzidas no momento de uso ou estocadas a -80 °C, foram mantidas durante pelo menos 10 min em banho de gelo. A cada 50 μ L de células DH5 α foram adicionados 2 μ L das reações de ligação. As misturas resultantes foram homogeneizadas gentilmente e mantidas em gelo por 30 min. Em seguida, estas misturas foram submetidas à incubação em banho de água a 42 °C por 2 min e subsequente banho de gelo por mais 2 min (choque térmico). Após este período, 250 μ L de meio LB foram adicionados a cada um dos tubos contendo as células transformadas com as reações de ligação. As novas misturas foram mantidas a 37 °C por 1 hora e meia sob agitação de 150 rpm. Estas foram plaqueadas em placas de Petri contendo 20 mL de meio seletivo Luria-Bertani sólido (LB; 15 g extrato de levedura, 10 g de bacto-triptona, 10 g de NaCl e 15 g de ágar por litro de solução a pH 7,5) suplementado com 10 μ L do antibiótico de seleção

ampicilina a 100 mg mL^{-1} , $40 \text{ }\mu\text{L}$ de X-Gal a 50 mg mL^{-1} e $30 \text{ }\mu\text{L}$ de IPTG a $0,5 \text{ M}$; e incubadas *overnight* a $37 \text{ }^\circ\text{C}$ em estufa.

II. 2.7 – Extração, Análise de Restrição e Seqüenciamento dos Plasmídeos Recombinantes pGEM-T easy de *Escherichia coli* DH5 α

Após o período de incubação, algumas colônias brancas resistentes (possíveis recombinantes ou clones positivos) foram selecionadas a fim de realizar a extração dos seus DNAs plasmidiais. A seleção definitiva dos recombinantes foi realizada através da análise de restrição e confirmada pelo do seqüenciamento de DNA.

As colônias previamente selecionadas foram repicadas, individualmente, em tubos estéreis para cultivo de bactérias contendo 5 mL de meio LB líquido (sem ágar) contendo $3 \text{ }\mu\text{L}$ de ampicilina a 100 mg mL^{-1} (pré-inóculo). Os tubos foram incubados a $37 \text{ }^\circ\text{C}$ por 16 horas, sob agitação de 250 rpm . A seguir, as células tiveram seus DNAs plasmidiais extraídos seguindo dois métodos:

i) por meio de lise alcalina (AUSUBEL *et al.*, 1995) para uma análise prévia de restrição enzimática a fim de verificar a presença dos produtos de PCR ligados aos vetores e confirmar os clones positivos;

ii) através do *kit Wizard[®] Plus SV Miniprep DNA Purification Systems* (Promega) para análise destes DNAs por seqüenciamento e obtenção dos genes a serem posteriormente subclonados nos vetores de expressão.

Ambos os métodos proporcionaram materiais adequados para os procedimentos seguintes, porém a extração por kit proporcionou materiais livres de contaminantes como DNA genômico, RNA, proteínas e macromoléculas; mais adequados aos procedimentos de seqüenciamento e subclonagem em plasmídeos de expressão. A

quantificação dos DNAs plasmidial foi realizada empregando-se o método de Sambrook e colaboradores (SAMBROOK *et al.*, 1989).

A primeira análise de restrição foi realizada empregando-se um volume final de 10 μL constituídos de 0,8 μL dos DNAs plasmidiais a 100 $\text{ng } \mu\text{L}^{-1}$, 1 μL da enzima *NdeI* (10 U), 1 μL da enzima *XhoI* (10 U) e 1 μL do tampão de enzima D (10X), estes últimos fornecidos pelo fabricante. Água Milli-Q foi utilizada para completar o volume total. Esta reação foi mantida a 37 °C em estufa por 4 horas. Os produtos de digestão foram analisados em gel de agarose preparativo 0,8% e visualizados sob luz UV.

Dos clones positivos identificados, dois clones contendo cada ORF de interesse foram selecionados e tiveram seu DNA plasmidial purificado por kit, quantificado e submetido ao seqüenciamento automático.

Uma segunda digestão também foi realizada com tais plasmídeos, porém empregando-se um volume de 50 μL . Este último procedimento foi necessário para a obtenção de quantidade suficiente dos insertos de interesse com extremidades coesivas adequadas para serem subclonados nos vetores de expressão, previamente digeridos com as mesmas endonucleases de restrição.

O seqüenciamento de DNA foi realizado utilizando-se o método do dideoxinucleotídeo marcado (*dideoxy chain termination method*) com auxílio de um seqüenciador automático *ABI Prism 377-96 Collection* da *Applied Biosystem* (Sarasota, FL, EUA), seguindo o protocolo do fabricante. Após o seqüenciamento, para verificar a integridade da seqüência obtida, esta foi submetida a uma busca em bancos de dados públicos como NCBI através da ferramenta BLAST (<http://www.ncbi.nlm.nih.gov/>).

II. 2.8 – Subclonagem dos Genes no Plasmídeo de Expressão pET28a

Plasmídeos do sistema pET utilizados possuem um gene que confere resistência a canamicina e podem apresentar uma seqüência nucleotídica na região N- ou C-terminal do sítio de múltipla clonagem que codifica seis histidinas denominada *His-tag*. Assim, as ORFs de interesse foram subclonados em fusão com esta seqüência facilitando a purificação das proteínas recombinantes através de cromatografia de afinidade em metal imobilizado. Os plasmídeos de expressão construídos foram denominados pET-XfHNL, pET-XfaroC e pET-XftatD, os quais codificam as proteínas putativas hidroxinitrila liase, corismato sintase e proteína D do sistema de transporte e secreção Tat, respectivamente.

As seqüências codificantes de interesse foram inseridas no vetor pET28a através de uma nova reação de ligação. Cada mistura de ligação foi preparada empregando-se um volume final de 10 μL constituídos de 1,5 μL do inserto a 100 $\text{ng } \mu\text{L}^{-1}$, 1,7 μL do vetor pET28a a 90 $\text{ng } \mu\text{L}^{-1}$, 5 μL do tampão de ligação para a enzima T4 DNA ligase (2X) e 1 μL da enzima T4 DNA ligase (6U). Água Milli-Q foi utilizada para completar o volume total. Esta reação foi mantida a 20 °C por 24 horas.

As novas misturas de ligação foram utilizadas na transformação por choque térmico de células competentes de *E. coli* DH5 α para nova propagação plasmidial.

II. 2.9 – Transformação da Bactéria *Escherichia coli* DH5 α com os Plasmídeos Recombinantes pET28a

O procedimento empregado na transformação de células competentes de *E.coli* DH5 α foi o mesmo descrito na seção II. 2.6, porém, com algumas modificações necessárias. Foram utilizados para transformar as células competentes por choque térmico os volumes totais das novas misturas de ligação. As células transformadas foram plaqueadas em placas de Petri contendo 20 mL de meio seletivo LB suplementado apenas com 12 μ L do antibiótico de seleção canamicina 50 mg mL⁻¹ e incubadas *overnight* a 37 °C em estufa.

II. 2.10 – Extração e Análise de Restrição dos Plasmídeos Recombinantes pET28a de *Escherichia coli* DH5 α

Após o período de incubação, algumas colônias resistentes foram selecionadas para realizar a extração dos seus DNAs plasmidiais. A seleção dos recombinantes foi realizada através da análise de restrição a fim de verificar a eficiência na transformação e propagação e obter quantidade suficiente dos plasmídeos para posterior transformação em linhagens de expressão.

As colônias selecionadas foram repicadas, individualmente, em 5 mL de meio LB líquido contendo 3 μ L de canamicina a 50 mg mL⁻¹. Os tubos foram incubados *overnight* a 37 °C em estufa sob agitação de 250 rpm até uma DO_{600 nm} entre 1 e 2.

As células tiveram seus DNAs plasmidiais extraídos através do *kit Wizard® Plus SV Miniprep DNA Purification Systems* e quantificados empregando-se o método de Sambrook e colaboradores (SAMBROOK *et al.*, 1989).

A análise de restrição foi realizada empregando-se um volume final de 10 μ L de acordo ao procedimento descrito na seção II. 2.7. Uma vez confirmada a presença dos DNAs plasmidiais nestas células, os mesmos foram utilizados na transformação por choque térmico de células competentes de *E. coli* BL21(DE3) para expressão heteróloga.

II. 2.11 – Transformação da Bactéria *Escherichia coli* BL21(DE3) com os Plasmídeos Recombinantes pET28a

O procedimento empregado na transformação de células competentes de *E. coli* BL21(DE3) foi o mesmo descrito na transformação de células DH5 α (seção II. 2.9). Porém, 10 μ L de cada solução de DNA plasmidial foram utilizados para transformar por choque térmico células competentes de *E. coli* BL21(DE3). A linhagem empregada nesta etapa do trabalho também passou pelo tratamento com cloreto de cálcio (AUSUBEL *et al.*, 1995).

II. 2.12 – Expressão Heteróloga das Proteínas Recombinantes de *Xylella fastidiosa* (rXfHNL, rXftatD e rXfaroC) em *E. coli*

Para realizar os testes de expressão foram feitos pré-inóculos das colônias resultantes da transformação de células BL21(DE3) com os plasmídeos pET-XfHNL, pET-XftatD e pET-XfaroC, selecionadas em meio sólido com antibiótico. Colônias isoladas contendo cada uma destas construções foram escolhidas aleatoriamente. Estas colônias foram inoculadas, individualmente, em 5 mL de meio LB líquido

contendo canamicina ($3 \mu\text{L}$ a 50 mg mL^{-1}) e crescidas *overnight* a 37°C sob 250 rpm de agitação.

As culturas foram então diluídas na proporção 1:100 com LB líquido contendo canamicina e crescidas, novamente, a 37°C até alcançar a OD_{600} de aproximadamente 0.6. Depois de atingido este crescimento celular, alíquotas de $100 \mu\text{L}$ de cada cultura foram armazenadas a -80°C em soluções estéreis de $100 \mu\text{L}$ de glicerol a 65 % e sulfato de magnésio a 100 mM na proporção 1:1. Antes da adição de IPTG no restante da cultura, alíquotas de $250 \mu\text{L}$ de cada fracso, representando os controles não induzidos foram centrifugadas a 10000 rpm por 2 min em microcentrífuga, ressuspendidas em $20 \mu\text{L}$ de água Milli-Q e congeladas a -20°C .

O restante das culturas de células foi induzido com IPTG a 0,5 M e o crescimento foi mantido por um período inicial de 4 horas a 37°C . Alíquotas de $250 \mu\text{L}$ da cultura induzida foram retiradas de hora em hora para avaliação do nível de expressão das proteínas recombinantes. Todos os precipitados (*pellets*) provenientes das alíquotas retiradas durante a etapa de expressão protéica foram adequadamente ressuspendidos e preparados de acordo com as subseqüentes análises em gel descritas na secção de caracterização físico-química.

As culturas induzidas restantes foram centrifugadas a 12000 rpm por 10 min a 4°C em centrífuga Sorvall modelo RC-5C Plus (Kendro, CA, EUA). Os *pellets* bacterianos obtidos de cada 250 mL de cultura induzida foram congelados a 20°C e descongelados antes do uso.

II. 2.13 – Purificação das Proteínas Recombinares de *Xylella fastidiosa* (rXfHNL, rXftatD e rXfaroC)

Métodos cromatográficos são rotineiramente utilizados para identificar e isolar proteínas de interesse através de sucessivos fracionamentos de acordo com suas propriedades físico-químicas como solubilidade, carga iônica, massa molecular, ponto isoelétrico, propriedades de adsorção e ligação a outras moléculas biológicas.

As técnicas cromatográficas mais comuns empregam colunas verticais onde as amostras são colocadas e as moléculas protéicas que as constituem se distribuem entre uma fase estacionária e outra móvel. A fase móvel corresponde ao solvente cujo fluxo arrasta a maioria das moléculas presentes nas amostras. No entanto, a fase estacionária é constituída de um material imóvel acondicionado (como resinas carregadas, camadas porosas, papel de filtro, entre outros) de modo a interagir com a mistura de proteínas e outras moléculas (ZAHA *et al.*, 1996, STRYER *et al.*, 1988).

Os métodos cromatográficos mais utilizados para isolamentos de proteínas são cromatografias de exclusão por tamanho, troca iônica e afinidade. Neste trabalho, os métodos aplicados para isolar as proteínas recombinantes foram cromatografias de afinidade em resina de níquel e exclusão por tamanho em coluna Superdex 75-HR 10/30.

A solução tampão utilizada na purificação das proteínas recombinantes foi tampão de lise Tris-HCl constituído de 20 mM de Tris base, 150 mM de NaCl e 5 % de glicerol, pH 8,0 ajustado com HCl concentrado.

Os *pellets* bacterianos descongelados obtidos a partir de 250 mL de cultura induzida foram ressuspensos em 12,5 mL de tampão de lise Tris-HCl e mantidos em gelo por 30 min.

As suspensões celulares mantidas em gelo foram submetidas à lise celular utilizando-se o sonicador 550 Sonic Dismembrator (Fisher Scientific International Inc., Hampton, NH, EUA) acoplado a uma sonda *microtip* e uma fonte de alta voltagem. O processo de sonicação foi realizado empregando-se 10 ciclos de 30 segundos com intervalos de 30 segundos. Os extratos protéicos brutos provenientes das células lisadas foram clarificados por centrifugação a 10000 rpm por 30 min a 4 °C. Os sobrenadantes foram empregados na purificação por cromatografias de afinidade e exclusão molecular.

II. 2.13.1 – Cromatografia de Afinidade em Resina de Níquel

A estratégia empregada para a purificação das proteínas recombinantes contendo cauda de hexahistidinas utiliza as vantagens da cromatografia de afinidade em metal. Os íons metálicos são imobilizados pelo uso de um agente quelante capaz de disponibilizar o metal para ligação à proteína. Certos aminoácidos, principalmente histidina, apresentam alta especificidade de ligação pelo metal imobilizado. Assim, proteínas ricas em histidinas ou com um grupo destes aminoácidos podem ser seletivamente eluídas da resina carregada com íons metálicos e, conseqüentemente, isoladas por este método. Uma representação da interação dos resíduos de histidinas com a matriz Ni-NTA é apresentada na Figura 14.

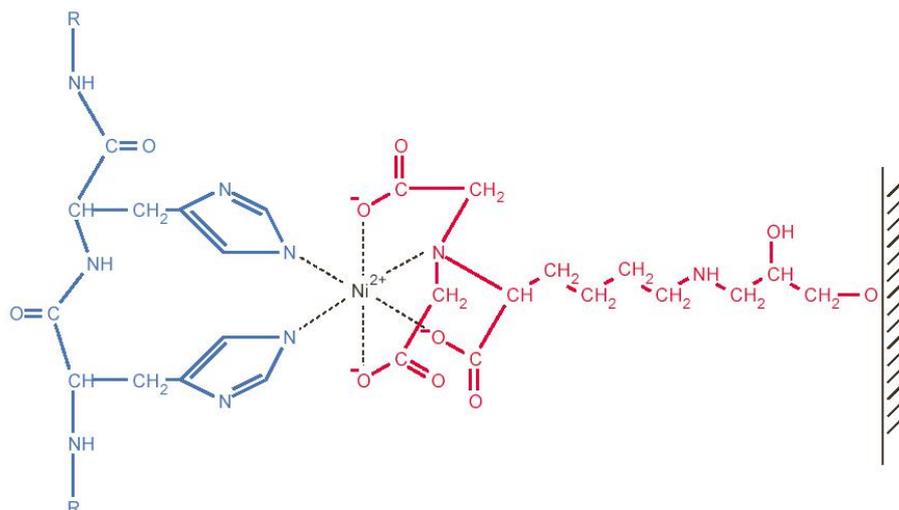


Figura 14. Interação entre resíduos vizinhos da cauda *His-Tag* e a matriz Ni-NTA. O quelante em vermelho, o ácido nitrilotriacético, coordena-se com quatro dos seis sítios ligantes do Ni (II). Os dois sítios livres da esfera de coordenação interagem com os anéis imidazol da cauda de 6X histidinas das proteínas recombinantes (em azul) (Qiagen).

A eluição das proteínas de interesse é realizada utilizando-se uma solução contendo o composto imidazol, que também possui o anel imidazólico, mas tem maior afinidade ao Ni (II) do que a histidina. A Figura 15 apresenta a estrutura molecular do imidazol e da histidina que se ligam aos íons níquel que estão imobilizados pelos grupos NTA na matriz.

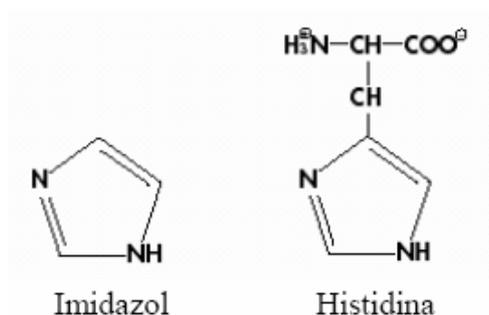


Figura 15. Estrutura molecular do composto imidazol e do aminoácido (Qiagen).

Os sobrenadantes foram inseridos em colunas contendo 3 a 5 mL de resina de níquel *Ni-NTA superflow* (QIAGEN) pré-equilibrada com tampão de lise. Após 30 min de repouso a 20 °C, os materiais não ligados de 12,5 mL de volume foram eluídos da coluna. As colunas foram então lavadas com um gradiente inicial de imidazol variando de 0 a 100 mM em tampão de lise Tris-HCl. As etapas de lavagem consistiram de 18 a 30 mL de tampão de lise sem imidazol, seguido de 18 a 50 mL de tampão de lise contendo 5 mM de imidazol, 24 a 50 mL de tampão contendo 40 mM de imidazol e 6 a 10 mL de tampão contendo 100 mM de imidazol. Todas as proteínas recombinantes foram eluídas com aproximadamente 10 mL de tampão contendo 250 mM de imidazol. Cada proteína recombinante deste estudo possuía uma coluna de afinidade individual para evitar eventuais contaminações.

Alíquotas de 20 µL de todas as frações de lavagem e eluição foram avaliadas e o grau de pureza das proteínas recombinantes avaliado por eletroforese em gel de poliacrilamida 15%, de acordo com as subseqüentes análises em gel descritas na secção de caracterização físico-química.

As frações provenientes das eluições com 250 mmol L⁻¹ de imidazol contendo as proteínas recombinantes purificadas foram reunidas e dialisadas contra o próprio tampão de lise (sem imidazol). Em seguida foram concentradas através de concentradores Centriprep Ultra 10 da Amicon (Millipore Corporation, Bedford, MA, EUA) contendo membranas de tamanho adequadas às suas massas moleculares. Cada proteína recombinante possuía também um concentrador individual para evitar eventuais contaminações.

As concentrações das proteínas recombinantes purificadas foram determinadas medindo-se suas absorvâncias em espectrofotômetro UV-Vis a 280 nm. As absorvâncias foram relacionadas aos seus respectivos coeficientes de extinção molar

btidos através do cálculo do conteúdo teórico de resíduos de triptofanos, tirosinas e fenilalaninas (GASTEIGER *et al.*, 2005). Este método proporcionou os coeficientes de extinção molar para rXfHNL, rXfaroC e rXfatD de 26030, 28930 e 21980 L mol⁻¹ cm⁻¹, respectivamente. Todos os estudos subseqüentes foram realizados com as proteínas recombinantes contendo a cauda de 6× histidinas.

II. 2.13.2 – Cromatografia de Exclusão por Tamanho em Colunas

Superdex

Esta técnica de separação utiliza colunas empacotadas com polímeros contendo esferas de tamanho específico nas quais formam uma rede tridimensional de poros proporcionando a separação das moléculas em solução de acordo com seus tamanhos. Dentre as colunas empacotadas utilizadas na cromatografia de exclusão por tamanho estão as colunas Superdex que são empregadas na separação de proteínas e outras biomoléculas. Por serem de tamanho selecionado, as matrizes destas colunas permitem que proteínas que apresentam maiores massas moleculares migrem mais rapidamente do que proteínas com menores massas moleculares. Devido a não interação com seus poros, estas são eluídas mais precocemente.

Como alternativa para a retirada do imidazol das frações protéicas provenientes da purificação em coluna de níquel e para prever o grau de oligomerização das proteínas recombinantes, a cromatografia de exclusão por tamanho foi realizada a temperatura ambiente empregando-se a coluna Superdex 75-HR 10/30, acoplada a um sistema HPLC preparativo (ÄKTA™-purifier 10, GE Healthcare BioSciences, Uppsala, Sweden). As colunas foram previamente lavadas com água Milli-Q e pré-equilibradas com tampão de lise Tris-HCl. Os eluatos foram monitorados por

absorbância no UV a 280 nm e coletados em frações de 1 mL empregando-se uma velocidade de fluxo de 0,5 mL min⁻¹. As frações contendo as proteínas recombinantes purificadas foram reunidas e concentradas através dos concentradores *Centriprep Ultra 10*.

Os padrões de massa molecular utilizados na calibração foram soro albumina bovina (BSA – 66 kDa), ovalbumina (OVA – 45 kDa), anidrase carbônica (CA –30 kDa), inibidor de tripsina de soja (ITS - 20,1 kDa) e α -lactoalbumina (a-Lacto - 14.4 kDa), separados individualmente.

As concentrações das proteínas recombinantes purificadas foram determinadas medindo-se suas absorbâncias por espectrofotômetro UV-vís a 280 nm. Todos os estudos subseqüentes foram realizados com as proteínas recombinantes contendo a cauda de 6X histidinas.

II. 2.14 - Análises Físico-química e Estrutural das Proteínas Recombinantes de *Xylella fastidiosa* (rXfHNL, rXftatD e rXfaroC)

II. 2.14.1 – Determinação Quantitativa do Nível de Expressão Protéica

As alíquotas derivadas da expressão heteróloga foram então avaliadas quanto à homogeneidade e tiveram suas massas moleculares e quantificações estimadas através de dois métodos eletroforéticos de análise: gel de poliacrilamida e microchip. Uma vez estabelecida às condições de expressão mais favoráveis, as proteínas recombinantes foram expressas em larga escala.

II. 2.14.1.1 - Eletroforese em Gel de Poliacrilamida (SDS-PAGE)

A separação, detecção e quantificação das amostras rXfHNL XfaroC e XftatD contidas nas alíquotas de cultura induzida foram realizadas segundo o método de Laemmli (LAEMMLI *et al.*, 1970). No sistema eletroforético foi utilizado gel de empilhamento de concentração igual a 5 % contendo bis-acrilamida:acrilamida: na proporção 0,8:30 (m/v) em Tris-HCl 0,5 M, pH 6,8 e SDS 0,1 % e gel de separação a 15% contendo bis-acrilamida:acrilamida na proporção 0,8:30 (m/v) em Tris-HCl 2,0 M, pH 8,8 e SDS 0,1 %. Os géis foram polimerizados em placas de vidro adaptáveis a sistemas verticais, as quais foram acopladas a seus respectivos reservatórios contendo tampão Tris-HCl a 25 mM contendo SDS a 0,1 % (m/v) utilizado como tampão de corrida.

Um volume de 10 μ L de tampão Tris-HCl a 0,0625 M, pH 6,8 contendo glicerol 10 %, β -mercaptoetanol 5 % (v/v), SDS 2% e azul de bromofenol a 0,001 % (tampão de amostra) foi adicionado a 20 μ L de todas as amostras submetidas a eletroforese. As soluções foram aquecidas a 100 °C por 5 min e aplicadas nos poços, num volume final de 20 μ L. As proteínas utilizadas como padrões de massa molecular foram albumina sérica bovina (BSA: 66 kDa), ovalbumina (Ova: 45 kDa), anidrase carbônica (CA: 30 kDa), inibidor de tripsina de soja (ITS: 20,1 kDa), citocromo C (cit C: 12,4 kDa) e α -lactoalbumina (α -Lacto: 14,4 kDa).

A eletroforese foi efetuada sob uma tensão de 120 V para a separação das amostras durante a migração no gel de empilhamento e de resolução. Após 2 horas de separação eletroforética, a presença das bandas protéicas foi detectada com a

imersão dos géis numa solução de corante *Comassie Brilliant Blue 250 R*, por 15 min sob agitação branda. Em seguida, os géis foram descorados com ácido acético 7 %.

A quantificação das proteínas recombinantes foi feita por densitometria do gel corado com *Coomassie* utilizando-se o software *Kodak 1D image analysis* (Rochester, New York, EUA) e anidrase carbônica (CA, 30 kDa) como padrão.

II. 2.14.1.2 - Eletroforese em Gel através de Microchip (Método GelChip-CE)

A separação, detecção e quantificação das amostras contidas nas alíquotas de cultura induzida também foram realizadas no instrumento *Agilent 2100 Bioanalyzer* - método GelChip-CE (Agilent Technologies, Waldbronn, Germany) em combinação com o *kit Protein 200 plus LabChip* e o software *Protein 200 Plus* (version A.02.12).

A preparação (*priming*) e o preenchimento dos *microchips* com as amostras foram realizados de acordo com o protocolo do *kit Protein 200 Plus LabChip*, como recomendado pelo fabricante. Cada *microchip* de vidro continha 16 reservatórios interconectados via rede de microcanais, que foram preenchidos com uma mistura de corante e gel. Dez dos 16 reservatórios foram preenchidos com as amostras experimentais, quatro reservatórios com a mistura gel-corante, um com a solução descorante (*destaining solution*) e um com o padrão de tamanho molecular (*ladder*).

Brevemente, os microcanais foram preenchidos pipetando-se 12 μ L da mistura gel-corante dentro dos reservatórios apropriados. A mistura foi inserida dentro dos microcanais aplicando-se uma pressão ao reservatório adequado via seringa de 1mL. Os *microchips* recentemente preparados (*priming*) foram utilizados imediatamente. Três reservatórios adicionais para tampão foram preenchidos com 12 μ L da mistura

gel-corante. As amostras foram desnaturadas misturando-se estas com tampão de amostra (*buffer* proporcionado pelo kit) na proporção 2:1 e aquecidas em banho de água fervente por 3 min. O tampão de amostra fornecido com o *kit Agilent Protein 200 plus Assay* continha 4 % SDS, um corante fluorescente apropriado, e duas proteínas controle de tamanho conhecido servindo como marcadores internos de tamanho. Após a etapa de desnaturação, as amostras foram diluídas em água desionizada na proporção 1:15 e inseridas no *microchip*. Os reservatórios das amostras e do padrão foram preenchidos com as soluções apropriadas. O *microchip* foi agitado em um agitador tipo *vortex* e inserido dentro do instrumento *Bioanalyzer* para análise.

As amostras foram eletroforéticamente separadas nos microcanais preenchidos com solução polimérica. A detecção foi baseada na fluorescência induzida a laser de um corante fluorescente que foi adicionado à solução polimérica e ligado não covalentemente às micelas proteína-SDS. O software *Bioanalyzer* automaticamente calcula o tamanho e a concentração de cada pico separado e expõe os resultados em tempo real. A quantidade de proteína recombinante nas amostras expressa como porcentagem de proteína total pode ser determinada diretamente da tabela dos resultados fornecida pelo software. Os resultados obtidos por GelChip-CE foram comparados com aqueles obtidos por gel de eletroforese convencional (15 % SDS-PAGE) (LAEMMLI *et al.*, 1970).

II. 2.14.2 - Determinação do Ponto Isoelétrico

As proteínas são macromoléculas anfóteras, ou seja, possuem cargas positivas ou negativas dependendo do pH. A rede de cargas nas proteínas é uma soma

algébrica de todas as cargas das cadeias laterais dos aminoácidos que as compõe. Cada proteína possui um pH específico na qual a somatória das cargas é nula ou zero e cujo valor é denominado ponto isoelétrico (pI) (STRYER, 1988).

A focalização isoelétrica (IEF, *isoelectric focusing*) é uma técnica eletroforética de alta resolução que separa proteínas em uma mistura de acordo com seus pontos isoelétricos. Ela é realizada em condições não desnaturantes e têm como princípio a movimentação destas macromoléculas carregadas em um gradiente de pH quando um campo elétrico é aplicado. Nesse gradiente, as proteínas se movimentam até a atingirem seus pIs. Em valores de pH acima do pI da proteína estas são negativamente carregadas e em valores abaixo do seu pI são positivamente carregadas.

A eletroforese de focalização isoelétrica com as amostras protéicas provenientes das cromatografias de exclusão por tamanho foi realizada através do sistema *PhastSystem* da GE Healthcare (Uppsala, Suécia). Como meio de separação foi utilizado o gel do kit *PhastGel IEF* com um gradiente de pH variando de 3 a 9. O padrão de pI empregado foi o kit *Broad pI* variando de 3 a 10. O gel foi corado com *Coomassie Blue G – 250* a 0,1% em solução aquosa de metanol a 25 % e ácido acético a 5 % . As amostras foram previamente dessalinizadas através das mini colunas de dessalinização *Centrisep spin columns* (Princeton Separations, Aldelphia, NJ, EUA), de acordo com o protocolo do fornecedor.

II. 2.14.3 - Determinação da Seqüência por Espectrometria de Massas

A espectrometria de massas é atualmente a técnica analítica mais utilizada para identificar compostos, elucidar algumas propriedades estruturais e químicas de moléculas e, muitas vezes, determinar as seqüências aminoacídicas de peptídeos que serão comparadas com massas e seqüências depositadas em bancos de dados mundiais. Esta técnica permite a determinação de moléculas com altas massas moleculares, uma vez que é capaz de produzir abundantemente íons multiplamente carregados.

Os espectrômetros de massas contam com três componentes básicos a fim de converter moléculas individuais em íons, separá-los e detectá-los: uma fonte de ionização, um analisador de massas e um detector. Uma das técnicas de ionização mais empregadas na determinação de massas moleculares é a de *electrospray*, na qual baseia-se na ionização de moléculas peptídicas, levando os íons a fase gasosa, sem degradação ou fragmentação das mesmas. No processo de ionização, a solução de interesse é bombeada num capilar metálico mantido a um potencial onde um campo elétrico é formado conduzindo a um acúmulo de carga na ponta do capilar. Este acúmulo de cargas leva à emissão de gotículas com altas densidades de carga que passam por um processo de evaporação do solvente e fissão eletrostática e, finalmente provocam a expulsão dos íons para a fase gasosa. Os íons são produzidos de forma contínua e acelerados.

Dentre os analisadores de massas, o quadrupolo tem sido o mais comumente empregado em equipamentos comerciais de uso geral. Neste tipo de filtro, os íons são injetados paralelamente a quatro barras metálicas, às quais são aplicadas

combinações de potenciais elétricos de corrente contínua (DC) e alternada (AC). A variação dos potenciais DC aplicados às barras metálicas faz com que somente íons com uma determinada relação massa/carga (m/z) tenham trajetórias estáveis e cheguem até o final do percurso (percurso).

Em seguida, os íons são direcionados para um detector que avalia seus tempos de vôo após sua produção e aceleração no processo de ionização.

A análise por espectrometria de massas foi utilizada para confirmar a expressão das proteínas recombinantes. A confirmação foi realizada através da seqüência de aminoácidos dos peptídeos gerados pela digestão triptica das amostras e de seus respectivos valores de massas. Esta análise foi feita junto ao Laboratório Nacional de Luz Síncroton (LNLS, Campinas, São Paulo, Brazil) com auxílio do doutorando Rogério de Campus Bicudo, usuário cadastrado nesta instituição.

Os digeridos tripticos foram analisados diretamente por cromatografia líquida acoplado em *tandem* a um espectrômetro de massas com ionização por *electrospray* no modo íon positivo (LC-nanoESI-MS/MS). A separação por LC-MS/MS foi realizada em um sistema *Micromass* (Waters Ltda, Manchester, UK) equipado com pré-coluna de desalinização e coluna analítica *C₁₈ Waters Symmetry 300* (com comprimentos de 5 mm e 150 mm, diâmetros internos de 350 μ m e 75 μ m e tamanhos de partículas de 5 e 3 μ m, respectivamente) e diretamente acoplado ao espectrômetro ESI-Q/Tof Ultima API. As soluções empregadas como fases móveis foram acetonitrila e ácido fórmico a 0,1 %, água e ácido fórmico a 0,1 % e água Milli-Q. Um padrão de íon interno foi empregado a fim de proporcionar alta exatidão nas medidas de massas em ambos os modos de aquisição MS e MS/MS.

O preparo das proteínas recombinantes provenientes das cromatografias de exclusão por tamanho foi feito utilizando o protocolo padrão de digestão triptica com

pequenas modificações quando necessário. Proteínas rXfHNL foram preparadas sem a etapa de redução e alquilação por não apresentarem resíduos de cisteínas. Todos os reagentes foram preparados imediatamente antes do uso. As soluções protéicas a $532 \mu\text{g mL}^{-1}$ foram previamente dessalinizadas, liofilizadas em *Speed-vac* e ressuspendidas em 1,5 mL de bicarbonato de amônio 0,05 mM a pH 8,0.

Quando necessário, a etapa de redução e alquilação se deu após a adição de apenas 100 μL de bicarbonato de amônio e pela adição de 5 μL de DTT a 45 mM e iodoacetamida a 100 mM. Em seguida, a mistura foi incubada por 1 hora a 56 °C em atmosfera de N_2 . Passado este período, iodoacetamida foi adicionada mantendo-se as amostras sob incubação a temperatura ambiente por mais 2 horas no escuro. Antes do início da digestão com tripsina, o restante do volume de bicarbonato de amônio foi adicionado às misturas.

Finalmente, a digestão das amostras foi realizada adicionando-se 399 μL de solução de tripsina a $20 \mu\text{g mL}^{-1}$, contendo 7,98 μg de tripsina modificada a uma proporção enzima:substrato de 1:50 (w/w). A digestão ocorreu a 37 °C *overnight*. As reações enzimáticas foram interrompidas adicionando-se 20 μL de TFA a 5%.

A seqüência de resíduos obtidos foi submetida ao sistema de busca NCBI-BLAST a fim de encontrar proteínas que apresentassem identidade seqüencial com as proteínas recombinantes.

II. 2.14.4 – Espalhamento de Raios-X a Baixos Ângulos (SAXS) e Análises dos Dados

A técnica de análise SAXS (*small-angle X-ray scattering*) fornece informações globais matemáticas quanto à forma, tamanho e estado de oligomerização de

macromoléculas em solução. Nesta técnica, um feixe de raios-X é incidido na amostra fazendo com que cada um de seus elétrons se torne uma fonte de onda espalhada, de modo que um padrão de espalhamento é observado. Este padrão contém informações sobre a estrutura espacial das partículas na amostra nas quais definem sua estrutura quaternária através de simulações com modelos.

O uso dos raios-X na caracterização estrutural de macromoléculas depende essencialmente de sua interação com a matéria através dos elétrons que constituem seus átomos e moléculas. Como o comprimento de onda de raios-X por radiação síncroton é próximo a 1 Å, que é correspondente a magnitude das ligações químicas, é possível detectar a configuração espacial dos átomos e moléculas que a compõe.

As análises de SAXS foram realizadas junto ao LNLS com auxílio do Dr Wânius Garcia e do Ms. Mario Oliveira Neto, usuário cadastrado nesta instituição. Os dados de SAXS foram feitos empregando-se feixes a baixos ângulos e detectores unidimensionais (KELLERMANN *et al.*, 1997). As curvas de espalhamento das soluções protéicas e dos correspondentes tampões foram monitoradas a comprimentos de onda de 0,148 nm com a distância amostra-detector de 732,1 mm, intervalos de medida (*momentum transfer range*) de $0.15 < q < 2 \text{ nm}^{-1}$ ($q = 4\pi\sin\theta/\lambda$, onde 2θ é o ângulo de espalhamento) e coletados em intervalos (*frames*) de 100 segundos a fim de monitorar danos na radiação e na estabilidade do feixe. Os dados obtidos foram normalizados à intensidade do feixe incidente e a resposta do detector corrigida. O espalhamento do tampão foi subtraído e as diferentes curvas foram re-escaladas para concentração.

Soluções de proteínas recombinantes provenientes da cromatografia de afinidade foram preparadas a concentrações de 3 e 12 mg mL⁻¹ em tampão de lise Tris-HCl.

A determinação dos parâmetros necessários para o refinamento dos dados experimentais foi realizada empregando-se diversas ferramentas computacionais. Os cálculos da inclinação molecular da proteína a baixas resoluções foram recuperadas dos dados experimentais de SAXS utilizando-se o procedimento *ab initio* implementado no programa GASBOR (PETOUKHOV *et al.*, 2003; DAMMIN *et al.*, 1998; SVERGUN *et al.*, 1992; SVERGUN *et al.*, 1988). Este programa gera modelos *ab initio* de baixa resolução com seqüências ao acaso C_α denominadas curvas *Dummy Atom Model* (DAM). A idéia principal é enovelar este modelo a fim de reduzir a discrepância (χ) entre as curvas DAM e as experimentais. O programa simula a estrutura interna das proteínas e não é necessário subtrair as constantes de intensidade de *Porod* (PETOUKHOV *et al.*, 2003; DAMMIN *et al.*, 1998; SVERGUN *et al.*, 1992; SVERGUN *et al.*, 1988). Os raios de giro (R_g) foram avaliados segundo o método *Guinier* ($\ln I(q) \times q^2$) (SVERGUN *et al.*, 1992; SVERGUN *et al.*, 1988) e pelo programa de transformada de Fourier indireta GNOM (SVERGUN *et al.*, 1992; SVERGUN *et al.*, 1988). As funções de distribuição de distâncias $p(r)$ também foram avaliadas pelo mesmo programa.

O banco de dados de estrutura de proteínas PDB (*Protein Data Bank*, Piscataway, NJ, EUA) foi utilizado a fim de encontrar proteínas de outros organismos com identidade de seqüência e função. Seus dados estruturais foram empregados como modelos comparativos para determinação das estruturas quaternárias das proteínas recombinantes deste estudo. O programa CRY SOL (PETOUKHOV *et al.*, 2003; DAMMIN *et al.*, 1998; SVERGUN *et al.*, 1992; SVERGUN *et al.*, 1988) gerou o espalhamento teórico das proteínas utilizadas como modelos, proporcionando parâmetros como raio de giro, distância máxima das partículas e o volume do envelope de partículas. Foi também possível comparar as

curvas experimentais com estes espalhamentos teóricos e calcular o fator de discrepância entre elas provando suas similaridades.

II. 2.14.5 – Espectroscopia de Dicroísmo Circular (CD) e Cálculo das Frações de Estrutura Secundária

A espectroscopia de dicroísmo circular (*Circular Dichroism*, CD) é a propriedade que uma molécula quiral tem de absorver diferentemente a luz circularmente polarizada à direita e à esquerda. (CAMPBELL *et al.*, 1984; CANTOR *et al.*, 1980).

Peptídeos e proteínas são amplamente estudados por CD, pois estão repletos de centros quirais (carbonos α , C_α) distribuídos ao longo da cadeia principal em hélices alfa, folhas beta, estruturas desordenadas e voltas denominadas estruturas secundárias. Portanto, é possível caracterizar a estrutura secundária dessas macromoléculas e quantificá-las por seus espectros de CD. A região do UV-distante nos espectros de CD, entre 180 e 260 nm, fornece informações capazes de caracterizar e discriminar as estruturas presentes nestas moléculas (SREERAMA *et al.*, 2000). A Figura 16 mostra os espectros de CD de proteínas compostas por diferentes tipos de estrutura secundária.

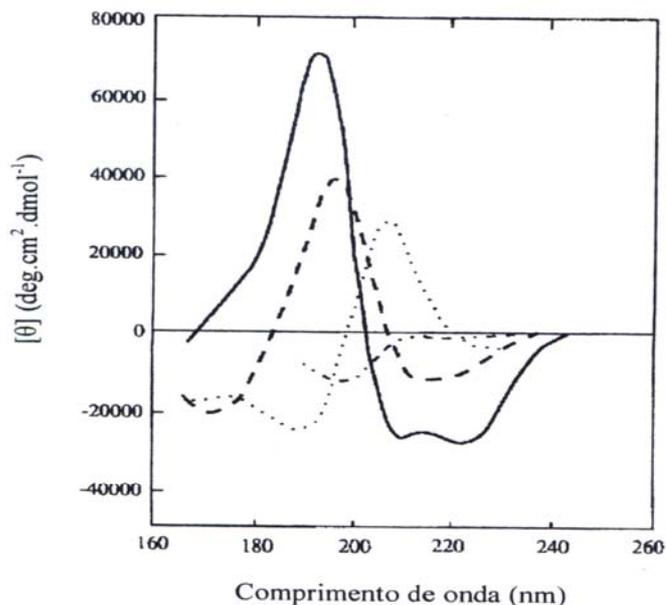


Figura 16. Espectro de CD de vários tipos de estrutura secundária. Os espectros representam α -hélice (—), folha- β (- - - -), voltas- β (····) e estrutura irregular (- · - ·). Modificada a partir de Kelly & Price, 1997.

A espectroscopia de CD em proteínas e peptídeos é uma ferramenta muito utilizada no estudo de mudanças conformacionais induzidas por interações da proteína com ligantes específicos e solventes, assim como verificar alterações ocorridas em função da temperatura, do pH, da força iônica e monitorar fenômenos de desnaturação e renaturação na interação com ligantes (CANTOR *et al.*, 1980).

A estimativa do conteúdo de estrutura secundária de uma proteína é feita com base num grupo de proteínas de referência que possuem estrutura secundária e espectros de CD conhecidos. Com o auxílio de programas computacionais, que utilizam diferentes métodos de desconvolução, são extraídas as componentes comuns dos espectros de CD da proteína analisada e a porcentagem que as mesmas representam no espectro (SREERAMA *et al.*, 2000).

As medidas de CD foram realizadas a temperatura controlada de 25 °C utilizando o espectropolarímetro JASCO J-720 (JASCO Corporation, Tóquio, Japão), na faixa de 198 a 250 nm em cubetas cilíndricas de quartzo de caminho óptico de 1 mm. Os espectros de CD foram tipicamente recuperados empregando-se médias de 16 varreduras. As contribuições dos tampões obtidos sob condições idênticas foram subtraídas e todos espectros de CD foram corrigidos a fim de eliminar quaisquer efeitos de ruído. Os espectros originais foram ainda filtrados com Transformadas de Fourier, preservando as bandas típicas de cada espectro. Em todos os casos, eles foram obtidos em elipticidade (θ) e, convenientemente transformados em elipticidade molar ($[\theta]$).

As soluções protéicas provenientes da purificação em coluna de afinidade, devidamente dialisadas, foram utilizadas para analisar a integridade das estruturas secundárias nas condições de extração e monitorar as possíveis mudanças conformacionais induzidas por influência do pH através das medidas de CD. A concentração das amostras foi de 10 μ M em tampão acetato-borato-fosfato de sódio 20 mM contendo NaCl 150 mM. As amostras foram incubadas em soluções tamponadas a pH variando de 3 a 10.

As contribuições de estrutura secundária pelo espectro de CD das proteínas recombinantes foram calculadas utilizando o pacote de desconvolução *CDPro software package*, que inclui três programas *SELCON3*, *CONTINLL* e *CDSSTR*, providos de um grupo de proteínas de referência contendo 43 espectros de CD (SREERAMA *et al.*, 2000; SREERAMA *et al.*, 1999; SREERAMA *et al.*, 1993). A partir dos resultados fornecidos pelos três programas, uma média destes resultados foi realizada para encontrar as frações de estruturas secundárias dos espectros de CD.

II. 2.14.6 - Espectroscopia de Emissão de Fluorescência

A absorção da radiação eletromagnética de um determinado comprimento de onda por um cromóforo, pode fazer com que seus elétrons passem do estado eletrônico fundamental para um estado eletrônico excitado. Muitas vezes, esta energia de excitação é perdida para o meio sob a forma de calor, promovendo o retorno do elétron para o estado fundamental. Quando este elétron retorna para o estado fundamental através do seu decaimento do menor nível de energia vibracional do estado eletrônico excitado (S_1) para o estado eletrônico fundamental (S_0) pela da emissão de um fóton, a fluorescência ocorre refletido por um comprimento de onda maior que o da excitação (CAMPBELL *et al.*, 1984; CANTOR *et al.*, 1980). A Figura 17 mostra o diagrama de Jablonski que representa este processo.

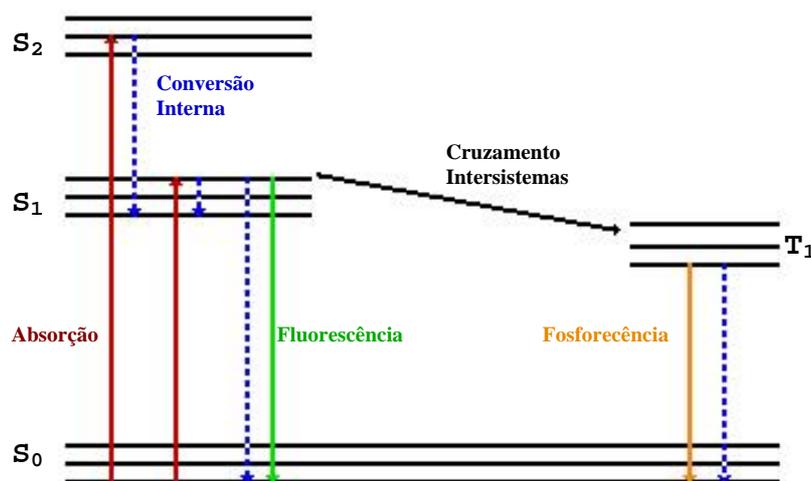


Figura 17. Diagrama de Jablonski. A absorção de determinada quantidade de energia faz com que um elétron passe do menor estado vibracional do estado fundamental (S_0) para um estado vibracional superior do estado excitado S_1 ou S_2 (setas vermelhas). Após passar por processos não radiativos (setas azuis), que conduzem o elétron ao menor nível vibracional do primeiro estado excitado, pode ocorrer emissão de fluorescência (seta verde), que é a emissão de um fóton devido ao decaimento do estado S_1 para o estado fundamental S_0 . A relaxação do elétron também pode ocorrer pela emissão de um fóton devido ao decaimento de um estado tripleto (T_1) para o estado eletrônico fundamental (S_0), processo chamado de fosforescência (seta laranja).

A espectroscopia de emissão de fluorescência estática também é largamente utilizada para caracterizar macromoléculas biológicas e verificar mudanças conformacionais discretas no microambiente dos grupos fluorescentes naturais (fluoróforos) como triptofano (Trp), tirosina (Tyr) e fenilalanina (Phe). A fluorescência destes resíduos é altamente sensível ao ambiente em que se encontram, portanto, é possível monitorar suas alterações por parâmetros como posição do máximo de emissão (λ_{\max}) e rendimento quântico (ϕ_F) devido à ação de solventes, ligantes, supressores, íons ou agentes solubilizantes (LAKOWICZ *et al.*, 1999).

As medidas de emissão de fluorescência das proteínas recombinantes foram realizadas no modo estático utilizando-se o espectrofluorímetro K2 (ISS Inc.,

Champaign, IL, EUA) acoplado a um sistema de banho refrigerador modelo RTE-111 (Neslab, Willow Brook Road, Freehold, NJ). Os espectros de fluorescência foram obtidos a temperatura controlada de 25 °C. Amostras foram excitadas a 295 nm e a emissão de fluorescência foi monitorada no intervalo de 300 a 450 nm. Cubetas retangulares de quartzo de 1 cm de caminho ótico foram utilizadas nas medidas. A fim de minimizar o efeito do espalhamento de luz, os espectros de emissão de fluorescência dos tampões foram subtraídos dos espectros das correspondentes amostras.

As proteínas recombinantes provenientes da purificação em coluna de afinidade também foram utilizadas nas medidas de fluorescência para analisar o microambiente dos resíduos de triptofano nas condições de extração. As possíveis mudanças conformacionais induzidas por influência do pH através das medidas de fluorescência também foram investigadas. A concentração das amostras foi de 37 µM em tampão de lise Tris-HCl.

II. 2.14.7 – Ensaios de Estabilidade Térmica e Química das Proteínas Recombinantes

II. 2.14.7.1 – Ensaio de Estabilidade Térmica Frente a Variações de pH

Os estudos de desnaturação térmica das proteínas recombinantes foram caracterizados por CD nos diferentes pHs medindo-se as mudanças de elipticidade a 222 nm induzidas pelo aumento de temperatura de 20 a 80 °C com taxa de aquecimento de 10 °C por hora.

Novos espectros de CD das amostras protéicas incubadas em soluções tamponadas a pH variando de 3 a 10 foram tomados após 12 horas de incubação à temperatura ambiente em cada um dos pHs escolhidos. A concentração das amostras protéicas para as medidas de CD foi mantida a 17 μ M.

O equipamento e a metodologia empregada nas análises de estabilidade térmica frente a variações nos pHs foram os mesmos descritos anteriormente.

II. 2.14.7.2 - Ensaio de Estabilidade Química Frente a Variações na Concentração de Uréia

O estudo da desnaturação química das proteínas recombinantes frente a diversas concentrações de um agente desnaturante também foram realizadas por ambas as técnicas espectroscópicas, CD e fluorescência. A desnaturação química foi monitorada empregando-se soluções protéicas a 5, 10 e 20 μ M de concentração em tampão de lise Tris-HCl contendo diferentes concentrações de uréia (0 a 6 M). Os estudos de desnaturação química, caracterizados por CD, foram medidos através das mudanças de elipticidade a 222 nm induzidas pelo aumento da concentração de uréia. Todos os experimentos foram levados a temperatura controlada de 25 °C. As contribuições dos tampões foram subtraídas de todos os experimentos.

Os equipamentos e as metodologias empregadas nas análises de estabilidade química frente à uréia foram os mesmos descritos anteriormente em cada técnica espectroscópica.

II. 2.14.8 – Estudos de Atividade Enzimática das Proteínas Recombinantes

O conhecimento teórico sobre cinética enzimática permitiu o desenvolvimento de ensaios quantitativos de atividade enzimática. Utilizando métodos analíticos relativamente simples como a espectrofotometria foi possível analisar a quantidade dos produtos formados ou consumidos pela adição das enzimas a uma solução tamponada com concentração conhecida de substrato. As reações ocorreram em tempos fixos, em condições determinadas de temperaturas e pHs. O uso de cofatores foi empregado quando necessário.

II. 2.14.8.1 – Hidroxinitrila Liase de *Xylella fastiosa* (rXfHNL)

A atividade enzimática da rXfHNL foi determinada seguindo a clivagem do substrato, mandelonitrila racêmica, em benzaldeído e ácido cianídrico (HCN) a 25 °C. A formação do benzaldeído foi monitorada espectrofotometricamente a 280 nm, comprimento de onda na qual a absorção da mandelonitrila pode ser negligenciada, já que não é o comprimento do máximo de absorção da mesma.

As medidas de atividade foram realizadas utilizando-se o espectrofotômetro UV-Vis Hitachi, modelo U-2800, acoplado a um computador para aquisição e tratamento dos dados. As cubetas utilizadas para monitorar a reação eram retangulares, providas de tampa, de quartzo, com 1 cm de caminho ótico e capacidade de 1,5 mL. Todas as soluções foram mantidas a 25 °C durante a dosagem e as amostras foram analisadas em triplicata.

A mistura reacional foi preparada utilizando o tampão fosfato de sódio a 50 mmol L⁻¹ contendo NaCl a 150 mM nos pHs 5 e 7. As concentrações finais de mandelonitrila e rXfHNL na mistura reacional foi de 0,5 mM e 0,05 µM, respectivamente. A velocidade da reação ou de formação dos produtos foi acompanhada por um intervalo de 8 min. Tampão fosfato de sódio contendo mandelonitrila racêmica foi utilizado como branco das amostras para zerar o espectrofotômetro. A atividade da enzima rXfHNL foi monitorada na ausência e na presença do cofator FAD.

II. 2.14.8.2 – Corismato Sintase de *Xylella fastiosa* (rXfaroC)

A atividade enzimática da rXfaroC foi determinada seguindo a formação do corismato a partir da do ácido corísmico em presença de NADH e FMN a 25 °C. A formação do corismato foi monitorada espectrofotometricamente a 340 nm.

As medidas de atividade foram realizadas utilizando-se o espectrofotômetro UV-Vis Hitachi, modelo U-2800, acoplado a um computador para aquisição e tratamento dos dados. As cubetas utilizadas para monitorar a reação eram retangulares, providas de tampa, de quartzo, com 1 cm de caminho ótico e capacidade de 1,5 mL. Todas as soluções foram mantidas a 25 °C durante a dosagem e as amostras foram analisadas em triplicata.

A mistura reacional foi preparada utilizando o tampão fosfato de potássio a 50 mM contendo NaCl a 150 mM no pH 7,8. As concentrações finais de ácido corísmico e rXfaroC na mistura reacional foi de 0,4 mM e 0,1 µM, respectivamente. A velocidade da reação ou de formação dos produtos foi acompanhada por um intervalo de 8 min. O Tampão fosfato de potássio contendo NADH e FMN foi

utilizado como branco das amostras para zerar o espectrofotômetro. A atividade da enzima rXfaroC foi monitorada na ausência e na presença do cofator NADH.

III – RESULTADOS E DISCUSSÃO

III. 1 – Hidroxinitrila Liase

III. 1.1 – Amplificação e Clonagem no Vetor pGEM-T *easy*

As suspensões bacterianas utilizadas na extração do DNA genômico foram do isolado 9a5c da bactéria *X. fastidiosa*. Inicialmente, a qualidade e a concentração do DNA genômico obtido foi visualmente medida através da eletroforese em gel de agarose 0,8 % e, subseqüentemente, confirmada através de leituras em espectrofotômetro (A_{260}/A_{280}), como descrito em materiais e métodos. A metodologia empregada segundo Ausubel *et al.* (AUSUBEL *et al.*, 1995), permitiu a obtenção de um DNA puro e de boa qualidade com rendimentos totais de 100 a 500 ng μL^{-1} suficiente para os experimentos seguintes.

Como mencionado anteriormente, os oligonucleotídeos utilizados na amplificação foram desenhados com base na seqüência da ORF SCJ21.16 e tiveram sítios de restrição para as enzimas *NdeI* e *XhoI* incorporados preservando seus códons originais de iniciação e terminação. Estes sítios de restrição nos iniciadores empregados para a amplificação foram inseridos de forma a permitir sua subclonagem nos vetores de expressão. Adicionalmente, a temperatura de anelamento do par de iniciadores (F e R) desenhado foi calculada teoricamente através do software Gene Runner 3.05 (Hastings Software Inc., EUA) a fim garantirmos que estes apresentassem temperaturas próximas e evitassem assim a formação de estruturas indesejáveis como dímeros e grampos que podem não só acarretar em suas auto-complementariedade, resultando na obtenção de

subprodutos, como também provocar perda de eficiência da reação de amplificação do DNA de interesse. A temperatura ótima encontrada ficou em torno dos 54 °C.

Com base na temperatura de hibridização calculada para o par de oligonucleotídeos desenhado foi possível a amplificação do fragmento correspondente à proteína hidroxinitrila liase, porém com baixa concentração. Nesta temperatura bandas inespecíficas estiveram presentes em ambas as extremidades do gel utilizado para avaliar as amostras amplificadas. Estas foram atribuídas a prováveis excessos de dNTPs, DNA genômicos e eventuais subprodutos na reação de amplificação. Diante deste primeiro resultado, fez-se necessário um processo de otimização da PCR a fim de garantir o máximo de eficiência na obtenção dos produtos desejados. Assim, somente quando foi empregada a temperatura de hibridização dos *primers* de 52 °C foi possível a amplificação desejada.

A Figura 18 ilustra o resultado desta amplificação a qual gerou o produto de 850 pares de bases. O produto de amplificação obtido pela PCR teve o seu tamanho concordante com o previsto, não apresentou nenhuma banda inespecífica, indicando o sucesso da otimização e resultou em quantidades suficientes para procedermos a posterior clonagem.

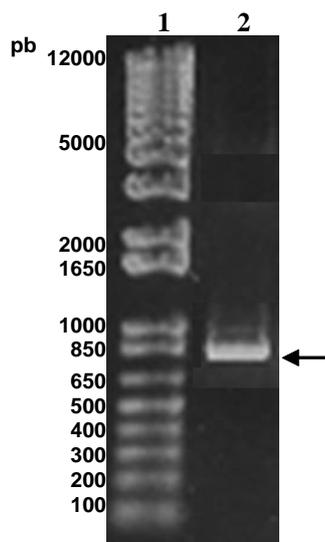


Figura 18. Amplificação da ORF codificadora da proteína hidroxinitrila liase. A coluna 1 é o padrão de massa molecular 1 kb Plus DNA Ladder da Gibco BRL. A coluna 2 é a ORF SCJ21.16 amplificada. O meio de separação foi gel de agarose 0,8 % em tampão 1X TAE, pH 8,3. A voltagem aplicada na separação foi 100 V por 30 min. Os fragmentos de DNA no gel foram visualizados em transiluminador UV a 365 nm.

Os produtos da amplificação por PCR foram devidamente purificados pelo *kit Wizard PCR Clean-up* da Promega, quantificados em espectrofotômetro e adenilados para inseri-los no vetor de clonagem pGEM-T *easy*. A concentração obtida para estes produtos variou de 500 até 700 ng μL^{-1} . A construção nos vetores de clonagem foi denominada de pGEM-XfHNL. A escolha deste sistema foi empregada para propagação baseada no fato deste vetor ser muito conveniente para clonar produtos de PCR facilitar a seleção dos clones positivos. Este vetor apresenta 3018 pb e contém duas marcas de seleção ao longo de sua seqüência, um gene que confere resistência ao antibiótico ampicilina e outro gene que codifica para a proteína β -galactosidase localizado no sítio de múltipla clonagem.

Os plasmídeos recombinantes pGEM-XfHNL recém construídos foram inseridos por choque térmico em bactérias de *E. coli* DH5 α competentes. As

bactérias transformadas foram semeadas em placas de Petri contendo meio LB, ágar, ampicilina, X-gal e IPTG e crescidas *overnight* em estufa por incubação a 37 °C. A Figura 19 mostra um exemplo dos resultados obtidos com a transformação das bactérias DH5 α .

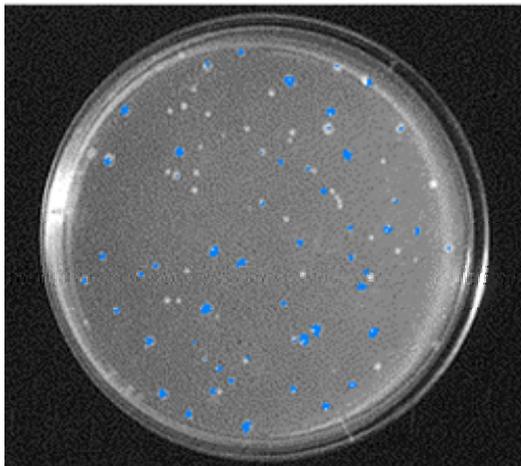


Figura 19. Colônias Recombinantes e Não Recombinantes de Células *E. coli* DH5 α . Células competentes DH5 α foram transformadas com os construídos recombinantes pGEM-XfHNL, plaqueadas em meio LB contendo ágar, ampicilina, X-gal e IPTG e incubadas *overnight* a 37 °C. As colônias que tiveram o gene da β -galactosidase interrompido apresentaram coloração branca e as colônias que tiveram o gene intacto apresentaram coloração azul. As colônias brancas são os prováveis clones positivos.

Diversas colônias brancas e azuis cresceram na placa contendo o meio selecionado descrito acima. A seleção dos transformantes foi realizada então pela capacidade de crescer em presença do antibiótico ampicilina e através da coloração branca de algumas colônias transformantes indicativa da presença do inserto dentro da região codificadora da enzima β -galactosidase. Dentre as colônias brancas resistentes, cinco foram escolhidas aleatoriamente, inoculadas em tubos estéreis contendo meio LB líquido com ampicilina e submetidas à incubação

overnight a 37 °C sob agitação. Após o crescimento individual destas colônias seus DNAs plasmidiais foram extraídos, quantificados, clivados por enzimas de restrição para checar a presença do inserto e, finalmente, seqüenciados.

III. 1.1.1 – Análise de Restrição e Seqüenciamento dos Plasmídeos Recombinantes

Os DNAs plasmidiais dos 5 prováveis clones positivos, extraídos através dos métodos de lise alcalina e pelo *kit Wizard Miniprep*[®] (descritos em II.2.7), foram inicialmente considerados de boa qualidade e adequados para as análises de restrição e seqüenciamento. É importante ressaltar que em algumas das análises de restrição ocorreu baixa eficiência da clivagem pela enzima *NdeI*, a qual requer geralmente DNAs livre de impurezas para que ocorra uma boa digestão enzimática. Sendo assim, sempre que necessário foi empregado para estas análises DNA obtido pelo protocolo de extração através kit. As amostras de DNA plasmidial extraídos pelo kit e pelo método de lise alcalina apresentaram em geral concentrações de 200 e 150 ng μL^{-1} , respectivamente.

A análise de restrição empregando as enzimas *NdeI* e *XhoI* foi realizada a fim de confirmar a presença do inserto nos clones selecionados. Nesta análise, o tamanho do inserto foi avaliado pela clivagem do plasmídeo em sítios de restrição adjacentes ao sítio de clonagem, com posterior eletroforese em gel de agarose. A Figura 20 apresenta o perfil de restrição destes cinco clones selecionados.

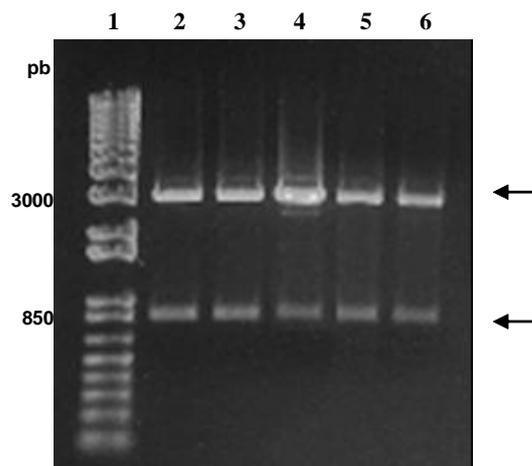


Figura 20. Separação dos plasmídeos digeridos pGEM-XfHNL. A coluna 1 é o padrão de massa molecular 1 kb Plus DNA Ladder da Gibco BRL. As colunas 2 a 6 são os clones positivos selecionados a serem utilizados na etapa de subclonagem nos vetores de expressão. O meio de separação foi gel de agarose 1 % em tampão 1X TAE, pH 8,3. A voltagem aplicada na separação foi 100 V por 30 min. Os fragmentos de DNA no gel foram visualizados em transiluminador UV a 365 nm. As bandas específicas dos genes XfHNL e dos plasmídeos digeridos estão representados por setas cujos tamanhos são aproximadamente 3000 e 850 pb, respectivamente.

A presença e o tamanho dos insertos confirmaram o sucesso das etapas de clonagem e transformação das bactérias. Os produtos da digestão de 850 pb correspondente a ORF XfHNL foram retirados do gel, purificados utilizando-se novamente o *kit Wizard PCR Clean-up*, quantificados e utilizados na etapa de subclonagem nos vetores de expressão. Estas amostras de DNA apresentaram em geral concentrações variando de 100 a 150 ng μL^{-1} .

As seqüências resultantes do gel de seqüenciamento foram analisadas através do programa *Sequencing Analysis V 3.4* que capturou automaticamente os dados coletados de cada amostra analisada no seqüenciador e extraiu as bases de cada seqüência arquivando-as em arquivos textos e em forma de eletroferogramas.

A Figura 21 ilustra a seqüência resultante de um dos clones submetidos ao seqüenciamento.

```

TCNATNNNTT GCGTCATATG CTCCNNGCCG CNTGGNNGCC GCGGGAANNC AATTCATAT GAACGCATAC
CGTCTATCG CTTTGCCTC ATCCGCAGCA TCTGATAAGC CAACCGTTT ACTGAAACCG ACAATTGTGC
TCGTACACGG TGCTTTGCT GATGGTTCTA CATGGAACAA AGTAATCCAT CAGCTCCAGG CAAAAGGGCT
GAATGCTGTG TCCGNTCAA ATCCGCTGAC GTCTTTCGAA GATGATGTAG CAGCAACAGC CCGTGTGCTT
GCTCTGAAA CGGGACCTGT AGTGCTCGTT GGCCACTCTT ATGGGGGAGC AGTGATTACT GAAGCTGGAC
AAGATGAACG AGTTAAAGCC CTAC-TTATA TTGCCGCATT TGCTCCATCC GAGGGAGAG TCGGTTGCTG
ATCTGAATAA AAAATATCCA TTACCTTCTG GATATAATCA TCTCAGCAGT GACAAAGAAG GGTTCTGAT
GCTAACGCCA GAAGGTGTGG AAAAATACTT GGCGCAGGAT ATCCCCCTG AGCAAACCTG CCTGATCATC
GCCACTCAGC ACCCTATACG CGGCGCCAAT TTTGAAAAA AGGTTTCAGC TGCCGCTTGG GAAACTAAAC
CATCATGGTA TTGNTGAGT GAGAATGATC TTATNCTTCA GCCAGCGCTG CAAAAAGAGA TGGCTCAAAA
AATAGGTGCT CATATCGTCA ATGTGGCAGC CAGCCATGTG CCACATCTT CACATGCTGC GGAAGTAGAA
AAAATCATCA GAGTGCAGT CAACGATGTT GAAGTATTAG AATAACTCGA GCGGAATCAC TAGTGATATC
GCGGACACCT GCAGCTCTAG CATATGGAGA GCTCCCATNG NGTTNGNN

```

Figura 21. Bases seqüenciadas do plasmídeo pGEM-XfHNL extraído do clone 1. As bases em vermelho correspondem a ORF SCJ21.16 que codifica para proteína hidroxinitrila liase. As bases em negrito e sublinhado são correspondentes às seqüências dos oligonucleotídeos iniciadores. Os nucleotídeos em negrito preto foram são os não identificados pelo seqüenciador.

As seqüências resultantes foram submetidas a um processo de busca por similaridade em diversos bancos de dados internacionais, através de servidores de busca disponíveis na rede. O principal programa utilizado foi o software público BLAST, cujos resultados mostraram similaridade de seqüência variando de 90 a 95% de identidade frente à cadeia nucleotídica correspondente a ORF deste estudo, confirmando a presença da região alvo desejada.

De modo geral os clones seqüenciados apresentaram pequenas variações entre si, que podem ser atribuídos a erros de leitura ou mesmo a erros gerados durante a marcação para o seqüenciamento ou na amplificação. No entanto, os

plasmídeos pGEM-XfHNL destes cinco clones apresentaram os genes de interesse inseridos na orientação e com fase de leitura corretas dentro dos vetores pGEM-T *easy* garantindo assim a integridade destes clones e a posterior leitura eficiente de todas as bases dos fragmentos na etapa de expressão protéica.

III. 1.2 – Subclonagem no Vetor de Expressão e Análise de Restrição dos Plasmídeos Recombinantes

Os fragmentos de DNA de 850 pb que codificam para a proteína rXfHNL foram subclonados nos vetores pET28a. Esta nova construção denominada pET-XfHNL foi escolhida para expressar a proteína de interesse fusionada com uma cauda de seis histidinas. Novamente, os plasmídeos recombinantes recém construídos foram inseridos por choque térmico em bactérias de *E. coli* DH5 α competentes, semeadas em placas semelhantes ao processo anterior, porém somente com adição de canamicina, o novo antibiótico de seleção, e finalmente incubadas para o crescimento. A seleção das colônias neste caso foi realizada unicamente pela análise de restrição, já que o vetor pET28a não apresenta genes que as tornem de coloração diferenciada.

Após o período de crescimento celular, novamente cinco colônias foram selecionadas aleatoriamente dentre as muitas colônias presentes na placa de Petri, e inoculadas seguindo o mesmo processo descrito anteriormente. Após o crescimento individual destas células, seus DNAs foram extraídos, quantificados e uma alíquota destes DNAs foi clivada por enzimas de restrição. Estas amostras de DNA também apresentaram concentrações variando de 100 a 150 ng μL^{-1} . A Figura 22 ilustra o novo perfil de restrição destes cinco clones selecionados.

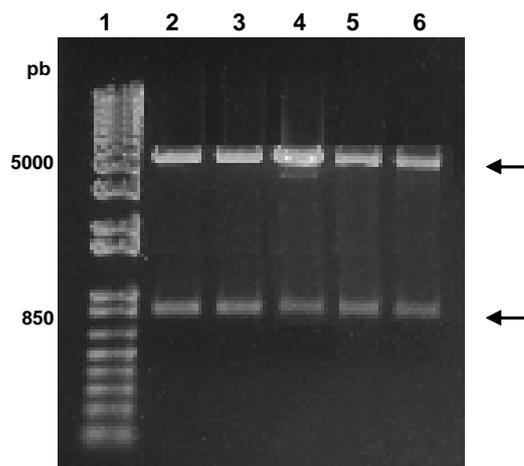


Figura 22. Separação dos plasmídeos digeridos pET-XfHNL. A coluna 1 é o padrão de massa molecular 1 kb Plus DNA Ladder da Gibco BRL. As colunas 2 a 6 são os clones positivos. O meio de separação foi gel de agarose 1 % em tampão 1X TAE, pH 8,3. A voltagem aplicada na separação foi 100 V por 30 min. Os fragmentos de DNA no gel foram visualizados em transiluminador UV a 365 nm. As bandas específicas dos genes XfHNL e dos plasmídeos digeridos estão representados por setas.

A presença dos insertos em todos os clones selecionados confirmou o sucesso das etapas de subclonagem e transformação das bactérias. Uma vez confirmada a presença do inserto, apenas os clones 1 e 2 (equivalentes as 2^o e 3^o canaletas do gel) os quais apresentaram maior intensidade de banda, foram utilizados para transformar as células competentes de expressão de *E. coli* BL21(DE3). A seleção de dois novos clones contendo as ORFs subclonadas nos vetores de expressão foi realizada após o plaqueamento das células em meio adequado, e estes clones foram empregados na produção da proteína expressa de forma heteróloga.

III. 1.3 – Expressão Heteróloga da Proteína rXfHNL a partir dos Vetores de Expressão

A produção da proteína recombinante foi realizada inicialmente por meio da indução da expressão por IPTG a 37 °C. Visando otimizar o período de expressão protéica, alíquotas foram retiradas de hora em hora para a quantificação da expressão de modo a obter quantidades adequadas da proteína recombinante a serem empregadas nas etapas de caracterização. Passado o período de indução de 4 horas, as células bacterianas resultantes foram recuperadas por centrifugação e lisadas na presença de tampão de lise tris-HCl. Os extratos protéicos foram avaliados quanto a sua solubilidade em SDS-PAGE, em tampão com condições próximas à fisiológica.

A Figura 23 mostra um gel de SDS-PAGE a 15% obtido após submeter às amostras a condições desnaturantes.

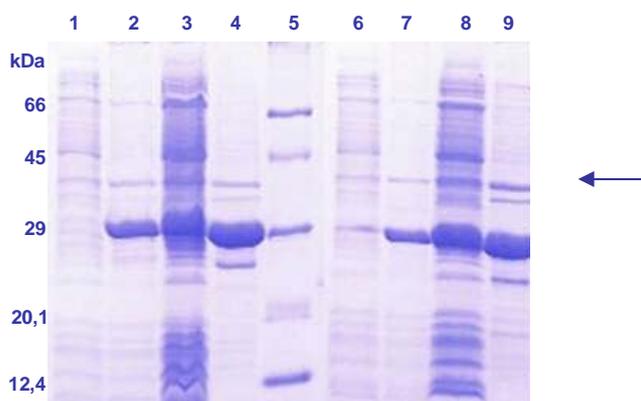


Figura 23. Expressão da proteína rXfHNL a partir dos clones 1 e 2. A análise foi realizada em SDS-PAGE a 15%. As colunas 1 a 4 e 6 a 9 representam as proteínas das expressões dos clones 1 e 2 em BL21(DE3) pET-rXfHNL, respectivamente. As colunas 1 e 6 são os extratos protéicos totais não induzido dos sistemas de expressão. As colunas 2 e 7 são os extratos protéicos totais induzidos. As colunas 3 e 8 são as frações insolúveis pós lise celular. E as colunas 4 e 9 são as frações solúveis pós lise celular. A coluna 5 é o marcador de massa molecular. A seta indica a expressão diferenciada esperada na região de 29 kDa.

A presença de banda diferenciada nas frações induzidas 2 e 7 tem o tamanho esperado para a proteína rXfHNL, aproximadamente 30 kDa, revelando sucesso na metodologia de expressão adotada.

Além disso, após a lise celular foi constatado que a proteína se apresentou em grande quantidade nas frações solúveis com o emprego do tampão tris-HCl, apesar de ter sido também observada a permanência da proteína na fração insolúvel do sistema de expressão, indicativo de que um processo de otimização da lise celular poderia ser empregado durante os experimentos.

Porém, com o rendimento aparentemente satisfatório da forma solúvel, apenas uma adaptação do procedimento de lise celular foi empregado para aumentar a quantidade da proteína recombinante na fração solúvel. Esta adaptação envolveu o aumento dos ciclos de sonicação e conseqüentemente de descanso do procedimento de lise celular. As frações solúveis foram, então, empregadas na etapa de purificação da proteína rXfHNL.

A produção em maior escala da proteína em *E.coli* foi realizada depois de otimizado o processo de indução, cuja quantidade foi suficiente tanto para as etapas de purificação quanto para os subseqüentes estudos estruturais e funcionais.

III. 1.4 – Purificação da Proteína rXfHNL por Cromatografia de Afinidade e Exclusão por Tamanho

As frações solúveis das culturas induzidas provenientes da lise celular foram submetidas à cromatografia de afinidade utilizando uma resina contendo níquel imobilizado. As frações eluídas durante o processo cromatográfico foram

analisadas por SDS-PAGE. A Figura 24 mostra a separação eletroforética em gel de poliacrilamida SDS-PAGE a 15% dos eluídos da coluna de Ni-NTA.

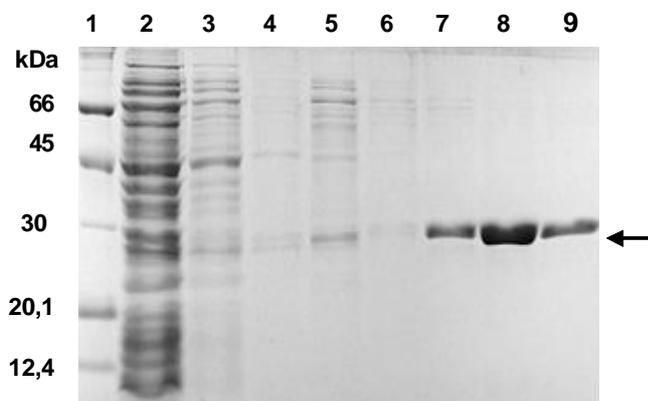


Figura 24. Purificação da rXfHNL por cromatografia de afinidade. A análise das frações eluídas da coluna Ni-NTA foi realizada em SDS-PAGE a 15%. A coluna 1 é o marcador de massa molecular. A coluna 2 representa a fração não ligada na coluna de afinidade. As colunas 3 e 4 representam as frações de lavagem com 5 mM de imidazol. As colunas 5 e 6, as frações de lavagem com 40 mM de imidazol. As colunas 7, 8 e 9, as frações de eluição com 100, 250 e 500 mM de imidazol, respectivamente. As frações eluídas representam a proteína rXfHNL purificada e estão indicadas pela seta.

A purificação da proteína recombinante rXfHNL foi alcançada em um único passo cromatográfico. A pureza foi confirmada pelas frações eluídas com tampão de lise trisHCl contendo 100, 250 e 500 mM de imidazol nas quais apresentaram uma única banda de aproximadamente 30 kDa, sob condições desnaturantes.

Em algumas das purificações realizadas neste estudo eventuais proteínas contaminantes foram observadas nas frações de eluição. Para tentar contornar a presença de contaminantes após a purificação em resina de afinidade, a estratégia de re-cromatografia de afinidade e cromatografia por exclusão molecular foram empregadas quando necessário.

Em algumas ocasiões a simples re-cromatografia por afinidade não foi suficiente para se obter a proteína pura o suficiente (resultados não mostrados). Desta forma, a cromatografia de exclusão por tamanho foi empregada não só a fim de permitir a obtenção de amostras mais puras como também prever o provável grau de oligomerização da proteína rXfHNL frente a calibração da coluna com três padrões de massa molecular.

O perfil de eluição da proteína rXfHNL comparado com os perfis dos padrões é mostrado na Figura 25: Os padrões empregados foram soro albumina bovina (66 kDa), ovalbumina (45 kDa), anidrase carbônica (30 kDa), α -lactoalbumina (20,1 kDa) e inibidor de tripsina (14,4 kDa).

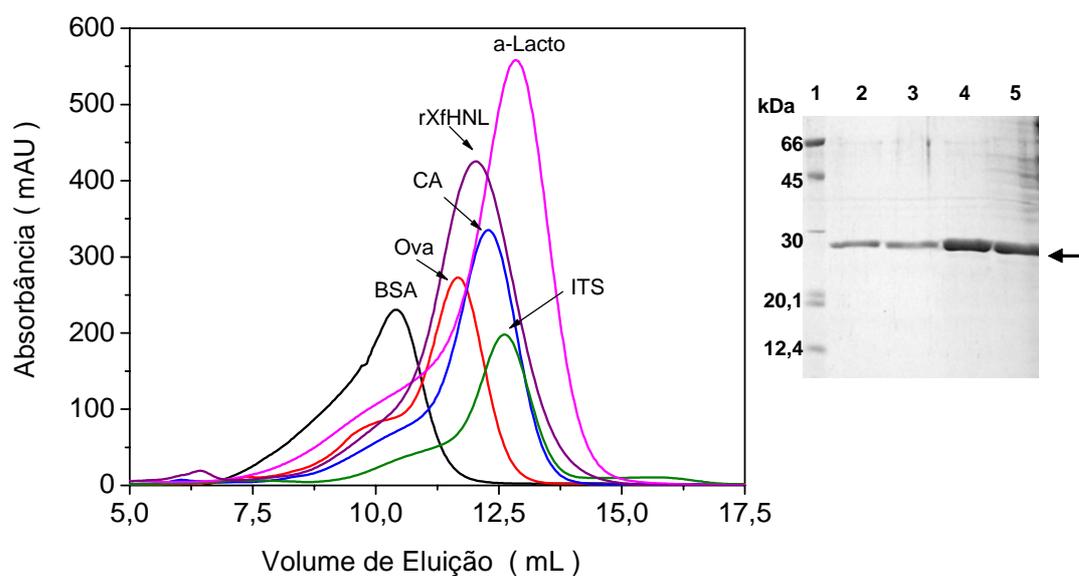


Figura 25. Perfil cromatográfico da rXfHNL por cromatografia de exclusão por tamanho e SDS-PAGE das frações eluídas. A separação de todas as amostras foi realizada individualmente em coluna Superdex 75 HR 10/30 a 25 °C empregando o tampão de lise tris-HCl a pH 8,0. As frações eluídas da cromatografia por exclusão molecular foram recolhidas e aplicadas em SDS-PAGE a 15% corado com azul de Comassie. A coluna 1 é o marcador de massa molecular. As colunas 2 a 5 são os volumes recolhidos de eluição de 10 a 13 mL. A seta aponta para a banda correspondente à forma monomérica de aproximadamente 30 kDa da proteína de fusão.

A proteína rXfHNL submetida da cromatografia de exclusão por tamanho foi eluída como um único pico representado pelo volume de eluição próximo a 12 mL. O emprego desta técnica mostrou-se capaz de separar a fração correspondente à proteína rXfHNL em fração distinta daquelas dos contaminantes e do imidazol, eluídos entre os volume de 6 a 8 mL e 16 a 25 mL, respectivamente. A eliminação do excesso de imidazol comumente presente após cromatografia de afinidade permitiu que a proteína de interesse fosse utilizada diretamente nas etapas de caracterização sem a necessidade de empregarmos a diálise para sua remoção, o que foi muito benéfico em termos de estabilidade protéica.

Além disso, a análise do perfil cromatográfico sugeriu inicialmente que o grau de oligomerização da proteína rXfHNL é aparentemente monomérico devido a presença de apenas um pico próximo a região de eluição do padrão anidrase carbônica (CA) de 30 kDa de massa molecular. Era esperado que a presença de prováveis formas diméricas da proteína em estudo apresentasse um pico nas regiões correspondente ao dobro desta massa, ou seja, próximo ao padrão soro albumina bovina de 60 kDa.

A eluição de uma macromolécula em experimentos de exclusão por tamanho é usualmente dado pelo coeficiente de partição ou de distribuição (Kd) da proteína entre as fases, móvel e estacionária. Na prática, Kd é difícil de determinar e é substituída pelo K_{av} uma vez que há uma relação constante entre K_{av} e Kd. K_{av} representa os volumes de eluição relativos e é um parâmetro calculado como

$$K_{av} = \frac{V_e - V_0}{V_T - V_0} \quad (1)$$

onde V_e é o volume de eluição da proteína, V_0 é o volume morto da coluna e V_T é volume total do gel na coluna (OHNO *et al.*, 1968; ACKERS *et al.*, 1967; LAURENT *et al.*, 1964). O V_0 foi determinado utilizando *Dextran Blue 2000*. Então, a massa molecular da proteína rXfHNL foi determinada pela interpolação na regressão linear da curva padrão através dos valores de K_{av} . A Figura 26 mostra o gráfico do coeficiente de partição K_{av} plotado como uma função da massa molecular dos padrões de proteínas.

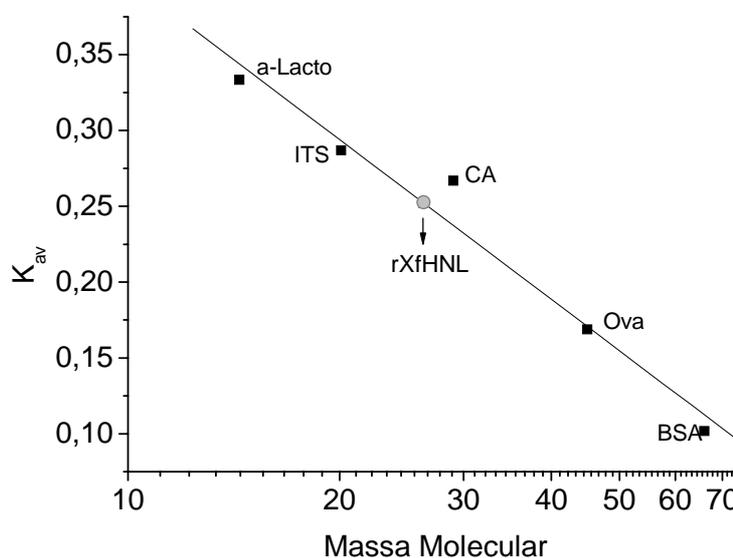


Figura 26. Determinação do grau de oligomerização aparente da rXfHNL por cromatografia de exclusão por tamanho. O peso molecular da rXfHNL foi determinado por regressão linear da curva de calibração dos padrões pelos valores dos seus respectivos K_{av} . Os volumes de eluição foram expressos em termos de K_{av} que foi determinado para cada proteína padrão utilizando a expressão $(V_e - V_0)/(V_T - V_0)$. A separação de todas as amostras foi realizada individualmente em coluna Superdex 75 HR 10/30 a 25 °C empregando o tampão de lise tris-HCl a pH 8,0.

A curva de calibração resultante revelou uma relação linear ($y = 0,74921 - 0,34994X$; $R = 0,984$) entre o coeficiente de partição e a massa molecular de cada

proteína padrão representado na Figura 25, como era esperado para um mecanismo de separação puramente baseado no tamanho. Em adição, o alto coeficiente de correlação sugere o potencial para se estimar a massa molecular baseada na curva de calibração. Estes resultados também sugerem que a proteína nativa existe como monômero. A estrutura monomérica está em acordo com o observado pra os membros da família das HNLs independentes de cofator. Durante a cromatografia de exclusão nenhum outro pico foi observado sugerindo a homogeneidade da proteína purificada.

Sendo assim, a purificação por cromatografia de exclusão por tamanho foi considerada um procedimento adotado sempre que necessário.

III. 1.5 – Análises Físico-química e Estrutural da Proteína rXfHNL

III. 1.5.1 – Determinação Quantitativa do Nível de expressão Protéica

Para testar o nível de expressão da rXfHNL, as alíquotas retiradas da cultura induzida em intervalos fixos de tempo foram analisadas por SDS-PAGE e GelChip-CE. A Figura 27 mostra os resultados analíticos desta expressão gerados por SDS-PAGE e pelo equipamento *Bioanalyzer* 2100 da Agilent. Os resultados fornecidos por este último foram também dispostos como eletroferogramas, porém não foram mostrados devido ao melhor efeito de comparação proporcionado entre pela imagem na forma de gel virtual.

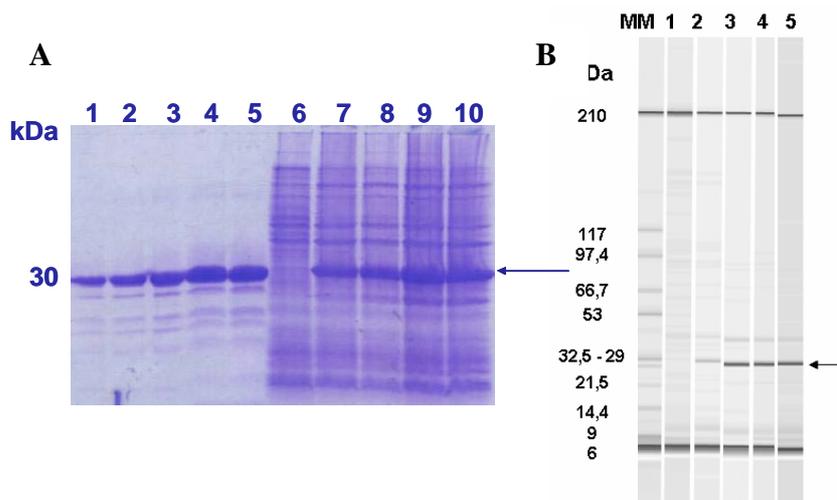


Figura 27. Géis representativos das análises por (A) SDS-PAGE e (B) GelChip-CE. A detecção em (B) foi baseada na fluorescência induzida a laser de um corante fluorescente, no qual foi adicionado na matriz de separação. As amostras foram separadas individualmente em uma solução polimérica apropriada fornecida pelo kit Protein 200 plus LabChip, como recomendado pelo fabricante. MM é o padrão de massa molecular. As colunas 6 (em A) e 1 (em B), representam os extratos protéicos totais não induzidos dos sistemas de expressão. As colunas 7 a 10 (em A) e 2 a 5 (em B) são as alíquotas dos extratos protéicos totais após 1, 2, 3 e 4 h de indução, respectivamente. As colunas 1 a 5 são as diversas concentrações do padrão anidrase carbônica. As bandas a 6 e 210 kDa no gel B são os marcadores que delimitam as separações das amostras por Gel-ChipCE. As setas indicam as bandas específicas da proteína rXfHNL.

Quando analisado qualitativamente o gel SDS-PAGE, o maior nível de expressão da rXfHNL aparentemente foi detectado após 2 horas do início do processo de indução. Já na análise por *microchip* uma observação visual não mostrou grandes diferenças com relação ao resultado prévio. Com a intenção de realizar uma comparação mais minuciosa entre estes dois resultados, a quantificação dos níveis de expressão da rXfHNL em SDS-PAGE foi analisada por densitometria ótica já que as análises por *microchip* forneceram a quantificação em tempo real de cada separação. A Figura 28 mostra o perfil da quantificação da expressão da proteína rXfHNL pelas duas técnicas eletroforéticas.

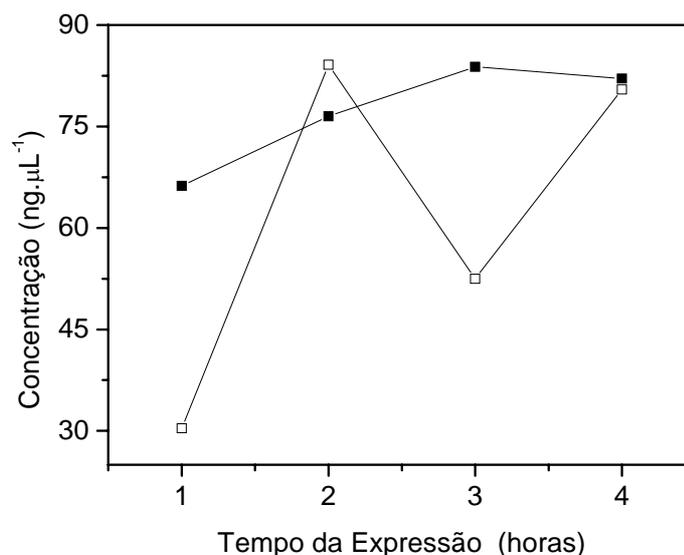


Figura 28. Resultado da expressão da rXfHNL em função do tempo por GelChip-CE e SDS-PAGE. As amostras analisadas por Gel Chip-CE foram separadamente eletroforesadas e a concentração e tamanho de cada banda separada foi automaticamente calculada pelo *software*. Os resultados foram fornecidos em tempo real. As intensidades das bandas apresentadas por SDS-PAGE foram quantificadas por densitometria empregando-se o *software Kodak 1D image analysis* e anidrase carbônica como padrão quantitativo. Os quadrados em branco representam os dados da SDS-PAGE, enquanto os quadrados em preto representam os resultados do Gelchip-CE.

A concentração protéica relativa de cada banda do GelChip-CE foi obtida baseado na medida da área de cada pico dos eletroferogramas simultaneamente gerados com o gel e comparadas a área de pico do marcador superior. O marcador superior, uma proteína de concentração conhecida, serviu como padrão interno. Para as análises do gel SDS-PAGE, a concentração protéica relativa foi baseada na curva de calibração obtida utilizando diferentes quantidades do padrão anidrase carbônica (CA).

Como pode ser visto na Figura 28, diferenças significativas nos níveis de expressão da rXfHNL foram observadas nestes métodos de quantificação. O gel de SDS-PAGE proporcionou resultados consistentes com as análises qualitativas

descritas anteriormente. O maior nível de expressão da rXfHNL (aproximadamente 84 ng μL^{-1}) foi alcançado duas horas após a indução para o gel SDS-PAGE. No entanto, as análises por GelChip-CE demonstraram que a expressão mais pronunciada foi após três horas de expressão protéica.

Esta diferença pode ser explicada pelo fato do gel SDS-PAGE ser um método semiquantitativo. Além disso, sob as condições deste método, a concentração de rXfHNL (principal banda) se apresenta fora da faixa dinâmica linear do ensaio. Quando a concentração dos contaminantes estava ainda em uma variação linear, a saturação na coloração da banda principal foi alcançada muito antes da saturação das outras bandas correspondentes aos contaminantes. Portanto, a porcentagem de rXfHNL encontrada pode ser considerada superestimada. O método Gel-Chip-CE permite uma quantificação mais exata através de parâmetros calculados automaticamente e proporcionados pelo *software*.

Assim, os tempos de indução para a produção em larga escala da rXfHNL foram limitados a três horas. A massa molecular aparente da banda correspondente a rXfHNL é de aproximadamente 30,4 kDa. Levando-se em conta a inclusão da cauda adicional de 6 histidinas, esta massa molecular foi considerada mais próxima daquela obtida por SDS-PAGE e coerente com a massa teórica de 29,1 kDa proporcionada pelo software *Protein Parameters Tools* (www.bo.expasy.org/tools/protparam.html) obtida através da seqüência dos aminoácidos.

Nenhum aumento no nível de expressão protéica foi observado quando diferentes temperaturas de indução foram testadas. A proteína recombinante rXfHNL foi sempre expressa na forma solúvel e a expressão como corpos de inclusão não foi observada.

III. 1.5.2 – Determinação do Ponto Isoelétrico

Inicialmente, o pI teórico para a proteína rXfHNL foi previsto pelo software *ProtParam* disponível online (<http://ca.expasy.org/tools/protparam.html>). Este software é uma ferramenta que permite calcular teoricamente vários parâmetros químicos e físicos através da seqüência das proteínas. O valor teórico encontrado para a rXfHNL foi de 5,85. A partir desta estimativa, o ensaio de focalização isoelétrica foi realizado com a proteína rXfHNL para confirmar seu pI experimental. A Figura 29 mostra o resultado experimental deste ensaio.

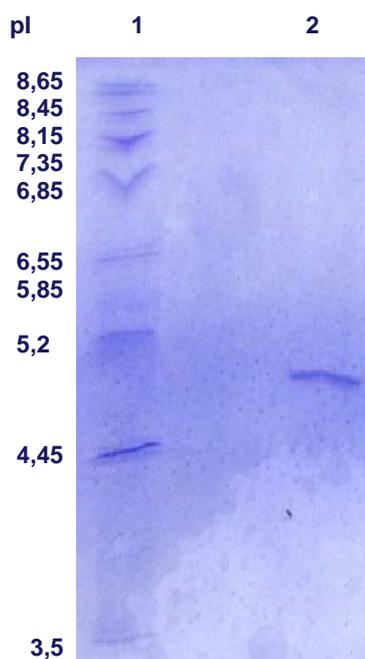


Figura 29 - Ensaio de focalização isoelétrica da rXfHNL. O ensaio foi realizado no *PhastSystem* gel numa faixa de pH de 9,0 a 3,0. A coluna 1 representa o padrão de pI e a 2 a amostra rXfHNL.

Como esperado, a presença da proteína foi observada na região próxima do pI teórico. No entanto, o ensaio realizado indicou que o pI experimental da

proteína rXfHNL é 4,6. Esta característica ácida no valor do pI desta proteína é uma característica apresentada para ambas as classes de hidroxinitrila descritas na literatura. Através deste resultado pode-se apenas prever que a rXfHNL apresenta uma das características inerentes a esta família de proteínas e pode estar incluída em qualquer umas das duas classes de proteínas.

III. 1.5.3 – Determinação da Seqüência por Espectrometria de Massas

A fim de verificar a identidade seqüencial da proteína recombinante, a rXfHNL foi analisada por espectrometria de massas. A metodologia para determinar a seqüência aminoacídica envolveu a digestão inicial da proteína com tripsina gerando peptídeos de várias massas diferentes nos quais foram fracionados da mistura através da cromatografia líquida. Após o fracionamento, os peptídeos gerados na digestão foram devidamente identificados pela espectrometria de massas.

As seqüências peptídicas resultantes foram utilizadas nas buscas frente ao banco de dados de proteínas não redundantes do Centro Nacional de Informações Biotecnológicas (NCBI) como expressão de busca. A Figura 30 mostra a representação gráfica da seqüência da rXfHNL e a localização dos peptídeos identificados.

```

1  MNAYRPIAFA SSAASDKPTV LLKPTIVLVH GAFADGSTWN KVIHQLQAK
51  GLNAVSVQNPL TSFEDDVAAT RRVLALQTGP VVLVGHSYGG AVITEAG
101 QDERVKALVYIAA FAPSEGESVA DLNKKYPLPS GYNHLSSDKE GFLMLT
151 PEGVEKYLAQDIPL EQTRLIATQ HPIRGANFEK KVSAAA WETK PSWYL
201 VSENDLMLQPALQKE MAQKIGAHIV NVAASHVPHL SHAAEVEKII MSAV
251 NDVEVLE

```

Figura 30. Cobertura da seqüência de aminoácidos da proteína rXfHNL obtida pelas análises por LC-nanoESI-MS/MS. Os aminoácidos em negrito azul foram identificados através dos peptídeos gerados na digestão com trpsina. 78% dos aminoácidos que compõe a seqüência primária de foram identificados.

Através das seqüências peptídicas reconhecidas, somente uma única proteína apresentou identidade de seqüência com os peptídeos identificados pelo seqüenciamento do digerido tríptico. Este resultados identificaram a proteína rXfHNL como sendo a proteína hipotética conservada de *X. fastidiosa* sob o código de acesso XFa0032 cujo código de acesso no NCBI, massa molecular nominal e pI estão de acordo com a proteína HNL deste estudo. Além disso, o perfil peptídico comparado ao banco de dados proporcionou uma porcentagem de cobertura de 78%, considerada satisfatória.

A alta identidade de seqüência encontrada na seqüência primária permitiu que fosse inferido o sucesso na expressão da proteína heteróloga, da clonagem em vetor pGEM-T *easy* e da subclonagem no vetor de expressão bacteriano pET28a.

III. 1.5.4 – Análise dos Dados de SAXS

Medidas de SAXS (LNLS/Brasil) foram realizadas para averiguar o estado de oligomerização da proteína rXfHNL em solução. As curvas de espalhamento da rXfHNL em solução foram obtidas a partir de duas concentrações distintas: 3 e 12 mg mL⁻¹. Curvas de espalhamento a altas concentrações protéicas podem mostrar um efeito de concentração não desprezível. Assim, foram combinadas a parte de baixa resolução da curva de espalhamento a 3 mg mL⁻¹ (< 0,5 nm⁻¹) com a parte de alta resolução da curva de espalhamento a 12 mg mL⁻¹. A curva de espalhamento resultante desta combinação foi utilizada para a obtenção do modelo *ab initio*. O programa CRY SOL foi utilizado para gerar a curva de espalhamento teórica do modelo DAM ajustada aos dados experimentais obtidos com a proteína rXfHNL. Curvas de espalhamento teóricas também foram obtidas com os modelos cristalográficos de alta resolução de HNLs monomérica e dimérica. As curvas obtidas para as três proteínas foram então comparadas. As duas proteínas HNL obtidas no PDB são de *Hevea brasiliensis* cujos códigos de acesso são 1QJ4 e 1YAS.

A curva de espalhamento experimental obtida na análise da proteína rXfHNL, o ajuste dos dados experimentais (curva teórica gerada) e as curvas teóricas obtidas com o programa CRY SOL para as HNLs monomérica e dimérica são mostradas na Figura 31 (A). Os valores encontrados para o raio do envelope destas estruturas são mostrados na Figura 31 (B). A função de variação de distância $p(r)$ e a dimensão máxima ($D_{\max} = 8,40 \pm 0,50$ nm) para a proteína rXfHNL foi determinada utilizando o programa GNOM.

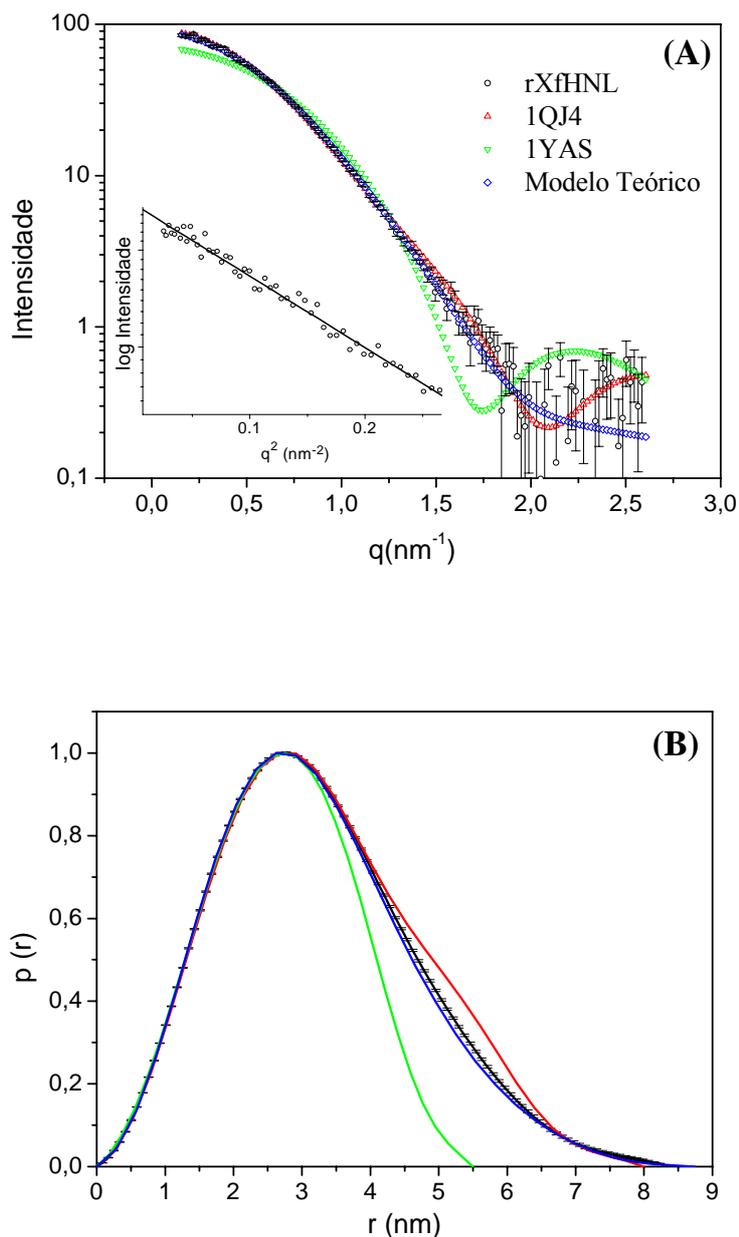


Figura 31. (A) Curvas de espalhamento de raios-X a baixos ângulos. O gráfico interno é a representação do gráfico de Guinier ($\log I$ versus q^2). A curva de espalhamento contendo as barras de erro em preto representam os dados experimentais da rXfHNL. A curva em azul representa o modelo teórico DAM produzido pelo programa GASBOR. As curvas em vermelho e azul representam o espalhamento das proteínas HNLs monomérica e dimérica depositadas no PDB sob os códigos 1QJ4 e 1YAS. **(B) Gráfico da função de variação de distância da rXfHNL.** Os valores $p(r)$ da rXfHNL foram calculados das curvas de espalhamento através do programa GNOM. O parâmetro de dimensão máxima (D_{\max}) encontrado é $8,40 \pm 0,50$ nm. A curva em azul corresponde ao $p(r)$ do modelo DAM. As curvas em vermelho e verde correspondem ao $p(r)$ das HNLs dimérica (PDB code: 1QJ4) e monoméricas (PDB code: 1YAS), respectivamente.

Um conjunto de parâmetros estruturais foi extraído desta curvas experimentais. A linearidade observada no gráfico de Guiner se mostrou característico de proteínas monodispersas em solução e constituídas principalmente por espécies oligoméricas únicas. Além disso, o raio de giro de 2,53 nm foi encontrado para rXfHNL através da região de Guinier da curva de SAXS no qual mostrou-se consistente com uma partícula dimérica.

A curva de espalhamento experimental da rXfHNL proporcionou ainda a construção do seu envelope molecular, utilizando o programa GASBOR, e com ele foi possível averiguar sua consistência com a estrutura cristalográfica resolvida da proteína HNL dimérica 1QJ4. A Figura 32 ilustra a sobreposição da estrutura 1QJ4 com o envelope molecular determinado da rXfHNL (sobreposição obtida com a utilização do programa SUBCOMP). Os dados do envelope molecular foram obtidos empregando-se uma resolução de espalhamento de raios-X de 2,4 nm.

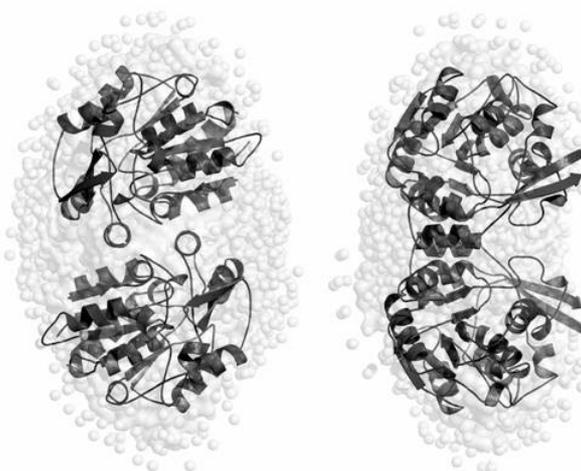


Figura 32. Modelo de SAXS para a rXfHNL. Os círculos em cinza representam o envelope molecular gerado das curvas de espalhamento da proteína rXfHNL. A estrutura cristalográfica sobreposta ao envelope é da HNL dimérica depositada no PDB sob os códigos 1QJ4. Estes modelos representam a visão superior e lateral das moléculas sobrepostas.

A superposição da estrutura cristalográfica da HNL dimérica com o envelope molecular obtido para rXfHNL mostra boa concordância entre as formas, independente da disposição espacial.

Por meio deste estudo, ao contrário do que foi previsto pela cromatografia de exclusão molecular, foi possível concluir que a proteína rXfHNL apresenta-se como um homo-dímero em solução e similar a HNL de *Hevea brasiliensis* (código de acesso 1QJ4) classificada na classe das HNLs independente de cofator.

Apesar dos resultados de SAXS serem bastante evidentes, apenas com dados cristalográficos seria possível obter um melhor esclarecimento das características estruturais da proteína em estudo, porém até o presente momento não há estrutura tridimensional resolvida para α -hidroxinitrila liase de *Xylella fastidiosa*, estudos estes que serão a continuidade deste trabalho.

III. 1.5.5 – Espectroscopia de Dicroísmo Circular

Com a definição da condição oligomérica da rXfHNL partiu-se para a investigação da estrutura secundária por espectroscopia de dicroísmo circular (CD) no UV-distante.

A Figura 33 mostra o espectro de CD da rXfHNL em tris-HCl (pH 8,0) contendo NaCl 150 mM e glicerol 5%, condição cuja estabilidade da proteína foi preservada.

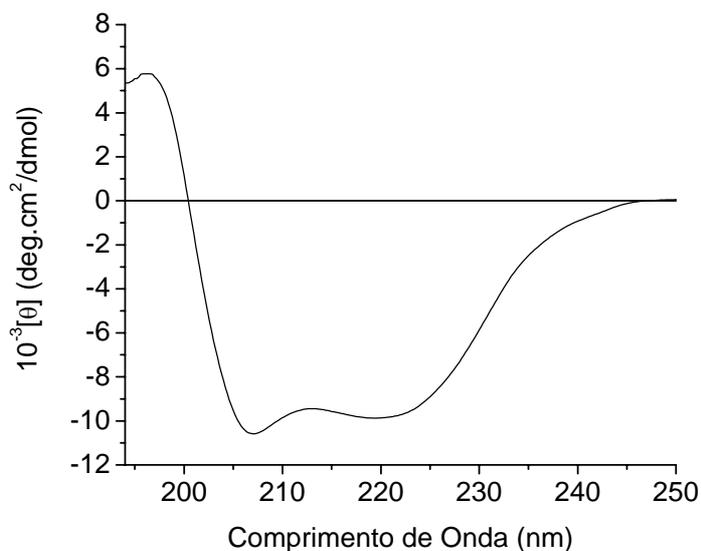


Figura 33. Espectro de dicroísmo circular da rXfHNL. O espectro de CD indica que a rXfHNL é constituída de estruturas α -hélices e folhas β . A rXfHNL foi incubada em tampão tris-HCl 20mM, NaCl 150 mM e glicerol 5%, pH 8,0. As medidas de CD foram realizadas a 25 °C. A concentração da proteína rXfHNL é 10 μ M.

A rXfHNL apresentou um espectro de CD caracterizado por dois mínimos pronunciados em torno de 222 e 207 nm, e um máximo positivo em torno de 196 nm. Estas bandas são características de proteínas contendo estruturas em α -hélice (ref). A desconvolução do espectro em pH 8,0 estimou 35,3 % de elementos helicoidais, 10,9 % de estruturas beta e 53,3 % de outras estruturas. A estabilidade estrutural de rXfHNL foi então avaliada frente a variação de pH, temperatura e na presença de um agente desnaturante.

III. 1.5.5.1 – Estudos de Estabilidade Térmica

Os estudos de estabilidade térmica da rXfHNL foram realizados para averiguar a natureza das transições de desenovelamento e renovelamento térmico,

monitoradas pela espectroscopia de dicroísmo circular. A Figura 34 mostra os espectros de dicroísmo circular da rXfHNL em diferentes pHs a 25 °C.

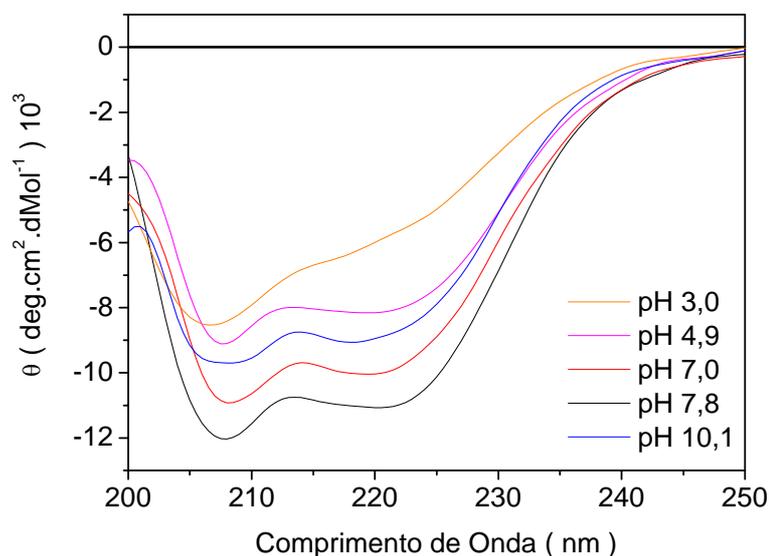


Figura 34. Espectros de CD da rXfHNL em diferentes pH. A concentração da proteína foi de 10 μ M em tampão acetate-borato-fosfato 20 mM, de pH 3,0 a 10,0; contendo NaCl 150 mM.

O espectro de CD da rXfHNL nativa, característico de proteína contendo elementos em α -hélice, foi preservado no intervalo de pH entre 4,9 e 10,1. Apenas pequenas variações foram observadas no valor da elipticidade molar dos mínimos negativos em 222 e 207 nm. Essas diferenças nos conteúdos de estruturas secundárias foram estimadas pela desconvolução dos espectros de CD. A desconvolução do espectro em pH 7,8 estimou 35,5 % de elementos helicoidais, 11,1 % de estruturas beta e 53,4 % de outras estruturas. Diminuindo o pH para 7,0 e 4,0, por exemplo, a desconvolução resultou em 29,1 e 23,6 % de elementos helicoidais, 15,2 e 17,5 % de estruturas beta, e 55,7 e 58,9 % de outras estruturas.

A redução da elipticidade molar provavelmente representa a diminuição de estruturas helicoidais.

No entanto, em pHs extremamente ácidos como o pH 3,0, foram observados um forte decréscimo na elipticidade molar e um pequeno deslocamento do mínimo em 207 para 206 nm, e a perda do mínimo em 222 nm do espectro nativo da rXfHNL. Neste pH, um notável rearranjo e/ou perda de estruturas helicoidais pode ter ocorrido, pois a proteína passou a apresentar um espectro de CD típico de proteínas com estruturas não-ordenadas; provavelmente perdendo parte da sua estrutura secundária ordenada. Neste caso, não foi possível calcular o conteúdo de estrutura secundária.

Conforme descrito na literatura (CANTOR *et al.*, 1980; HAVEL *et al.*, 1996; VENYAMINOV, *et al.*, 1996) estas observações podem ser explicadas através do comportamento protéico no processo básico de enovelamento. Cerca de 80 % das cadeias laterais apolares dos resíduos de aminoácidos hidrofóbicos Ala (13,1 %), Val (10 %), Leu (9,0 %), Ile (5,2 %), Phe (2,4 %), Met (2,0 %), Trp (1,2 %) que constituem a seqüência primária da rXfHNL, estão localizados no interior da molécula, fora do contato com água (meio hidrofílico), e esta é assumida a conformação mais estável, geralmente, o estado nativo (enovelado).

A Figura 35 mostra a isoterma de desnaturação da proteína rXfHNL em tampão acetato-borato-fosfato (pH 8,0) em diversas temperaturas. O processo de desnaturação da rXfHNL foi monitorado pela mudança na elipticidade em 222 nm, que é a banda dominante característica de estruturas helicoidais e está dentro da região do espectro onde a razão sinal/ruído é de alta qualidade.

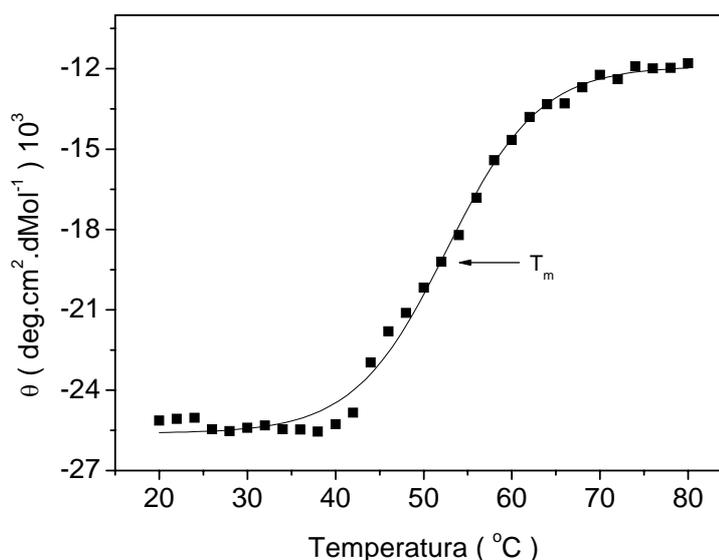


Figura 35. Transição da desnaturação térmica da rXfHNL. A concentração da proteína foi de 20 μM em tampão acetato-borato-fosfato de sódio 20 mM contendo NaCl 150 mM. A curva de desnaturação térmica foi construída a partir dos espectros de CD monitorados a 222 nm. A T_m , temperatura de transição, foi de 52,6 $^{\circ}\text{C}$, obtida através da regressão sigmoideal.

A desnaturação térmica de rXfHNL mostrou-se um processo irreversível, uma vez que as amostras aquecidas até 80 $^{\circ}\text{C}$ não voltaram a apresentar espectro de CD semelhante ao seu estado nativo, mesmo após rápido ou lento resfriamento para 20 $^{\circ}\text{C}$. A curva sigmoideal indica que este processo envolve aparentemente uma transição de dois estados. A temperatura de transição do estado enovelado para o desenovelado da proteína, T_m , foi de 52,6 $^{\circ}\text{C}$.

A curvas de desnaturação da rXfHNL em função da temperatura foi realizada em vários pHs. A Figura 36 mostra a dependência da temperatura de transição da enzima rXfHNL com a variação do pH.

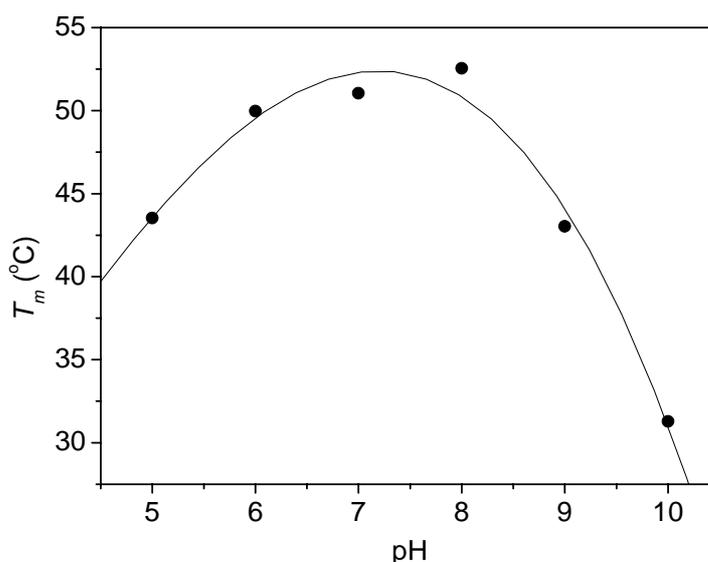


Figura 36. Dependência do T_m de rXfHNL com o pH, monitorado por CD. O tampão utilizado foi acetato-borato-fosfato de sódio 20 mM contendo NaCl 150 mM, variando em pH de 5,0 a 10,0. A concentração da proteína foi de 20 μ M. A temperatura de transição (T_m) de 52,6 °C para o pH de maior estabilidade (pH 8,0) foi obtida através da regressão representada pela curva polinomial.

Os resultados mostram que a proteína rXfHNL é relativamente sensível a alteração de sua estrutura secundária em pHs ácidos ou básicos, em função do aumento da temperatura. A temperatura de transição apresentou pequena variação para os valores de pHs entre 6,0 e 8,0. Contudo, um maior decaimento nos valores de T_m é observado em pH abaixo de 6,0 e acima de 8,0.

III. 1.5.5.2 –Estudos da Estabilidade Química

A estabilidade da estrutura secundária de rXfHNL, na presença de um agente químico, também foi estudada por CD. As medidas de CD foram realizadas a 25°C empregando concentrações de 0 a 6 M do agente caotrópico uréia adicionado em tampão de lise tris-HCl a pH 8,0. Três concentrações distintas da proteína rXfHNL

foram empregadas neste estudo (5, 10 e 20 μM). A Figura 37 mostra os espectros de CD da rXfHNL (20 μM) em diferentes concentrações de uréia.

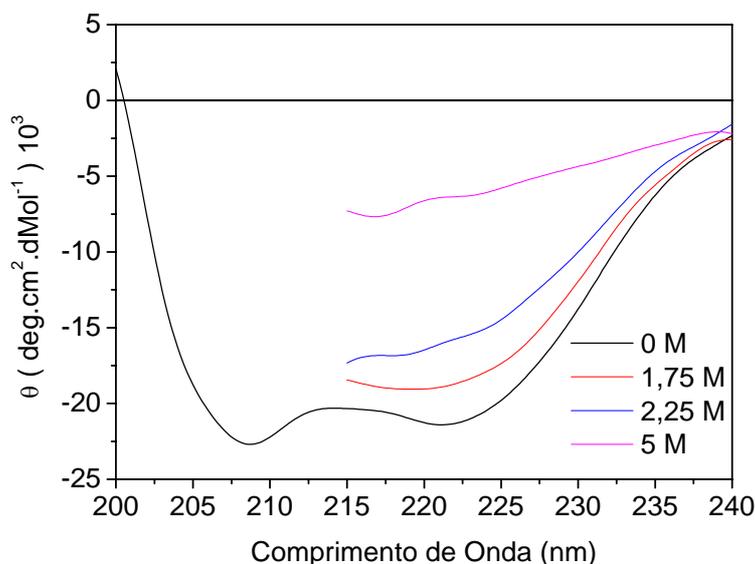


Figura 37. Desnaturação da rXfHNL induzida por uréia. A concentração da proteína foi de 20 μM em tampão tris-HCl contendo NaCl 150 mM, glicerol 5% e concentrações crescentes de uréia. Os espectros de CD foram obtidos individualmente.

O espectro de CD da rXfHNL em uréia a 0 M manteve os dois mínimos negativos a 207 e 222 nm. Com o aumento da concentração de uréia na solução protéica pode ser observado um concomitante decréscimo na elipticidade a 222 nm, sugerindo uma perda cooperativa no conteúdo de α -hélice da estrutura secundária. A completa desnaturação da enzima ocorreu em 5,0 M de uréia, uma vez que o espectro de CD da rXfHNL nesta condição corresponde ao espectro de uma proteína sem estrutura secundária ordenada.

A partir dos resultados de desnaturação química obtidos por CD foram construídas as isotermas de desnaturação, onde a elipticidade, expressa em termos

da fração de proteína desnaturada, f_D , foi colocada no eixo das ordenadas e a variação da concentração do agente desnaturante no eixo das abscissas. A fração de proteína desnaturada foi calculada pela relação:

$$f_D = \frac{\theta_n - \theta_{obs}}{\theta_n - \theta_d} \quad \text{e} \quad f_n + f_d = 1 \quad (2)$$

onde θ_{obs} é a elipticidade ou centro de massa da área do pico (no caso de espectroscopia de fluorescência) da amostra em uma determinada condição, θ_d e θ_n são os valores de elipticidade para os estados desnaturado e nativo.

A análise das isotermas foi feita assumindo que a proteína é um homodímero e que a desnaturação ocorre como um processo reversível de dois estágios, sem a formação de intermediários passíveis de serem detectados. Assim, o ajuste das curvas foi realizado como descrito por Mallam, e colaboradores (MALLAM *et al.*, 2005; HUANG *et al.*, 2005), aplicando o modelo homodimérico de dois estados, em que duas populações de proteína existem em equilíbrio, denominadas homodímeros enovelados (D) e monômeros desenovelados (M). Esse modelo é descrito pelas equações 3 e 4.



$$K_D = \frac{[M]^2}{[D]} \quad (4)$$

A concentração de proteínas monomérica, P_t , para qualquer concentração de uréia pode ser descrito em termos da fração de proteína desnaturada como descrito pela equação 5.

$$P_t = 2D + M = (1 - f_D)P_t + f_D P_t \quad (5)$$

A partir das equações 4 e 5, K_U pode ser expresso em termos de P_t and f_D .

$$K_D = \frac{2P_t(f_D)^2}{(1 - f_D)} \quad (6)$$

Partindo das equações anteriores, a energia livre de desnaturação (ΔG_U) para o modelo de dois estágios pode ser expressa como uma função linear da concentração de uréia, onde $\Delta G_U^{H_2O}$ é:

$$\Delta G_U = \Delta G_U^{H_2O} + m[U] = -RT \ln(K_D) \quad (7)$$

extrapolando a energia livre do desnovelamento na ausência de uréia, m é a inclinação do gráfico ΔG_U versus a concentração de uréia [U], e R e T são a constante dos gases e a temperatura absoluta, respectivamente. Combinando as equações 6 e 7 e rearranjando os valores de $\Delta G_U^{H_2O}$ em termos da fração de proteínas desenoveladas, a concentração de uréia pode ser obtida diretamente das curvas de desnaturação ajustando-as com a equação 8:

$$f_D = \frac{\sqrt{K_D^2 + 8K_D P_t} - K_D}{4P_t} \quad (8)$$

$$K_D = e^{\frac{-(\Delta G_U^{H_2O} + m[\text{denaturant}])}{RT}} \quad (9)$$

A equação 10 pode ser rearranjada e resolvida para determinar $U_{1/2}$.

$$U_{1/2} = \frac{-[RT \ln(P_t) + \Delta G_U^{H_2O}]}{m} \quad (10)$$

onde $U_{1/2}$ é a fração da concentração de uréia em que 50% da proteína está na forma desenovelada ($f_U = 0,5$)

Assim, os dados de CD foram tratados seguindo este modelo e as curvas para cada concentração de proteína foram ajustados individualmente com a equação 8, utilizando para isso o programa Origin 7.0.

Realizando os cálculos termodinâmicos descritos anteriormente, o processo de desnaturação da rXfHNL foi avaliado pela construção das isotermas de desnaturação (*unfolding*). Adicionalmente, o processo inverso de reenovelamento (*refolding*) também foi acompanhado, utilizando como ponto de partida, a proteína rXfHNL desnaturada a 8 M de uréia. O processo de reenovelamento foi realizado através da adição da proteína totalmente desnaturada em soluções contendo concentrações decrescentes do agente caotrópico. A Figura 38 mostram os ajustes com os parâmetros termodinâmicos associados aos processos termodinâmicos de desnaturação e renaturação protéica.

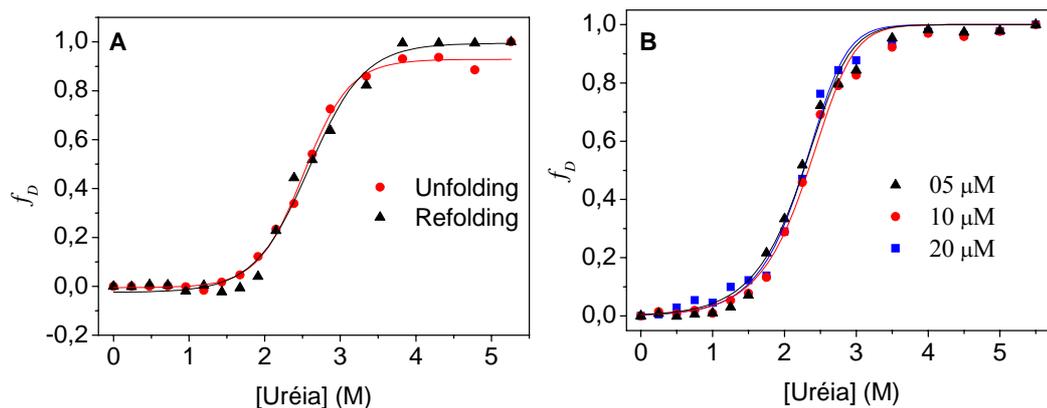


Figura 38. Desnaturação e renaturação da rXfHNL induzida por uréia. Cada curva foi monitorada a 222 nm e expressas em termos da fração de proteínas desnaturada em função da concentração de agente desnaturante. A concentração da proteína empregada em (A) foi de 20 μM em tampão de lise tris-HCl. As isotermas em (B) são representativas do processo de desnaturação química da rXfHNL a diferentes concentrações.

As isotermas construídas indicaram que não há formação de intermediários detectáveis durante a transição do desenovelamento, independente das concentrações protéicas empregadas (5, 10 e 20 μM). A desnaturação ocorreu como um processo reversível, no qual cerca de 90% do sinal da elipticidade inicial foi restaurada.

Através dos espectros de CD da rXfHNL, analisados de acordo com o modelo de dois estágios, foi possível calcular os valores termodinâmicos descritos acima a fim de verificar a independência destes mediante diferentes valores de concentração protéica. A Tabela 3 apresenta alguns destes valores como $m_{N_2 \leftrightarrow 2D}$ e $\Delta G_{H_2O}^{N_2 \leftrightarrow 2D}$. $\Delta G_{H_2O}^{N_2 \leftrightarrow 2D}$ é a diferença da energia livre entre 1 mol de dímero e 2 mols de monômero desenovelado na ausência do desnaturante.

Tabela 3. Dados termodinâmicos para diferentes concentrações da rXfHNL por CD.

P_t (μM)	$[U]_{1/2}$ (M)	$\Delta G_{H_2O}^{N_2 \leftrightarrow 2D}$ (kJ/mol)	$m_{N_2 \leftrightarrow 2D}$ (kJ/mol.K)
5	$2,25 \pm 0,03$	$71,9 \pm 1,9$	$10,9 \pm 0,8$
10	$2,32 \pm 0,06$	$71,4 \pm 2,1$	$11,1 \pm 0,9$
20	$2,27 \pm 0,02$	$70,3 \pm 2,1$	$11,6 \pm 0,9$

A dependência da concentração protéica para o equilíbrio das curvas de desenovelamento em um sistema homodimérico pôde ser explicada em termos da constante de equilíbrio K_D , definida como $K_D = [M]^2/[D]$. Assim, para uma dada concentração de agente desnaturante, K_D e ΔG permaneceram constantes para todas as concentrações protéicas e somente a fração de cada espécie presente no equilíbrio mudou com P_t . Um sistema homodimérico, $[U]_{1/2}$ é definido como a concentração de desnaturante onde a fração de monômeros desenovelados é igual à fração de monômeros presentes como dímeros. Isto é $K_D = P_t$, em que a concentração de $[U]_{1/2}$ ocorrerá dependente de P_t , assim uma mudança na concentração de agente desnaturante com a concentração total de proteína pôde ser esperada, mas a variação de energia livre, $\Delta G_{H_2O}^{N_2 \leftrightarrow 2D}$, permaneceu praticamente constante independentemente da mudança no valor de P_t e $[U]_{1/2}$. Como o valor de m é uma constante para cada proteína este não foi afetado pela concentração protéica.

Deste modo, o modelo de dois estágios homodimérico descreve adequadamente os resultados experimentais. Entretanto, existe a possibilidade que a forma dimérica da rXfHNL possa dissociar-se produzindo monômeros antes de atingir a transição. Contudo, deve-se destacar que estas possíveis formas não puderam ser detectadas pelo método de análise empregado.

III. 1.5.6 – Espectroscopia de Emissão de Fluorescência

O ambiente dos três resíduos de triptofano presentes na seqüência primária da rXfHNL foi monitorado pela espectroscopia de emissão de fluorescência visando complementar e confirmar os experimentos de CD quanto a integridade estrutural e a estabilidade da proteína em diferentes pHs. Mudanças na emissão de fluorescência fornecem informações sobre a acessibilidade do solvente e sobre a hidrofobicidade do meio ao redor destes resíduos. O resíduo de triptofano atua como uma sonda para a estrutura terciária da proteína (LAKOWICZ *et al.*, 1983). Na seqüência primária da rXfHNL, dois resíduos de triptofano estão presentes perto da porção C-terminal da proteína, e o terceiro resíduo está perto da região N-terminal. O espectro de emissão de fluorescência intrínseca de rXfHNL em diferentes pH está mostrado na Figura 39.

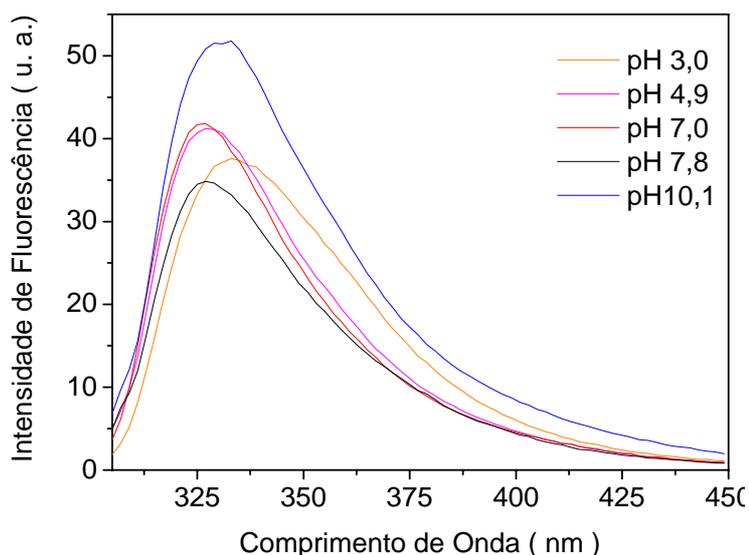


Figura 39. Espectros de emissão de fluorescência da rXfHNL a vários pH. A solução tampão empregada foi acetate-borato-fosfato de sódio 20 mM contendo NaCl 150 mM. As medidas foram realizadas com excitação em 295 nm. A concentração utilizada para as amostras foi 10 μ M.

O máximo de emissão de fluorescência da rXfHNL foi em 327 nm (λ_{em}) no pH 8,0, indicativo de resíduos de triptofano internalizados na proteína. No entanto, o pico de emissão foi assimétrico, sugerindo que o ambiente dos três resíduos de Trp pode não ser idêntico, isto é, um ou mais resíduos provavelmente podem estar expostos a um ambiente mais polar (mais hidrofílico) da estrutura da molécula. Porém, isto poderá ser confirmado apenas pela estrutura cristalográfica da proteína, um ensaio que ainda está em progresso.

Conforme esperado, foi observado que a emissão de fluorescência de rXfHNL foi uma boa estratégia para monitorar mudanças estruturais induzidas pela concentração hidrogeniônica. Nos pH 4,9, 7,0 e 7,8, a rXfHNL mostrou um pico de emissão de fluorescência em 327 nm, para uma excitação em 295 nm. Este resultado indicou que os resíduos de triptofano estão mais protegidos do meio polar e, provavelmente, a enzima se mantém estruturada. Em pHs mais extremos (3,0 e 11,0) ocorreu um deslocamento para o vermelho no pico de emissão (333 e 331, respectivamente), indicando que o ambiente dos resíduos de triptofano tornou-se mais hidrofílico. Como os resíduos de triptofano estão mais expostos à água, a enzima provavelmente começou a se desenovelar (desenovelamento parcial).

Além disso, foi possível observar que a intensidade de fluorescência sofreu mudanças com a variação do pH do meio. Uma supressão de fluorescência intrínseca pôde ser atribuída onde a redução do rendimento de fluorescência foi resultante da distribuição de cargas e da mudança da carga líquida da proteína rXfHNL. Por exemplo, em proteínas monoméricas, a protonação progressiva dos resíduos de aminoácidos pelo decréscimo do pH dá origem a interações repulsivas que conduzem à perda das estruturas secundária e terciária. A estabilização das

pontes salinas é removida e as cadeias laterais internalizadas na estrutura tornam-se expostas. Proteínas que não foram desenoveladas completamente, portanto, podem adotar estados conformacionais relativamente compactos, que diferem da estrutura nativa, conforme indicado por suas propriedades espectroscópicas (LAKOWICZ *et al.*, 1983).

III. 1.5.6.1 – Estudos da Estabilidade Química

A estabilidade química da rXfHNL empregando uréia como desnaturante também foi estudada pela espectroscopia de emissão de fluorescência intrínseca a 25 °C. A Figura 40, mostra o espectro de emissão de fluorescência da rXfHNL em diferentes concentrações de uréia (0; 2,75 e 5 M).

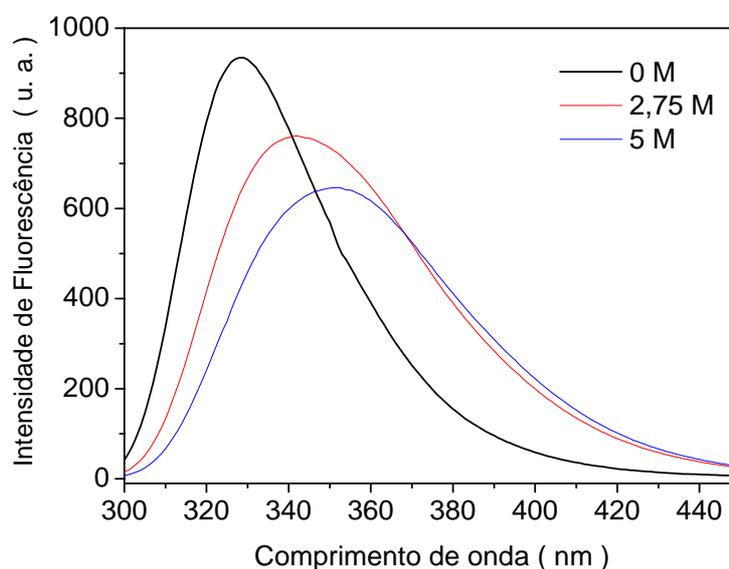


Figura 40. Efeitos da desnaturação induzida por uréia. As amostras da rXfHnl foram expostas a concentrações de uréia variando de 0 a 6 M. As medidas de emissão de fluorescência foram monitoradas de 300 a 450 nm com comprimento de excitação em 295 nm a 25 °C. Os dados foram analisados através do centro de massa de cada espectro.

Com excitação em 295 nm, a rXfHNL em 0 M de uréia também exibiu um máximo de emissão em torno de 327 nm, típico de triptofano internalizado no interior da proteína. Adicionalmente, a incubação da rXfHNL com concentrações crescentes de uréia resultou numa grande supressão da intensidade de fluorescência e num deslocamento significativo do máximo de emissão fluorescência. O comportamento observado à 5 M de uréia pôde, então, ser atribuído a maior exposição dos três triptofanos ao solvente indicado pelo deslocamento para o vermelho (~350 nm). Isto permitiu que a proteína fosse considerada como completamente desenovelada. Com 5 M de uréia, o espectro de fluorescência da rXfHnl é característico de proteínas que apresentam maior porcentagem de estruturas não ordenadas.

Assim, como para as medidas de CD, a presença de dois extremos característicos foi observada, colaborando para sua caracterização como um processo de desnaturação constituído de, pelo menos, dois estados. Desta forma, as isotermas de desnaturação e renaturação puderam ser calculadas.

Para a construção destas isotermas, o centro de massa (λ_{cm}) de cada espectro teve que ser calculado a partir dos espectros de fluorescência. Os centros de massa representam um índice do valor de energia médio do espectro. Os centros de massa dos espectros de fluorescência também denominada comprimento de onda de emissão médio foram calculados de acordo com a equação 10.

$$\lambda_{cm} = \frac{\sum \lambda I(\lambda)}{\sum I(\lambda)} \quad (10)$$

onde λ é o comprimento de emissão, and $I(\lambda)$ representa a intensidade de fluorescência no comprimento de onda λ .

A Figura 41 mostra a fração desnaturada da rXfHNL, calculada através do centro de massa λ_{cm} em função da concentração de uréia.

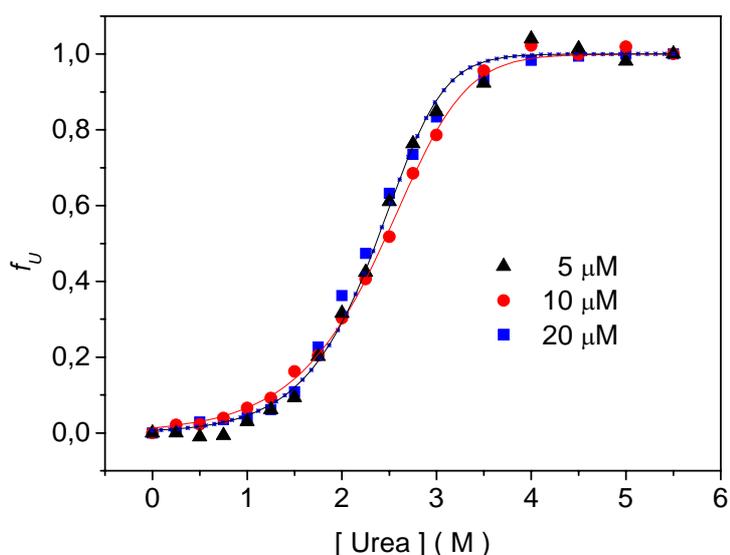


Figura 41. Desnaturação induzida por uréia da rXfHNL por Fluorescência. Cada curva foi calculada através do centro de massa λ_{cm} e expressas em termos da fração de proteínas desnaturada em função da concentração de agente desnaturante. As concentrações de proteína empregadas neste estudo foram de 5, 10 e 20 μM em tampão de lise tris-HCl. As isotermas são representativas do processo de desnaturação química da rXfHNL.

As curvas obtidas por fluorescência também sugerem que o desenovelamento da rXfHNL comporta-se como um processo de dois estados. O reenovelamento da rXfHNL confirmou que a desnaturação química com uréia é um processo reversível. Como observado anteriormente, mudanças no centro de massa do espectro de emissão refletem as mudanças conformacionais induzidas na estrutura da proteína influenciadas pelos resíduos aromáticos, como triptofano e a tirosina, sensíveis à polaridade de seus micro-ambientes. Esta curva sugere que não há

estados intermediários presentes em quantidades detectáveis durante a desnaturação química.

Como a emissão de fluorescência intrínseca fornece uma medida da mudança na estrutura terciária, um teste bem conhecido da transição de dois estados consiste em examinar se as mudanças na estrutura secundária ocorrem paralelamente com as modificações na fluorescência. Quando examinamos a estrutura secundária da rXfHNL pela espectroscopia de fluorescência intrínseca em função da concentração de uréia, as mudanças foram coincidentes com a espectroscopia de CD (Figura 38). Analisando a estrutura de α -hélice em 222 nm, observa-se uma transição sigmoide semelhante. Neste caso, o ajuste dos dados de acordo com o modelo de *Mallam et al.*, também sugeriu ausência aparente de intermediários. Adicionalmente, o reenovelamento da rxfHnl revelou que a desnaturação química medida por fluorescência também foi um processo reversível (resultados não mostrados).

Alguns resultados do ajuste dos dados da desnaturação com uréia pela equação 8 e 10 estão resumidos na Tabela 4.

Tabela 4. Dados termodinâmicos para diferentes concentrações da rXfHNL por Fluorescência.

P_t (μM)	$[U]_{1/2}$ (M)	$\Delta G_{H_2O}^{N_2 \leftrightarrow 2D}$ (kJ/mol)	$m_{N_2 \leftrightarrow 2D}$ (kJ/mol.K)
5	$2,20 \pm 0,07$	$70,6 \pm 1,2$	$10,9 \pm 0,5$
10	$2,21 \pm 0,06$	$67,3 \pm 0,5$	$9,8 \pm 0,3$
20	$2,27 \pm 0,05$	$68,7 \pm 1,2$	$10,9 \pm 0,5$

Os valores de ΔG_{H_2O} foram determinados calculando os valores médios de ΔG_{H_2O} calculados em cada concentração de proteína. Os valores de ΔG_{H_2O} entre as concentrações da proteína ficaram dentro do erro experimental (~5%). Isto indica que os valores foram efetivamente invariantes sobre uma escala de concentrações da proteína, como se esperaria para o modelo de desnaturação de dois estados.

Portanto, o modelo de dois estados para a desnaturação de uma proteína homodimérica apenas o dímero nativo e o monômero desnaturado estão significativamente populados no equilíbrio, é consistente com os dados de desnaturação por uréia, descritos neste estudo. Duas condições deste modelo são que o desenovelamento do dímero é reversível e que as curvas do desenovelamento demonstraram dependência com a concentração da proteína. Foi estabelecido, portanto, a total reversibilidade da desnaturação da rxfHnl. Pela lei de ação das massas, o aumento da concentração da proteína aumentará a proporção da molécula nativa homodimérica em cada concentração do desnaturante, e o ponto médio da curva sigmoide de desnaturação aumentará com o aumento da concentração da proteína.

III. 1.5.7 – Ensaio de Atividade Enzimática

Os estudos estruturais descritos anteriormente permitiram avaliar a atividade específica da rXfHNL nos pH 4,9 e 7,0. A atividade da hidroxinitrila liase foi determinada estimando a produção do benzaldeído durante a decomposição do substrato em 280 nm. Todas as atividades enzimáticas foram medidas a 25 °C. A

reação enzimática no pH 7,8, condição em que a rXfHNL é mais estável, não foi avaliada devido a significativa degradação da mandelonitrila. O substrato se decompõe espontaneamente neste pH. Reações de controle, sem rXfHNL, para as condições do ensaio permitiram uma estimativa da quebra não-enzimática do substrato e os valores da atividade foram ajustados de maneira adequada.

As atividades enzimáticas específicas observadas em pH 4,9 e 7,0 foram de 6,8 e 33,8 U mg⁻¹, respectivamente. A proteína recombinante produziu uma quantidade mais elevada de benzaldeído por minuto em pH 7,0 do que em pH 4,9. A atividade da liase aqui representada foi definida como a quantidade de enzima (em uma concentração 113 µg mL⁻¹) que catalisa a produção de 76 and 380 nmol de benzaldeído por minuto em pH 4,9 e 7,0, respectivamente. A atividade da enzima mostrou ser independente do cofactor FAD, já que as atividades específicas encontradas na presença de FAD foram praticamente iguais (6,6 e 33,2 U mg⁻¹).

Estes resultados demonstram claramente que a proteína recombinante foi obtida na forma ativa. Uma vez que a reação de cianogênese é descrita como um processo reversível, estes resultados podem ser correlacionado com a hipótese que a rXfHNL está envolvida em uma via bioquímica para a detoxicação ao cianeto secretado pela planta como um mecanismo de defesa contra a *X. fastidiosa*.

O desenovelamento parcial da rXfHNL nos pHs mais ácidos, anteriormente proposto pelos resultados de CD e de fluorescência, foi acompanhado por uma perda da atividade enzimática.

A caracterização cinética e a determinação da estrutura tridimensional da rXfHNL foram iniciadas recentemente a fim alcançar a completa caracterização da

proteína e conseqüentemente proporcionar maiores contribuições no conhecimento de seu papel biológico.

III. 1.5.8 – Conclusões Parciais

As etapas de clonagem e a expressão de heteróloga em *E. coli* BL21(DE3) resultaram em quantidades satisfatórias de rXfHNL na forma solúvel e facilmente purificáveis.

As técnicas de SDS-PAGE, Gelchip-CE e cromatografia de exclusão molecular mostraram que a proteína recombinante apresentou massas moleculares próximas a esperada.

A proteína recombinante de *X. fastidiosa* demonstrou ser uma molécula estável, homodimérica e de forma alongada em solução, como demonstrada pelos resultados de espalhamento de raios-X a baixo ângulo.

Baseado nos experimentos de CD e fluorescência, os resultados sugerem que rXfHNL é possuem uma estrutura secundária que é mais estável entre pHs 4,9 a 7,8. A ausência do mínimo observado em 220 nm no espectro de CD em pH 3,0 e a mudança para comprimentos de onda mais elevados (~350 nm) dos máximos de emissão dos espectros de fluorescência em pH mais extremos (3,0 e 10,1), indicam que a estabilidade estrutural da enzima foi afetada. Porém, a máxima estabilidade de rxfHnl foi observada entre os pH 6,0 e 7,8.

O processo de desnaturação química induzido pela uréia se dá através de um processo reversível clássico, de dois estados, onde somente os estados nativo e desnaturado estão significativamente presentes em quantidades detectáveis, sem intermediários. Este processo é descrito adequadamente por um modelo

homodimérico. Entretanto, o desenovelamento térmico foi um processo irreversível nas circunstâncias aqui descritas. Contudo, a forma dimérica, provavelmente antes de alcançar o estado desnaturado, é esperada dissociar para produzir monômeros e é interessante que estas possíveis formas não foram detectadas para a rXfHNL nos dois métodos usados no presente estudo.

Durante a anotação do gene Xfa0032 do projeto genoma, foi levantada a hipótese de que a rXfHNL estivesse envolvida no processo de detoxicação da bactéria. Este estudo reforça esta hipótese, uma vez que a proteína rXfHNL é ativa e apresenta dependência com o pH, podendo perfeitamente estar envolvida na reação inversa do processo de cianogênese. De alguma maneira, a liberação do cianeto pela planta pode servir como substrato para a produção de glicosídeos cianogênicos, atóxicos, comumente encontrados no metabolismo de bactérias. Isto explicaria a sobrevivência desta bactéria no hospedeiro.

Uma vez que HNLs de outros organismos são, atualmente, enzimas muito empregadas industrialmente na síntese orgânica estereoespecífica de compostos ativos, estes resultados podem ser muito úteis como uma fonte de referência na manipulação futura da enzima rXfHNL em aplicações biotecnológicas. Assim como, poderá servir como guia no desenho de moléculas inibidoras alvo contra a colonização da *Xylella fastidiosa* em citros.

III. 2 – Corismato Sintase

III. 2.1 – Amplificação e Clonagem no Vetor pGEM-T *easy*

As suspensões bacterianas utilizadas na extração do DNA genômico também foram do isolado 9a5c da bactéria *X. fastidiosa*. A metodologia empregada para avaliar a qualidade e a concentração do DNA genômico obtido foi descrita na seção III. 1.1.

Como descrito anteriormente, os oligonucleotídeos utilizados na amplificação foram desenhados com base na sequência da ORF *aroC* e foram incorporados com sítios de reconhecimento das enzimas de restrição *NdeI* e *XhoI* preservando seus códons originais de iniciação e terminação. Estes sítios de restrição nos iniciadores a fim de permitir sua subclonagem nos vetores de expressão.

Como para o estudo anterior, a temperatura de hibridização do par de iniciadores (F e R) desenhado foi calculada teoricamente através do software Gene Runner 3.05, buscando garantir a eficiência da reação de amplificação do DNA de interesse. A temperatura ótima encontrada neste caso ficou em torno dos 69 °C.

Com base na temperatura calculada para o par de oligonucleotídeos desenhado foi possível a amplificação do fragmento correspondente ORF que codifica a proteína corismato sintase, porém com baixa concentração e na presença de algumas bandas inespecíficas. Diante deste resultado, se fez necessária a otimização da PCR a fim de garantir o máximo de eficiência na obtenção dos produtos desejados. Assim, somente quando foi empregada a temperatura de hibridização dos *primers* de 58 °C, foi possível a amplificação desejada.

A Figura 42 ilustra o resultado desta amplificação a qual gerou um produto de aproximadamente 1125 pb. Resultados excelentes foram obtidos na amplificação do gene de interesse. O produto de amplificação obtido pela PCR teve o seu tamanho concordante com o previsto, não apresentou nenhuma banda inespecífica, indicando o sucesso da otimização e resultou em quantidades suficientes para proceder à clonagem posterior.

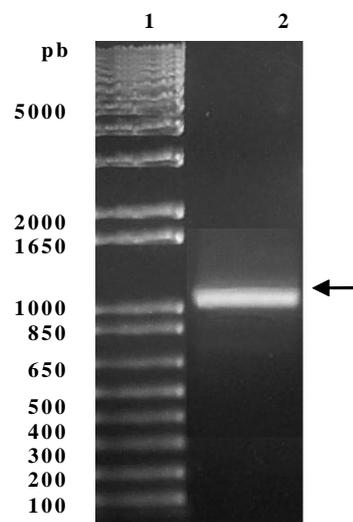


Figura 42. Amplificação do fragmento de DNA que codifica a proteína corismato sintase. A coluna 1 é o padrão de massa molecular 1 kb Plus DNA Ladder da Gibco BRL. A coluna 2 é a ORF *aroC* amplificada. O meio de separação foi gel de agarose 0,8 % em tampão 1x TAE, pH 8,3. A voltagem aplicada na separação foi 100 V por 30 min. Os fragmentos de DNA no gel foram corados com brometo de etídeo e visualizados em transiluminador UV a 365 nm.

Os produtos da amplificação por PCR foram devidamente purificados pelo *kit Wizard PCR Clean-up* da Promega, quantificados em espectrofotômetro e adenilados a fim de inseri-los no vetor de clonagem pGEM-T easy. A concentração obtida para estes produtos variou de 300 até 500 ng μL^{-1} . A construção nos vetores de clonagem foi denominada de pGEM-X*faroc*.

Os plasmídeos recombinantes pGEM-XfaroC recém construídos foram inseridos por choque térmico em bactérias de *E. coli* DH5 α competentes. As bactérias transformadas foram semeadas em placas de Petri contendo meio LB sólido, ampicilina, X-gal e IPTG e crescidas *overnight* a 37 °C.

Como para a HNL, diversas colônias brancas e azuis cresceram na placa contendo o meio selecionado descrito acima. A seleção dos transformantes foi realizada então pela capacidade de crescer em presença do antibiótico ampicilina e através da coloração branca de alguns transformantes indicativa da presença do inserto dentro da região codificadora da enzima β -galactosidase. Dentre as colônias brancas resistentes, novamente quatro colônias foram escolhidas aleatoriamente, inoculadas em tubos estéreis contendo meio LB líquido com ampicilina e submetidas à incubação *overnight* a 37 °C sob agitação. Após o crescimento individual destas colônias, seus DNAs plasmidiais foram extraídos, quantificados, clivados por enzimas de restrição e finalmente seqüenciados.

III. 2.1.1 – Análise de Restrição e Seqüenciamento dos Plasmídeos Recombinantes

O DNA plasmidial dos 4 prováveis clones positivos extraídos através dos métodos descritos pelo *kit Wizard Miniprep*[®] e de lise alcalina empregados de acordo com o descrito em II.2.7, foram inicialmente considerados de boa qualidade e adequados para as análises de restrição e seqüenciamento. As amostras de DNA plasmidial extraídos pelo kit e pelo método de lise alcalina apresentaram neste caso concentrações de 150 e 80 ng μL^{-1} , respectivamente.

A análise de restrição empregando as enzimas *NdeI* e *XhoI* foi realizada a fim de confirmar a presença do inserto nos clones selecionados. A Figura 43 apresenta o perfil de restrição destes nove clones selecionados.

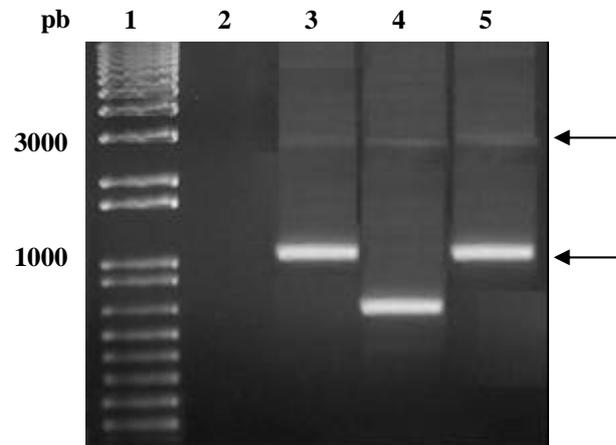


Figura 43. Análise de restrição dos plasmídeos pGEM-aroC. A coluna 1 é o padrão de massa molecular 1 Kb Plus DNA Ladder da Gibco BRL. As colunas 2 a 5 são os clones selecionados. O meio de separação foi gel de agarose 1 % em tampão 1x TAE, pH 8,3. A voltagem aplicada na separação foi 100 V por 30 min. Os fragmentos de DNA no gel foram corados com brometo de etídeo e visualizados em transiluminador UV a 365 nm. As bandas específicas dos genes *XfaroC* e dos plasmídeos digeridos estão representados por setas cujos tamanhos são aproximadamente 3000 e 1125 pb, respectivamente.

A presença e o tamanho do inserto correto foram confirmados apenas em dois dos 4 clones selecionados. Assim, os clones 2 e 4 localizados na 3^a e 5^a canaleta do gel puderam confirmar o sucesso das etapas de clonagem e transformação das bactérias. Os produtos da digestão de aproximadamente 1000 pb correspondente a ORF *aroC* foram retirados do gel, purificados utilizando-se novamente o *kit Wizard PCR Clean-up*, quantificados e utilizados na etapa de subclonagem nos vetores de expressão. Estas amostras de DNA apresentaram em geral concentrações variando de 150 a 200 ng μL^{-1} .

Os DNAs plasmídias selecionados para o seqüenciamento também foram os clones 2 e 4. Os resultados do gel de seqüenciamento para estes clones foram de baixa qualidade apresentando pouca eficiência na leitura das bases. Assim, não permitiram a certificação de que o gene correspondia exatamente a ORF de interesse. Em vista disso, o seqüenciamento dos plasmídeos contendo o gene ligado no vetor de expressão pET28a foi posteriormente realizado a fim de garantir sua identidade.

III. 2.2 – Subclonagem no Vetor de Expressão, Análise de Restrição e Seqüenciamento dos Plasmídeos Recombinantes

Os fragmentos de DNA de 1125 pb que codificam para a proteína rXfaroC foram subclonados nos vetores pET28a. Esta nova construção denominada pET-XfaroC também foi escolhida devido a expressar a proteína de interesse fusionada com uma cauda de seis histidinas na região N-terminal. Novamente, os plasmídeos recombinantes recém construídos foram inseridos por choque térmico em bactérias de *E. coli* DH5 α competentes, semeadas em placas semelhantes ao processo anterior, porém somente com adição de canamicina como agente seletivo, e finalmente incubadas para o crescimento. A seleção das colônias neste caso foi realizada pela análise de restrição.

Após o período de crescimento celular, quatro colônias foram selecionadas aleatoriamente dentre as inúmeras colônias presentes na placa de Petri, e inoculadas seguindo o mesmo processo descrito anteriormente. Após o crescimento individual destas células seus DNAs foram extraídos, quantificados e alíquotas destes DNAs foram clivados pelas enzimas de restrição e submetidas

novamente ao sequenciamento. Estas amostras de DNA também apresentaram concentrações variando de 120 a 150 ng μL^{-1} . A Figura 44 ilustra o novo perfil de restrição destes quatro clones selecionados.

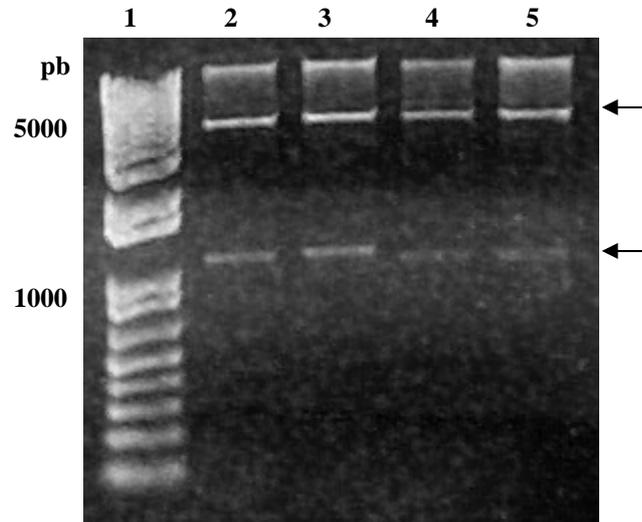


Figura 44. Análise de restrição dos plasmídeos pET-XfaroC. A coluna 1 é o padrão de massa molecular 1 Kb Plus DNA Ladder da Gibco BRL. As colunas 2 a 5 são os clones selecionados. O meio de separação foi gel de agarose 1 % em tampão 1x TAE, pH 8,3. A voltagem aplicada na separação foi 100 V por 30 min. Os fragmentos de DNA no gel foram corados com Brometo de etídeo e visualizados em transiluminador UV a 365 nm. As bandas específicas da XfaroC e dos plasmídeos digeridos estão representados por setas.

Uma vez confirmada a presença do inserto nos quatro clones selecionados, as etapas de subclonagem e transformação das bactérias foram consideradas bem sucedidas. Apenas os clones 1 e 2 (equivalentes as 2^a e 3^a canaleta do gel) que apresentaram maior intensidade de banda, foram utilizados para transformar as células competentes de expressão de *E. coli* BL21(DE3). A seleção de dois novos clones contendo as ORFs subclonadas nos vetores de expressão foi realizada após o plaqueamento das células em meio adequado. Finalmente, um inóculo destes

clones foi empregado na produção da proteína recombinante em células de *E. coli* BL21(DE3).

As seqüências obtidas a partir do seqüenciamento foram analisadas através do programa *Sequencing Analysis V 3.4*. A Figura 45 ilustra a seqüência resultante de um dos clones submetidos ao seqüenciamento.

```

1  TCAATGCNTT GCGTCGCATG CTCCGGGCCG CNTGGCGGCC GCGGGAATTC GATTGGAATT CCATATGGGG
2  GCGAACACGT TTGAAAAGTT ACTGGCGGTC ACGACCTTCG GTGAATCACA TGGCCCTGCC ATTGGCTGTG
3  TGATTGATGG CTGTCCTCCC GGATTGGAAC TCGCTGCTGA GGAGTTTGCC CACGACTTGC AGCGTCGCGC
4  TACCGGTAGG AGCCGTCATA CCTCGGCGCG GCGTGAGGCC GATGAAGTCG AGATTCTCTC CGGCGTCTAC
5  GAAGGCCGTA CCACGGGCAC TCCTATTGCA CTGTTGATCC GCAATACCGA CCAGCGTAGT AAGGATTATG
6  CGACGATTGC CCGGCAGTTC CGTCCGGGCC ATGCCGATTA CACCTATTGG CAGAAATATG GTATCCGTGA
7  TCCTCGTGGT GGTGGGCGTT CTCGGCCCCG TGAGACAACA ATGCGTGTGG CTGCTGGGGT GGTGCTAAA
8  AAGTGGCTGA AGCAGCGTTA TGGTGTGATT GTGCGCGGTT TTTTATCCA GCTTGGTGAG ATCCGCCCGG
9  AAGGCTTTCG TTGGGATGCG GTCGAAGATA ATCCTTTTTT CTGGCCGCAG GCGGCTCAGG TGCCAGAGCT
10 GGAAGCCTAC ATGGATGCAT TGCGCAAATC TGAAATTCN GTCGGTGAC GCTGGATGTG GTTGCCGAAG
11 GCGTGCCGCC AGGGTGGGGT GAGCCGATTT ACGGCAAGCT TGATGGTGAA CTTGCTGCTG CGTTGATGAG
12 TATCAACGCA GTAAAAGGTG TGGAGATCGG CGCTGGTTTC GGCAGCACTG TACAAAAAGG AACCGAGCAC
13 CGTGATTGA TGACACCGTT AGGTTTTCTC AGTAACCATG CTGGCGGCAT CATCGGTGGG ATCACCCTG
14 GTCAGCCGAT TATTGTTTCG ATTGCTCTGA AACCAACATC TAGCCTACGT TTGCCTGGCG AGACGGTGGA
15 TGTTGATGGC CACCCTGTCC AAGTGATCAC CAAGGGACGT CATGATCCTT GTGTGGGGAT TCGTGCCCT
16 CCGATTGCTG AAGCCATGGT CGCGTTAGTA CTGATGGACC AGGCGGTGCG CCACCGTGCC CAGTGTGGCG
17 ATGTCGGTGA GATGTCTCCG TGCATCCTTG AGAATGTGGG TTTCAGGAAT GCCGATGATT GACTCGAGCC
18 GAATCACTAG TGAATTCGCG GCCGCCTGCA GGTGACCAT ATGGAGAGCT CCCAANGCGT TNGAT

```

Figura 45. Seqüência obtida do plasmídeo pET-XfaroC extraído do clone 3. As bases em vermelho correspondem a ORF aroC que codifica para proteína corismato sintase. As bases em negrito e sublinhado são correspondentes às seqüências dos oligonucleotídeos iniciadores. Os nucleotídeos em negrito preto foram os não identificados pelo seqüenciador.

As seqüências resultantes foram submetidas a um processo de busca por similaridade em diversos bancos de dados internacionais, através de servidores de busca disponíveis na rede. O principal programa utilizado foi o software público BLAST onde os resultados mostraram similaridade de seqüência variando de 99%

de identidade frente à cadeia nucleotídica correspondente a ORF deste estudo, confirmando a presença da região alvo desejada.

De modo geral os clones seqüenciados apresentaram pequenas variações entre si, que podem ser atribuídos a erros de leitura ou mesmo a erros gerados durante a marcação para o seqüenciamento. No entanto, os plasmídeos pET28a-XfaroC seqüenciados apresentaram os genes de interesse inseridos na orientação e com fase de leitura corretas dentro dos vetores pET28a garantindo assim a integridade destes clones e a leitura eficiente de todas as bases dos fragmentos na etapa de expressão protéica.

III. 2.3 – Expressão Heteróloga da Proteína rXfaroC a partir dos Vetores de Expressão

Como no estudo de expressão heteróloga anterior, a produção da proteína rXfaroC foi realizada inicialmente por meio do teste de expressão na presença de IPTG a 37 °C. Porém, após o período de indução de 4 horas e a subsequente lise celular na presença de tampão tris-HCl, nenhuma expressão diferenciada na região de aproximadamente 45 kDa foi observada a partir dos clones selecionados.

Diante da expressão negativa da proteína de interesse, uma etapa de otimização do processo de indução foi necessária neste estudo. A mudança da temperatura de expressão para 20 °C foi inicialmente avaliada. Conseqüentemente, um período de indução de 12 horas se fez necessário para alcançar quantidade suficiente de proteína para sua visualização por SDS-PAGE.

A Figura 46 mostra um gel de SDS-PAGE a 15% obtido após submeter às amostras resultantes da expressão a condições desnaturantes.

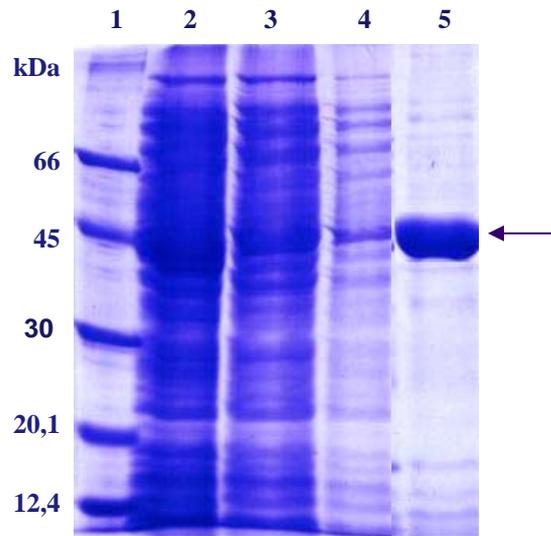


Figura 46. Expressão piloto da proteína rXfaroC a partir do clone 1. A análise foi realizada em SDS-PAGE a 15%. A coluna 1 é o marcador de massa molecular. A coluna 2 é o extrato protéico total não induzido do sistema de expressão. A coluna 3 é o extrato protéico total induzido. As colunas 4 e 5 são as frações insolúvel e solúvel pós lise celular. A seta indica a expressão diferenciada esperada na região de 45 kDa.

Apesar da expressão diferenciada na região esperada não ter sido claramente visualizada no extrato protéico induzido, sua presença foi confirmada pela banda diferenciada na fração solúvel (coluna 5) a qual concorda com o tamanho esperado para a proteína rXfaroC de aproximadamente 45 kDa, revelando a eficiência da metodologia de clonagem e expressão da proteína recombinante.

Além disso, após a lise celular foi constatado que a proteína se apresentou em grande quantidade nas frações solúveis com o emprego do tampão tris-HCl, apesar de ter sido também observada a permanência de quantidade redundante da proteína na fração insolúvel do sistema de expressão.

As frações solúveis foram, então, empregadas na etapa de purificação da proteína rXfaroC.

Neste caso, a produção em maior escala da proteína em *E.coli* foi realizada depois de avaliado o processo de purificação uma vez que a quantidade analítica foi considerada satisfatória para os subseqüentes estudos estruturais e funcionais.

III. 2.4 – Purificação da Proteína rXfaroC por Cromatografia de Afinidade.

As frações solúveis provenientes da lise celular das culturas induzidas foram submetidas à cromatografia de afinidade utilizando uma resina com níquel imobilizado. As frações eluídas durante a etapa de purificação foram analisadas por SDS-PAGE. A Figura 47 mostra a separação eletroforética em gel de poliacrilamida SDS-PAGE a 15% dos eluatos da coluna de Ni-NTA.

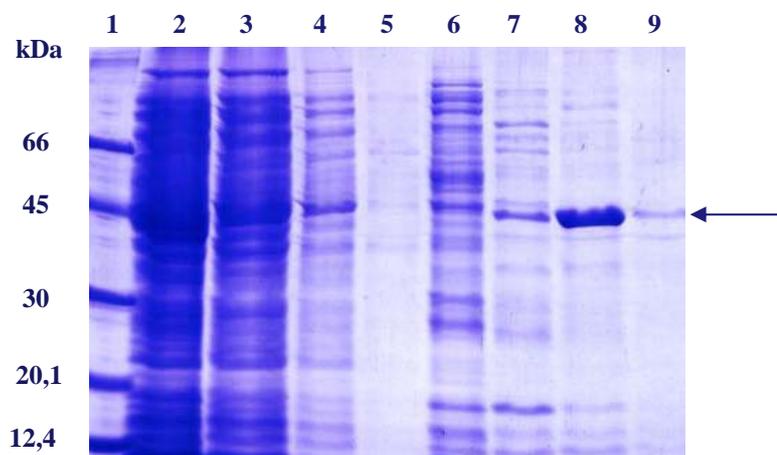


Figura 47. Purificação da rXfaroC por cromatografia de afinidade. A análise das frações eluídas da coluna Ni-NTA foi realizada em SDS-PAGE a 15%. Colunas: 1) marcador de massa molecular; 2) fração insolúvel após a lise celular; 3) fração solúvel após a lise celular; 4) fração não ligada na coluna de afinidade; 5) lavagem com 5 mM de imidazol; 6) lavagem com 40 mM de imidazol; 7, 8 e 9) eluição com 100, 250 e 500 mM de imidazol, respectivamente. A banda correspondente a proteína rXfaroC está indicada pela seta.

A purificação da proteína recombinante rXfaroC não foi alcançada em um único passo cromatográfico como para a HNL. Proteínas contaminantes em maior quantidade foram observadas nas frações de eluição.

O processo de otimização da purificação envolveu o aumento dos volumes de tampão contendo 5 e 40 mM de imidazol empregados no processo de lavagem. Assim, o dobro de seus volumes iniciais foi empregado na purificação por afinidade. A Figura 48 mostra a separação eletroforética em gel de poliacrilamida SDS-PAGE a 15% dos novos eluídos da coluna de Ni-NTA.

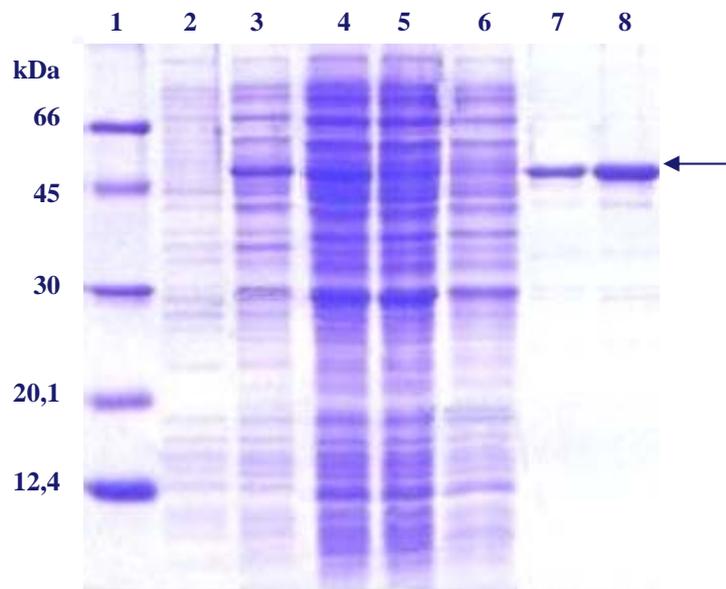


Figura 48. Otimização da purificação da rXfaroC por cromatografia de afinidade. A análise das frações eluídas da coluna Ni-NTA foi realizada em SDS-PAGE a 15%. Colunas: 1) marcador de massa molecular; 2) fração não ligada na coluna de afinidade; 3 e 4) representam as frações de lavagem com 5 mM de imidazol; 5 e 6) frações de lavagem com 40 mM de imidazol; 7 e 8) frações de 100 e 250 mM de imidazol, respectivamente. A proteína rXfaroC purificada está indicada pela seta.

Como mostra Figura 48, a purificação da proteína rXfaroC foi alcançada após a otimização do processo. Em algumas ocasiões, quando as frações eluídas a

250 mM de imidazol se apresentavam ainda com alguns contaminantes, a simples re-cromatografia por afinidade foi suficiente para se obter a proteína na forma mais adequada. É importante ressaltar que para as eventuais etapas de re-cromatografia foi necessário retirar o excesso de imidazol presente nas frações recém eluídas da coluna através da diálise.

A pureza foi confirmada pela análise das frações eluídas com tampão de lise tris-HCl contendo 100 e 250 mM de imidazol, as quais apresentaram uma única banda de aproximadamente 45 kDa, sob condições desnaturantes.

Numa tentativa de prever o grau de oligomerização da proteína rXfaroC através de cromatografia de exclusão molecular foi realizada, porém sua instabilidade não permitiu que tais experimentos fossem conclusivos.

III. 2. 5 – Análises Estrutural da Proteína rXfaroC

III. 2.5.1 – Espectroscopia de Dicroísmo Circular

A fim de investigar a estrutura secundária e estimar seu pH ótimo de estabilidade, a rXfaroC purificada foi estudada por espectroscopia de dicroísmo circular (CD).

A Figura 49 ilustra o espectro de CD observado para a rXfaroC em tampão de lise tris-HCl a pH 8,0 contendo NaCl e glicerol, condição cuja estabilidade da proteína foi preservada.

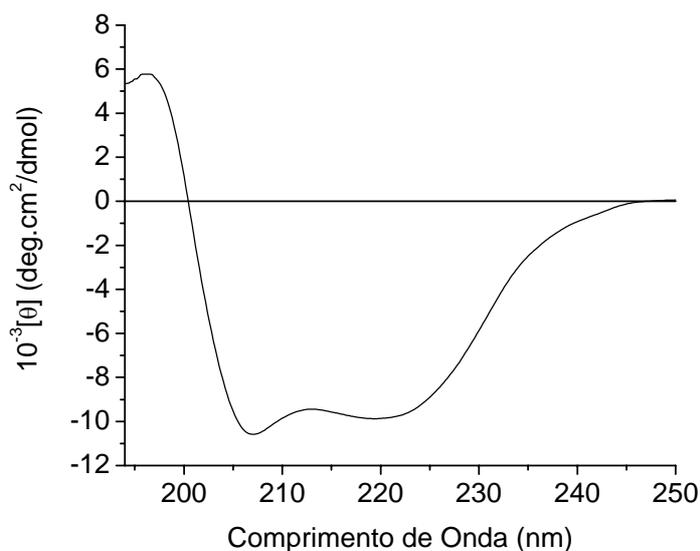


Figura 49. Espectro de dicroísmo circular da proteína recombinante rXfaroC. O espectro de CD mostra que provavelmente a rXfaroC é também constituída predominantemente de estruturas α -hélices e folhas β . A rXfaroC foi preparada em tampão tris-HCl a 20mM, NaCl a 150 mM e glicerol a 5%, pH 8,0. As medidas de CD foram realizadas a temperatura de 25 °C.

A rXfaroC também apresentou um espectro de CD caracterizado por dois mínimos pronunciados em torno de 220 e 207 nm, e um máximo positivo em torno de 196 nm. Estas bandas são características de proteínas contendo estruturas em α -hélice. A estabilidade estrutural de rXfaroC foi então avaliada frente a variação da concentração hidrogeniônica (Figura 50).

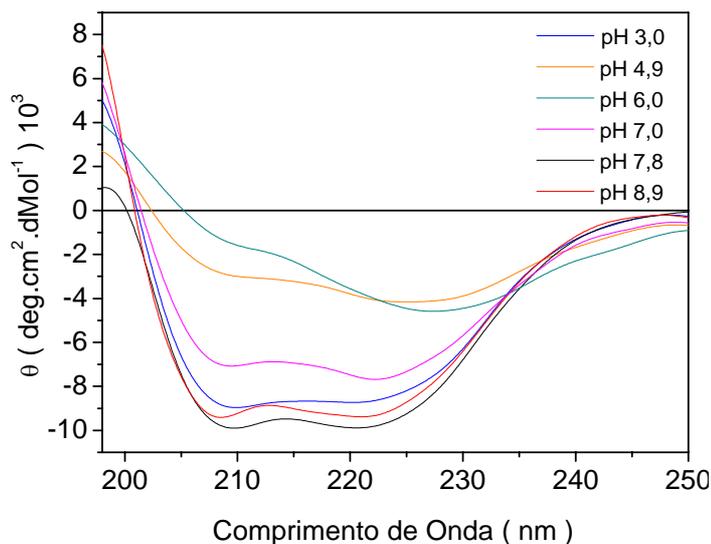


Figura 50. Espectros de CD da rXfHNL em diferentes pHs. A concentração da proteína foi de 10 μM em tampão acetate-borato-fosfato a 20 mM contendo NaCl a 150 mM.

Os espectros de CD, característicos de estruturas contendo elementos do tipo em α -hélice, foram relativamente preservados nos pHs 3,0; 7,8 e 8,9. As presenças dos mínimos negativos em torno de 220 e 207 nm foram observadas com pequenas variações no valor da elipticidade molar. Além disso, análises de desconvolução dos espectros de CD mostraram diferenças nos conteúdos de estruturas secundárias. A desconvolução do espectro em pH 8,9 estimou 33,4 % de elementos helicoidais, 14,1 % de estruturas beta e 52,5 % de outras estruturas. Diminuindo o pH para 7,8 e 3,0 por exemplo, a desconvolução do espectro resultou em 28,1 e 21,5 % de elementos helicoidais, 16,8 e 17,9 % de estruturas beta, e 55,1 e 58,6 % de outras estruturas. A redução da elipticidade molar provavelmente representa a diminuição de estruturas helicoidais.

No entanto, em pHs intermediários como o pH 4,9 e 6,0, um forte decréscimo na elipticidade molar e notáveis deslocamentos dos mínimo em 207 e

220 nm foram observados. Nestes pHs, um notável rearranjo e/ou perda de estruturas helicoidais pode ter ocorrido. A proteína provavelmente se tornou desnaturada ou desenovelada, caracterizada pelo espectro típico de proteínas com estruturas não-ordenadas. É importante mencionar que o pI teórico calculado para a rXfaroC é de 6,33. Esta desestabilidade estrutural em pHs intermediários pode ser devido ao seu pI. Neste caso, não foi possível calcular o conteúdo de estrutura secundária.

O estudo de estabilidade química e térmica por CD frente à variação do pH e da temperatura e também sob diversas concentrações de uréia foi realizado recentemente e, porém não foram incluídos, pois ainda necessitam de tratamento mais refinado.

III. 2.5.2 – Espectroscopia de Emissão de Fluorescência

O ambiente dos cinco resíduos de triptofano presentes na seqüência primária da rXfaroC foram monitorados pela espectroscopia de emissão de fluorescência a fim de complementar e confirmar os experimentos de CD quanto a integridade estrutural e a estabilidade em diferentes pH. Como mencionado anteriormente, mudanças na emissão de fluorescência revelam informações sobre a acessibilidade do solvente e sobre a hidrofobicidade do meio ao redor destes resíduos. O resíduo de triptofano atua como uma sonda espectral para a estrutura terciária da proteína (LAKOWICZ, 1983). Na seqüência primária da rXfaroC, os cinco resíduos de triptofano se localizam na parte central da seqüência proteína. O espectro de emissão de fluorescência intrínseca de rXfaroC em diferentes pH é mostrado na Figura 51.

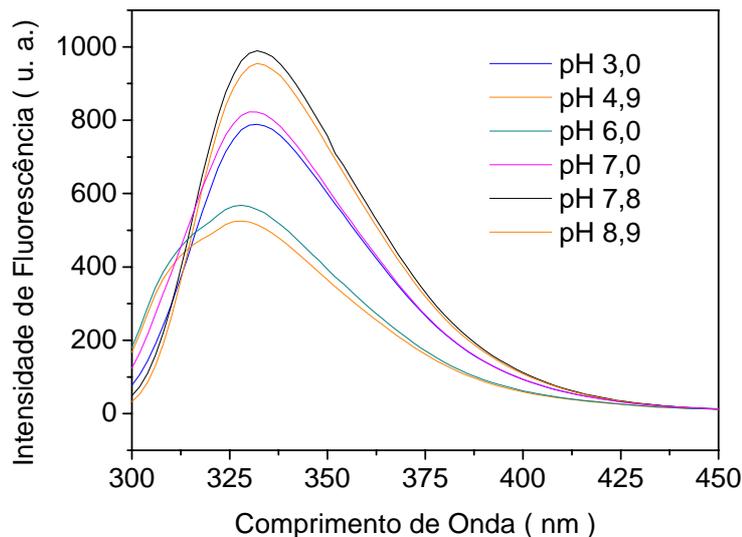


Figura 51. Espectros de emissão de fluorescência da rXfaroC a vários valores de pH. A solução tampão empregada foi acetato-borato-fosfato de sódio a 20 mM contendo NaCl a 150 mM. As medidas foram realizadas com excitação a 295 nm. A concentração utilizada para as amostras foi 10 μ M.

O máximo de emissão de fluorescência da rXfaroC foi em 332 nm (λ_{em}) no pH 7,8, indicativo de resíduos de triptofano internalizados na proteína. No entanto, o pico de emissão foi assimétrico, sugerindo que o ambiente dos cinco resíduos de Trp pode não ser idêntico, isto é, um ou mais resíduos provavelmente podem estar expostos a um ambiente mais polar (mais hidrofílico) na estrutura da molécula. Porém, isto pode ser confirmado apenas pela estrutura cristalográfica da proteína, um ensaio que ainda será objeto de estudo.

Conforme esperado, foi observado que a emissão de fluorescência de rXfaroC foi capaz de monitorar mudanças estruturais induzidas por pH. Em todos os pHs, a rXfaroC mostrou picos de emissão de fluorescência em torno de 330 nm, para uma excitação em 295 nm. Estes resultados indicaram que os resíduos de triptofano devem estar mais protegidos do meio polar e provavelmente a enzima se

mantém estruturada. Ao contrário das medidas de CD, as pequenas diferenças observadas na fluorescência em pHs intermediários (4,9 e 6,0), indicam que os resíduos de triptofano tendem a estar protegidos na estrutura interna da proteína mesmo sofrendo mudança conformacional de sua estrutura secundária, evidenciadas pelos espectros de dicroísmo circular.

Além disso, foi possível observar que mudanças na intensidade de fluorescência aconteceram com mudanças no pH do meio. Uma supressão de fluorescência intrínseca pôde ser atribuída onde a redução do rendimento de fluorescência foi resultante da distribuição de cargas e da mudança da carga líquida da proteína rXfaroC. Por exemplo, em proteínas monoméricas, a protonação progressiva dos resíduos de aminoácidos pelo decréscimo do pH dá origem a interações repulsivas que conduzem à perda das estruturas secundária e terciária. A estabilização das pontes salinas é removida e as cadeias laterais internalizadas no interior da estrutura tornam-se expostas. Proteínas que não foram desnoveladas completamente, no entanto, podem adotar estados conformacionais relativamente compactos, que diferem da estrutura nativa, conforme indicado por suas propriedades espectroscópicas (LAKOWICZ, 1983).

O estudo de estabilidade química por fluorescência frente a diversas concentrações de uréia foi realizado recentemente e ainda está em processo de análise.

III. 2.5.3 – Ensaio de Atividade Enzimática

A atividade da corismato sintase foi determinada estimando a produção do corismato em 340 nm pela via do ácido chiquímico a 25 °C. Esta foi uma via

alternativa para a medida da atividade da proteína rXfaroC devido a não disponibilidade comercial do principal substrato natural desta proteína, o 5-enolpiruvil chiquimato-3-fosfato (EPSP). O ácido corísmico proveniente da junção do ácido chiquímico e uma molécula de fosfoenolpiruvato gera o corismato que é utilizado na síntese de aminoácidos aromáticos (triptofano, fenilalanina e tirosina) (HERRMANN *et al.*, 1983).

Assim, a reação enzimática foi monitorada no pH 7,8, condição em que a rXfaroC se apresentou mais estável nos estudos estruturais, porém os demais pHs não foram avaliados. O interesse deste ensaio foi verificar se a rXfaroC apresentava atividade enzimática, o que é o aspecto mais interessante da tecnologia do DNA recombinante.

A atividade enzimática específica observada para rXfaroC foi $10,3 \text{ U mg}^{-1}$. A atividade aqui representada foi definida como a quantidade de enzima (em uma concentração $53,3 \text{ } \mu\text{g mL}^{-1}$) que catalisa a produção de 37 nmol de corismato por minuto. A atividade da enzima mostrou ser dependente dos cofactores NADH e FMN, já que não ocorreu nenhuma reação na ausência destes cofactores.

Estes resultados demonstram claramente que a proteína recombinante foi obtida na forma ativa. No entanto, o estudo enzimático nos demais pHs bem como a caracterização cinética e a determinação da estrutura tridimensional da rXfaroC serão concluídas brevemente a fim alcançar a completa caracterização da proteína e conseqüentemente proporcionar maiores contribuições no conhecimento de seu papel biológico.

III. 2.6 – Conclusões

As etapas de clonagem e a expressão heteróloga em *E. coli* BL21(DE3) resultaram em quantidades satisfatória de rXfaroC na forma solúvel e facilmente purificáveis.

A massa molecular para a proteína recombinante medida por SDS-PAGE foi de aproximadamente 45 kDa o que está de acordo com a massa molecular esperada.

A proteína recombinante de *X. fastidiosa* demonstrou ser uma molécula aparentemente estável através das medidas espectroscópica, evidenciando uma estrutura secundária contendo elementos em α -hélice e estruturalmente mais estável nos pHs extremos. A mudança de conformação com aparentes ausências dos mínimos no espectro de CD e a mudança nos máximos de emissão dos espectros de fluorescência, indica provavelmente que a estabilidade estrutural da enzima foi afetada.

O estudo da enzima corismato sintase visou sua contribuição como uma possível aplicação alternativa na área biotecnológica, uma vez que outras enzimas desta via são muito empregadas industrialmente na obtenção de compostos como o glifosato, muito utilizado na fabricação de importantes herbicidas agrícolas. O glifosato é um conhecido inibidor da via do chiquimato presente em plantas, fungos e bactérias que inibe a formação de aminoácidos como o triptofano e fenilalanina considerados essenciais no metabolismo destes organismos.

A possível inativação desta via pela inibição da corismato sintase deste estudo pode ser muito útil como uma fonte de referência na sua manipulação

futura como defensivos agrícolas. Assim como, poderá servir como um guia no desenho de moléculas alvo contra a colonização da *Xylella fastidiosa* em citrus.

III. 3 – Proteína D do Sistema de Transporte e Secreção Tat

III. 3.1 – Amplificação e Clonagem no Vetor pGEM-T *easy*

Como para as duas proteínas descritas nos capítulos anteriores, as suspensões bacterianas utilizadas na extração do DNA genômico também foram do isolado 9a5c da bactéria *X. fastidiosa*. A metodologia empregada para avaliar a qualidade e a concentração do DNA genômico obtido foi descrita na seção III. 1.1.

Como descrito anteriormente, os oligonucleotídeos utilizados na amplificação foram desenhados com base na sequência da ORF *mttC* e incorporados com os mesmos sítios de restrição das demais proteínas, *NdeI* e *XhoI*, preservando também seus códons originais de iniciação e terminação e permitindo posteriormente sua subclonagem nos vetores de expressão.

Como para os estudos anteriores, a temperatura de hibridização do par de iniciadores (F e R) desenhado foi calculada teoricamente através do software Gene Runner 3.05, buscando garantir a eficiência da reação de amplificação do DNA de interesse. A temperatura ótima encontrada neste caso ficou em torno dos 60 °C.

Com base na temperatura calculada para o par desenhado de oligonucleotídeos iniciadores foi possível a amplificação da ORF correspondente à proteína D do sistema de transporte e secreção Tat, porém com baixa concentração e presença de algumas bandas inespecíficas. Diante deste resultado, também se fez necessário o processo de otimização da PCR a fim de garantir o máximo de eficiência na obtenção dos produtos desejados. Assim, somente quando foi empregada a temperatura de hibridização dos *primers* de 58 °C foi possível a amplificação desejada (resultados não mostrados).

Resultados excelentes foram obtidos na amplificação da ORF de 807 pb. O produto de amplificação obtido pela PCR teve o seu tamanho concordante com o previsto, não apresentou nenhuma banda inespecífica, indicando o sucesso da otimização e resultou em quantidades suficientes para proceder à clonagem posterior.

Os produtos da amplificação por PCR foram devidamente purificados pelo *kit Wizard PCR Clean-up* da Promega, quantificados em espectrofotômetro e adenilados a fim de inseri-los no vetor de clonagem pGEM-T *easy*. A concentração obtida para estes produtos variou de 100 até 200 ng μL^{-1} . A construção nos vetores de clonagem foi denominada de pGEM-XftatD.

Os plasmídeos recombinantes pGEM-XftatD recém construídos foram inseridos por choque térmico em bactérias de *E. coli* DH5 α competentes. As bactérias transformadas foram semeadas em placas de Petri contendo meio LB com ágar, ampicilina, X-gal e IPTG e crescidas *overnight* em estufa a 37 °C.

Diversas colônias brancas e azuis cresceram na placa contendo o meio selecionado descrito acima. A seleção dos transformantes foi realizada então pela capacidade de crescer em presença do antibiótico ampicilina e através da coloração branca de alguns transformantes indicativa da presença do inserto dentro da região codificadora da enzima β -galactosidase. Dentre as colônias brancas resistentes, cinco colônias foram escolhidas aleatoriamente, inoculadas em tubos estéreis contendo meio LB líquido com ampicilina e submetidas à incubação *overnight* a 37 °C sob agitação. Após o crescimento individual destas colônias seus DNAs foram extraídos, quantificados, clivados por enzimas de restrição e finalmente seqüenciados.

III. 3.2 – Análise de Restrição e Seqüenciamento dos Plasmídeos Recombinantes

O DNA plasmidial dos cinco prováveis clones positivos extraídos através dos métodos descritos pelo *kit Wizard Miniprep*[®] e de lise alcalina empregados de acordo com o descrito em II.2.7, foram inicialmente considerados de boa qualidade e adequados para as análises de restrição e seqüenciamento. As amostras de DNA plasmidial extraídos pelo kit e pelo método de lise alcalina apresentaram em geral concentrações de 200 e 150 ng μL^{-1} , respectivamente.

A análise de restrição empregando as enzimas *NdeI* e *XhoI* foi então realizada a fim de confirmar a presença do inserto nos clones selecionados. A Figura 52 apresenta o perfil de restrição destes cinco clones selecionados.

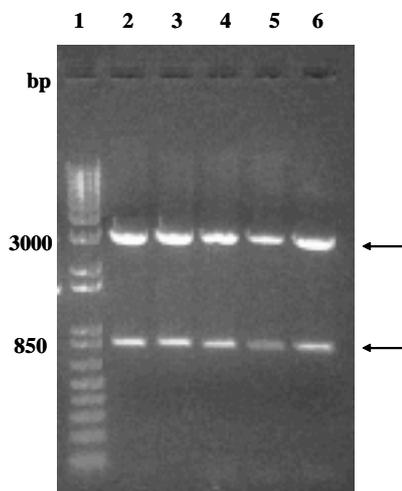


Figura 52. Análise de restrição dos plasmídeos pGEM-XftatD. A coluna 1 é o padrão de massa molecular 1 Kb Plus DNA Ladder da Gibco BRL. As colunas 2 a 6 são os clones positivos selecionados a serem utilizados na etapa de subclonagem nos vetores de expressão. O meio de separação foi gel de agarose 1 % em tampão 1x TAE, pH 8,3. A voltagem aplicada na separação foi 100 V por 30 min. Os fragmentos de DNA no gel foram visualizados em transiluminador UV a 365 nm. As bandas específicas da ORF XftatD e dos plasmídeos digeridos estão representados por setas cujos tamanhos são aproximadamente 3000 e 850 pb, respectivamente.

A presença e o tamanho dos insertos confirmaram o sucesso das etapas de clonagem e transformação das bactérias. Os produtos da digestão de 850 pb correspondente a ORF XftatD foram retirados do gel, purificados utilizando-se novamente o *kit Wizard PCR Clean-up*, quantificados e utilizados na etapa de subclonagem nos vetores de expressão. Estas amostras de DNA apresentaram em geral concentrações variando de 150 a 200 ng μL^{-1} .

Os DNAs plasmidiais de cinco clones foram selecionados para o seqüenciamento. As seqüências resultantes do gel de seqüenciamento destes clones também foram analisadas através do programa *Sequencing Analysis V 3*.

As seqüências resultantes foram submetidas a um processo de busca por similaridade em diversos bancos de dados internacionais, através de servidores de busca disponíveis na rede. O principal programa utilizado foi o software público BLAST cujos resultados mostraram similaridade de seqüência variando de 97 % de identidade frente à cadeia nucleotídica correspondente a ORF deste estudo, confirmando a presença da região alvo desejada.

De modo geral os clones seqüenciados apresentaram pequenas variações entre si, que podem ser atribuídos a erros de leitura ou mesmo a erros gerados durante a marcação para o seqüenciamento. No entanto, os plasmídeos pGEM-XftatD destes cinco clones apresentaram os genes de interesse inseridos na orientação e com fase de leitura corretas dentro dos vetores PGEM-T *easy* garantindo assim a integridade destes clones e a posterior leitura eficiente de todas as bases dos fragmentos na etapa de expressão protéica.

III. 3.3 – Subclonagem no Vetor de Expressão e Análise de Restrição dos Plasmídeos Recombinantes

Os fragmentos de DNA de 850 pb que codificam para a proteína rXftatD foram subclonados nos vetores pET28a. Esta nova construção denominada pET-XftatD também foi escolhida devido a expressar a proteína de interesse fusionada com uma cauda de seis histidinas. Novamente, os plasmídeos recombinantes recém construídos foram inseridos por choque térmico em bactérias de *E. coli* DH5 α competentes, semeadas em placas semelhantes ao processo anterior, porém somente com adição de canamicina, o novo antibiótico de seleção; e finalmente incubadas para o crescimento. A seleção das colônias neste caso foi realizada unicamente pela análise de restrição, já que o vetor pET28a não apresenta genes que as tornem de coloração diferenciada.

Após o período de crescimento celular, novamente cinco colônias foram selecionadas aleatoriamente dentre as inúmeras colônias presentes na placa de Petri, e inoculadas seguindo o mesmo processo descrito anteriormente. Após o crescimento individual destas células seus DNAs foram extraídos, quantificados e uma alíquota destes DNAs foram clivados por enzimas de restrição. Estas amostras de DNA também apresentaram concentrações variando de 100 a 150 ng μL^{-1} . Os DNAs extraídos restantes foram armazenados a $-20\text{ }^{\circ}\text{C}$ para eventuais necessidades.

A Figura 53 ilustra o novo perfil de restrição destes cinco clones selecionados.

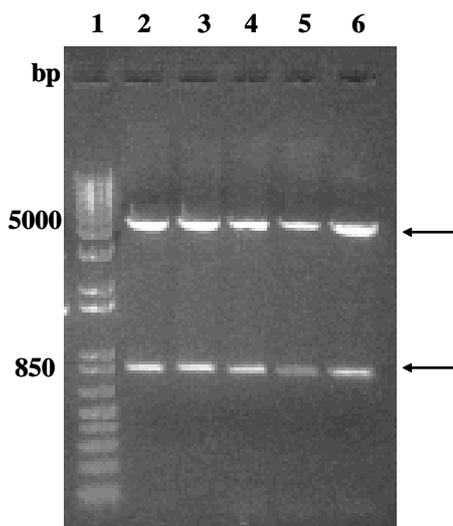


Figure 53. Análise de restrição dos plasmídeos pET-XftatD. A coluna 1 é o padrão de massa molecular 1 Kb Plus DNA Ladder da Gibco BRL. As colunas 2 a 6 são os clones positivos. O meio de separação foi gel de agarose 1 % em tampão 1x TAE, pH 8,3. A voltagem aplicada na separação foi 100 V por 30 min. Os fragmentos de DNA no gel foram visualizados em transiluminador UV a 365 nm. As bandas específicas das ORFs XftatD e dos plasmídeos digeridos estão representados por setas.

A presença dos insertos em todos os clones selecionados confirmou o sucesso das etapas de subclonagem e transformação das bactérias. Uma vez confirmada a presença do inserto, apenas os clones 1 e 2 (equivalentes as 2^o e 3^o canaletas do gel), os quais apresentaram maior intensidade de banda, foram utilizados para transformar as células competentes de expressão de *E. coli* BL21(DE3). A seleção de dois novos clones contendo as ORFs subclonadas nos vetores de expressão foi realizada após o plaqueamento das células em meio adequado. Finalmente, os inóculos destes clones foram empregados na produção da proteína recombinante em células adequadas à expressão.

III. 3.4 – Expressão Heteróloga da Proteína rXftatD a partir dos Vetores de Expressão

A produção da proteína recombinante foi realizada inicialmente por meio da indução da expressão por IPTG, a 37 °C. Passado o período de indução de 4 horas, as células bacterianas resultantes foram coletadas por centrifugação e lisadas na presença de tampão de lise tris-HCl. Os extratos protéicos resultantes foram avaliados quanto a sua solubilidade por SDS-PAGE.

A Figura 54 mostra o SDS-PAGE 15% no qual foram aplicadas as amostras para análise.

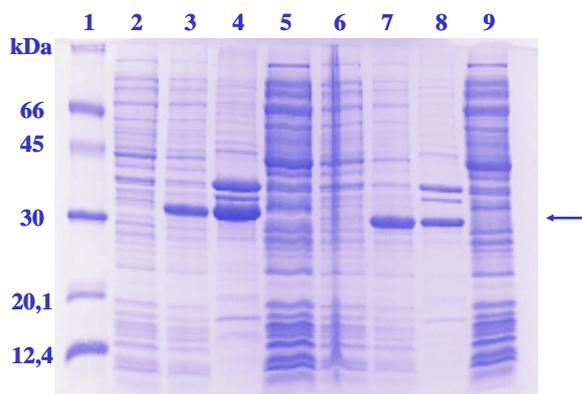


Figura 54. Expressão Piloto da proteína rXftatD a partir dos clones 1 e 2. A análise foi realizada em SDS-PAGE a 15%. Colunas 2 a 5 e 6 a 9 representam as expressões dos clones 1 e 2 contendo os construídos pET-rXftatD, respectivamente. As colunas 2 e 6 são os extratos protéicos totais não induzido dos sistemas de expressão. As colunas 3 e 7 são os extratos protéicos totais induzidos. As colunas 4 e 8 são as frações insolúveis após a lise celular. E as colunas 5 e 9 são as frações solúveis após a lise celular. A coluna 1 é o marcador de massa molecular. A seta indica a banda de expressão diferenciada na região de 30 kDa.

A presença de bandas adicionais nas frações induzidas 3 e 7 concordam com o tamanho esperado de 30 kDa para a proteína rXftatD, revelando a eficiência

da metodologia de expressão adotada. Porém, a rXftatD não foi observada na fração solúvel.

Como a cinética de produção das proteínas heterólogas está sujeita a variados fatores, incluindo tamanho e estrutura da proteína, sua solubilidade, transporte e tamanho, a expressão heteróloga da rXftatD foi realizada *overnight* a temperatura de 20 °C para averiguar se a proteína recombinante era expressa na forma solúvel.

Além disso, novas células de expressão como a Rosetta (DE3) pLys também foram transformadas com o plasmídeo pEt-XftatD a 20 °C. Estas células permitem a captação do indutor de forma lenta, permitindo que a expressão da proteína de interesse seja expressa lentamente favorecendo assim seu enovelamento correto.

A mudança de temperatura, do tempo de indução e da célula de expressão não foram suficientes para expressar a proteína rXftatD na forma solúvel.

Assim, como uma segunda alternativa de solubilização da rXftatD, a subclonagem da ORF *mttC* no vetor de expressão pET 32a foi realizada na tentativa de possibilitar sua expressão solúvel. Este vetor permite a expressão da proteína alvo em fusão com a proteína tireorredoxina na com a região N-terminal, uma proteína altamente solúvel, que catalisa a formação de pontes de dissulfeto. Porém, esta nova construção expressou a proteína novamente na forma insolúvel.

Certas proteínas alvo podem se acumular em agregados insolúveis na forma de corpos de inclusão. A formação de corpos de inclusão bacterianos são formalmente considerados ocorrer via associação não específica das superfícies hidrofóbicas promovida pela interação intermolecular específica sendo um fator muitas vezes dependente da seqüência (CLARK, 2001). Isto nem sempre é um

problema, pois a formação de corpos de inclusão pode proteger a proteína contra a degradação bacteriana e também facilitar a purificação, uma vez que são corpos densos, precipitados à baixa velocidade de centrifugação, enquanto a maior parte das proteínas bacterianas permanece no sobrenadante.

Diante destes resultados, estudos aplicados para solubilizar e renaturar a proteína rXftatD insolúvel se faziam necessários e foram inicialmente alvos deste trabalho. Porém, dada a incerteza do resultado e por requererem longos períodos de tempo para alcançar a renaturação adequada da proteína, de forma a mantê-la solúvel e funcional, estes estudos foram considerados inviáveis.

II. 3.5 – Conclusões

As etapas gerais de clonagem e a expressão heteróloga em *E. coli* foram bem sucedidas, porém a rXftatD somente foi obtida na forma insolúvel.

V – PERSPECTIVAS

O prosseguimento do trabalho com as proteínas de *Xylella fastidiosa* pode objetivar as seguintes tarefas:

- ✓ Realização de ensaios de cristalização, e uma vez obtidos cristais, utilizar-se da difração de raios-X para resolver suas estruturas tridimensionais.
- ✓ Ensaios de cinética enzimática com o intuito de se calcular os parâmetros cinéticos envolvidos para cada uma dessas enzimas;
- ✓ Determinação dos parâmetros termodinâmicos (ΔH , ΔG , ΔS , ΔC_p) envolvidos na interação destas mesmas enzimas, por ensaios de calorimetria, num microcalorímetro VP-ITC Microcal e VP-DSC.

V - REFERÊNCIAS BIBLIOGRÁFICAS

ACKERS, G.K. Molecular sieve studies of interacting protein systems. I: equations for transport of associating systems. **Journal of Biological Chemistry**, v. 242, p.3026–3034, 1967.

AHN; H.J.; YOON, H.; LEE, B.I.; SUH, S.W. Crystal structure of chorismate synthase: a novel FMN-binding protein fold and functional insights. **Journal of Molecular Biology**, v. 336, n.4, p.903-915, 2004.

ARCURI, H.L.; CANDURI, F.; PEREIRA, J.H.; SILVEIRA, N.J.F.; CAMERA JUNIOR, J.C.; OLIVEIRA, J.S.; BASSO, L.A.; PALMA, M.S.; SANTOS, D.S.; AZEVEDO JUNIOR, W.F.A. Molecular models for shikimate pathway enzymes of *Xylella fastidiosa*. **Biochemical and Biophysical Research Communications**, v. 320, p.979-991, 2004.

ASANO, Y.; TAMURA, K.; DOI, N.; UEATRONGCHIT, T.; H-KITTIKUN, A.; OHMIYA, T. Screening for new Hidroxynitrilases from plants. **Bioscience, Biotechnology and Biochemistry**, v. 69, n.12, p.2349-2357, 2005.

AUSUBEL, F.M.; BRENT, R.; KINGSTON, R.E.; MOORE, D.D.; SEIDMAN, J.G.; SMITH, J.A.; STRUTHL, K. **Short Protocols in Molecular Biology**. New York: John Wiley, 1995. 1512 p.

AZEVEDO, J.L.; MACCHERONI JUNIOR, W.; PEREIRA, J.O.; ARAÚJO, W.L. Endophytic microorganism: a review on insect control and recent advances on tropical plants. **Electronic Journal of Biotechnology**, v. 3, n.1, p.40-65, 2000.

BALSERA, M.; MENENDEZ, M.; SAIZ, J.L.; RIVAS, J.; AANDREU, J.M.; ARELLANO, J.B. Structural hability of the PsbQ protein of higher plant photosystem II. **Biochemistry**, v. 43, p.14171-14179, 2004.

BAUER, M.; GRIENGL, H.; STEINER, W. Parameters influencing stability and activity of a (S)-hydroxynitrile lyase from *Hevea brasiliensis* in two-phase systems. **Enzyme and Microbial Technology**, v. 24, p.514-522, 1999.

BAUER, M.; GRIENGL, H.; STEINER, W. Kinetics studies of the enzyme (S)-Hidroxynitrile lyase from *Hevea brasiliensis* using initial rate methods and progress curve analysis. **Biotechnology and Bioengineering**, v. 62, n.1, p.20-29, 1999.

- BURSTEIN, E.A.; VENDENKINA, N.S.; IVKOVA, M.N. Fluorescence and the location of tryptophan residues in proteins molecules. **Photochemistry and Photobiology**, v. 18, p.263-279, 1973.
- CAMPBELL, I.D.; DWEK, R.A. **Biological Spectroscopy**. California, Benjamin Cummings Publish, 1984. 404 p.
- CANTOR, C.R.; SCHIMMEL, P. R. **Biophysical Chemistry, part II: techniques for the study of biological structure and function**, New York: W. H. Freeman, 1980. 365 p.
- CASTRONUOVO, G. Proteins in aqueous solutions: calorimetric studies and thermodynamic characterization. **Thermochimica Acta**, v. 193, p.363-390, 1991.
- CLARK, E.D.B. Protein refolding for industrial processes. **Current Opinion in Biotechnology**, v. 12, p.202-207, 2001.
- CONN, E.E. Cyanogenic glycosides. **The Biochemistry of Plants**, New York:Academic Press, 1981. 798 p.
- COSTES, D.; WEHTJE, E.; ADLERCREUTZ, P. Hydroxynitrile liase-catalyzed of cyanohydrins in organic solvents; parameters influencing activity and enantiospecificity. **Enzyme and Microbial Technology**, v. 25, p.384-391, 1999.
- DAMMIN, J.; SVERGUN, D. I. **Journal of Biophysics**. v. 76, p. 2879-2886, 1999.
- DINAMARCA, M.A.; RUIZ-MANZANO, A.; ROJO, F. Inactivation of cytochrome o ubiquinol oxidase relieves catabolic repression of the *Pseudomonas putida* GPO1 alkane degradation pathway. **Journal of Bacteriology**, v. 184, n.14, p.3785-3793, 2002.
- EBBS, S. Biological degradation of cyanide compounds. **Current Opinion in Biotechnology**, v. 15, p.231-236, 2004.
- EFTINK, M.R. The use of fluorescence methods to monitor unfolding transitions in proteins. **Biophysical Journal**, v. 66, p.482-501, 1994.
- FECHTER, M.H.; GRIENGL, H. Hydroxynitrile lyases: biological sources and applications as biocatalysts. **Food Technology and Biotechnology**, v. 42, n.4, p.287-294, 2004.

FILLOUX, A.; MICHEL, G.; BALLY, M. GSP – dependent protein secretion in Gram-negative bacteria: the Xcp system of *Pseudomonas aeruginosa*. **FEMS Microbiology Reviews**, v.22, p.117-198, 1998.

FITZPATRICK, T.; RICKEN, S.; LANZER, M.; AMRHEIN, N.; MACHEROUX, P.; KAPPES, B. Subcellular localization and characterization of chorismate synthase in the apicomplexan *Plasmodium falciparum*. **Molecular Microbiology**, v. 40, n.1, p.65-75, 2001.

GALPERIN, M.Y.; KOOONIN, E.V. “Conserved hypothetical” proteins: prioritization of targets for experimental study. **Nucleic Acid Research**, v. 32, n.18, p.5452-5463, 2004.

GASTEIGER, E.; HOOGLAND, C.; GATTIKER, A.; DUVAUD, S.; WILKINS, M.R.; APPEL, R.D.; BAIROCH, A. **The proteomics protocols handbook**. New Jersey: Humana Press, 2005. p.571-607

GERRITS, P.J.; ZUMBRAGEL, F.; MARCUS, J. analyzing the hydrocyanation reaction: chiral HPLC and the synthesis of racemic cyanohydrins. **Tetrahedron**, v. 57, p.8691-8698, 2001.

GIRIGOSWAMI, A. S. A fluorimetric and circular dichroism of hemoglobin : effect of pH and anionic amphiphiles. **Journal of Colloid and Interface Science**, v. 296, p.324-331, 2006.

GLIEDER, A.; WEIS, R.; SKRANC, W.; POECHLAUER, P.; DREVENY, I.; MAJER, S.; WUBBOLTS, M.; SCHWAB, H.; GRUBER, K. Comprehensive step-by-step engineering of an (R)-Hydroxynitrile liase for large-scale asymmetric synthesis. **Angewandte Chemie**, v. 42 n.39, p.4815-4818, 2003.

GOLOVANOV, A.P.; HAUTBERGUE, G.M.; WILSON, S.A.; LIAN, L. A simple method for improving protein solubility and long-term stability. **Journal of the American Chemical Society**, v. 126, p.8933-8939, 2004.

GRIENGL, H.; KLEMPIER, N.; POCHLAUER, P.; SCHMIDT, M.; SHI, N.; ZABELINSKAJA-MACKOVA, A.A. Enzyme catalysed formation of (S)-cyanohydrins derived from aldehydes and ketones in a biphasic solvent system. **Tetrahedron**, v. 54, p.14477-14486, 1998.

GRIENGL, H.; HICKEL, A.; JOHNSON, D. V.; KRATKY, C.; SCHMIDT, M.; SCHWAB, H. Enzymatic cleavage and formation of cyanohydrins: a reaction of biological and synthetic relevance. **Chemical Communications**, v. 20, p.1933-1940, 1997.

GRÖGER, H. Enzymatic routes to enantiomerically pure aromatic α -hydroxy carboxylic acids: a further example for the diversity biocatalysis. **Advanced Synthesis & Catalysis**, v. 343, p.547-558, 2001.

HANEFELD, U.; STRAATHOF, A.J.J.; HEIJNEN, J.J. Study of the (S)- Hydroxynitrile lyase from *Hevea brasiliensis*: mechanistic implications. **Biochimica et Biophysica Acta**, v. 1432, p.185-193, 1999.

HARTUNG, J.S.; BERETTA, J.; BRLANSKY, R. H.; SPISSO, J.; LEE, R.F. Citrus variegated chlorosis bacterium: axenic culture, pathogenicity, and serological relationships with other strains of *Xylella fastidiosa*, **Phytopathology**, v. 84, p.591–597, 1994.

HASAN, N.; NESTER, E.W. Purification and properties of Chorismate synthase from *Bacillus subtilis*. **Journal of Biological Chemistry**, v. 253, n.14, p.4993-4998, 1978.

HASSLACHER, M.; SCHALL, M.; HAYN, M.; BONA, R.; RUMBOLD, K.; LUCKL, J.; GRIENGL, H.; KOHLWEIN, S.D.; SCHWAB, H. High level intracellular expression of Hydroxynitrile lyase from the tropical rubber tree *Hevea brasiliensis* in microbial hosts. **Protein Expression and Purification**, v. 11, p.61-71, 1997.

HASSLACHER, M.; SCHALL, M.; HAYN, M.; GRIENGL, H.; KOHLWEIN, S.D.; SCHWAB, H. Molecular cloning of the full-length cDNA of (S)-hydroxynitrile lyase from *Hevea brasiliensis*. **Journal of Biological Chemistry**, v. 271, n.10, p.5884-5891, 1996.

HAVEL, A.H. **Spectroscopic methods for determining protein structure in solution**. New York: John Wiley, 1996. 497 p.

HERRMANN, K.M. The shikimate pathway. **Annual Review of Plant Physiology and Molecular Biology**, v. 50, p.473-503, 1999.

HICKEL, A.; HASSLACHER, M.; GRIENGL, H. Hydroxynitrile lyases: functions and properties. **Plant Physiology**, v. 98, p.891-898, 1996.

HILLEBRAND, S.; GARCIA, W.; CANTÚ, M.D.; ARAÚJO, A.P.U.; TANAKA, M.; TANAKA, T.; GARRAT, R.C.; CARRILHO, E. In vitro monitoring of GTPase activity and enzyme kinetics studies using capillary electrophoresis. **Analytical and Bioanalytical Chemistry**, v. 383, p.92-97, 2005.

HUANG, D.T.; KAPLAN, J.; MENZ, R.I.; KATIS, V.L.; WAKE, R.G.; ZHAO, F.; WOLFENDEN, R.; CHRISTOPHERSON, R.I. Thermodynamic analysis of catalysis by the dihydroorotases from hamster and *Bacillus caldolyticus*: as compared with the un catalyzed reaction. **Biochemistry**, v. 45, p.8275-8283, 2006.

HUGHES, J.; CARVALHO, J.P.C.; HUGHES, M.A. Purification, characterization, and cloning of α -Hydroxynitrile lyase from Cassava (*Manihot esculenta* Crantz). **Archives of Biochemistry and Biophysics**, v. 311, n.2, p.496-502, 1994.

HUGHES, J.; KERESZTESSY, Z.; BROWN, K.; SUHANDONO, S.; HUGHES, M.A. Genomic organization and structure of α -Hydroxynitrile lyase in cassava (*Manihot esculenta* Crantz). **Archives of biochemistry and biophysics**, v. 356, n.2, p.107-116, 1998.

HUGHES, J.; LAKEY, J.H.; HUGHES, M.A. Production and characterization of a plant α -Hydroxynitrile lyase in *Escherichia coli*. **Biotechnology and bioengineering**, v. 53, n.3, p.332-338, 1997.

JEVSEVAR, S.; GABERC-POREKAR, V.; FONDA, I.; PODOBNIK, B.; GRDADOLNIK, J.; MENART, V. Production of nonclassical inclusion bodies from which correctly folded protein can be extracted. **Biotechnology**, v. 21, p.632-639, 2005.

KELLERMANN, G.; VICENTIN, F.; TAMURA, E.; ROCHA, M.; TOLENTINO, H.; BARBOSA, A.; CRAIEVICH, A.; TORRIANI, I. Journal Applied. Crystallography, v. 30, 880-883, 1997.

KITZING, K.; AUWETER, S.; AMRHEIN, N.; MACHEROUX, P. Mechanism of chorismate synthase. **Journal of Biological Chemistry**, v. 279, n.10, p.9451-9461, 2004.

KITZING, K.; MACHEROUX, P.; AMRHEIN, N. Spectroscopic and kinetic characterization of the bifunctional chorismate synthase from *Neurospora crassa*. **Journal of Biological Chemistry**, v. 276, n.46, p.42658-42666, 2001.

KOBAWILA, S.C.; LOUEMBE, D.; KELEKE, S.; HOUNHOUIGAN, J.; GAMBA, C. Reduction of the cyanide content durin fermentatios of cassava roots and leaves to produce blkedi and ntoba mbodi, two food products from Congo. **African Journal of Biotechnology**, v. 4, n.7, p.689-696, 2005.

KRUSE, C.G. Chiral cyanohydrins: their manufacture and utility as chiral building blocks. In: COLLINS, A.N.; SHELDRAKE, G.N.; CROSBY, Y.J. **Chirality in Industry**. New York: John Wiley, 1992. p.279-299.

LADOKHIN, A.S. Fluorescence spectroscopy in peptide and protein analysis. In: **ENCYCLOPEDIA of analytical chemistry**. Chichester: John Wiley, 2000. p.5762-5779

LAEMMLI, U.K.; MOLBERT, E.; SHOWE, M.; KELLENBERGER, E. Form-determining function of the genes required for the assembly of the head of bacteriophage T4, **Journal of Molecular Biology**, v. 49, p.99-113, 1970.

LAKOWICZ, J.R. **Principles of fluorescence spectroscopy**. New York: Plenum Press, 1983. 698 p.

LAURENT, T.C.; KILLANDER, J. A theory of gel filtration and its experimental verification. **Journal of Chromatography**. v. 14, p.317–330, 1964.

LEWIN, B. **Genes V**. Oxford. Oxford University Press and Cell Press, 1994. 1272p.

LEWIN, B. **Genes VI**. New York. Oxford University Press and Cell Press, 1997. 1260p.

LORCK, H. Production of hydrocyanic acid by bacteria. **Physiologia Plantarum**, v. 1, p.142-146, 1948.

MACHEROUX, P.; SCHMID, J.; AMRHEIN, N.; SCHALLER, A. A unique reaction in a common pathway: mechanism and function of chorismate synthase in the shikimate pathway. **Planta**, v. 207, p.325-334, 1999.

MALLAM, A. L.; JACKSON, S. E. Folding Studies on a Knotted *Protein* **Journal of Molecular Biology**, 346, p. 1409–1421, 2005.

MCCONKEY, G.A.; PINNEY, J.W.; WESTHEAD, D.R., PLUECKHAHN, K.; FITZPATRICK, T.B.; MACHEROUX, P.; KAPPES, B. Annotating the *Plasmodium* genome and the enigma of the shikimate pathway. **Trends in Parasitology**, v. 2, n.2, p.60-65, 2004.

MCINTOSH, K.A.; CHARMAN, W.N.; CHARMAN, S.A. The application of capillary electrophoresis for monitoring effects of excipients on protein confirmation. **Journal of Pharmaceutical and Biomedical Analysis**, v. 16. p.1097-1105, 1998.

MEIDANIS, J.; BRAGA, M.D.V.; VERJOVSKI-ALMEIDA, S. Whole-genome analysis of transporters in the plant pathogen *Xylella fastidiosa*. **Microbiology and Molecular Biology Reviews**, v. 66, n.2, p.272-299, 2002.

OHNO, H.; BLACKWELL, J.; JAMIESON, A.M.; CARRINO, D.A.; CAPLAN, A.I. Calibration of the relative molecular mass of proteoglycan subunit by column chromatography on Sepharose CL-2B. **Biochemical Journal**, v. 235, p.553-557, 1986.

PACOLLA-MEIRELES, L. D.; VALARINI, M. J.; AZEVEDO, J. L.; ALFENAS, A. C. **Manual de técnicas eletroforéticas em microorganismos**. Piracicaba: Fundação de Estudos Agrários Luiz de Queiroz, 1988. 54p.

PAIN, R.H. Determining the fluorescence spectrum of a protein. In: **Current Protocols in Protein Science**, New York: John Wiley, 1996. p.7.7.1-7.7.19

PETOUKHOV, M. V.; SVERGUN, D. I. **Journal of Applied Crystallography**. v 36, p. 540-544, 2003.

POLLARD, J. W.; WALKER, J. M. **Methods in Molecular Biology**, New Jersey: Humana Press, 1990. 597p.

PONGDEE, R.; LIU, H. Elucidation of enzyme mechanisms using fluorinated substrate analogues. **Bioorganic Chemistry**, v. 32, n.5, p.393-437, 2004.

PURCELL, H. A.; HOPKINS, D. L. Fastidious xylem limited bacterial plant pathogens, **Annual Review of Phytopathology**. v. 34, p.131-151, 1996.

PURDIE, N.; SWALLOWS, K.A.; MURPHY, L.H.; PURDIE, R.B. Circular dichroism-II: analytical applications. **Trends in Analytical Chemistry**, v. 9, n.4, p.136-141, 1990.

PURDIE, N.; SWALLOWS, K.A. Circular dichroism-I: theory and practice. **Trends in Analytical Chemistry**, v. 9, n.3, p.94-97, 1990.

RAMJEE, M.K.; COGGINS, J.R.; THORNELEY, N.F. A continuous, anaerobic spectrophotometric assay for chorismate synthase activity that utilizes photoreduced flavin mononucleotide. **Analytical Biochemistry**, v. 220, p.137-141, 1994.

ROBERTS, F.; ROBERTS, C.W.; JOHNSON, J.J.; KYLE, D.E.; KRELL, T.; COGGINS, J.R.; COOMBS, G.H.; MILHOUS, W.K.; TZIPORI, S.; FERGUSON, D.J.P.; CHAKRABARTI, D.; MCLEOD, R. Evidence for the shikimate pathway in apicomplexan parasites. **Nature**, v. 393, p.801-805, 1998.

SAMBROOK, J.; FRITSCH, E.F.; MANIATIS, T. **A Laboratory Manual in Molecular Cloning**. New York: Cold Spring Harbor, 1989. pp.10.6-10.12, 10.66-10.67.

SARGENT, F. et al. Overlapping functions of components of a bacterial Sec-independent protein export pathway. **Embo Journal**, v.17, p.3640-3650, 1998.

SCHALLER, A.; AFFERDEN, M.V.; WINDHOFER, V.; BULOW, S.; ABEL, G.; SCHMID, J.; AMRHEIN, N. Purification and characterization of chorismate synthase from *Euglena gracilis*. **Plant Physiology**, v. 97, p. 1271-1279, 1991.

SELMAR, D.; CARVALHO, F.J.P., CONN, E.E. A colorimetric assay for α -Hydroxynitrile lyase. **Analytical Biochemistry**, v. 166, p.208-211, 1987.

SVERGUN, D.I.; SEMENYUK, A. V.; FEIGIN, L. A. **Acta crystallography**, v. 44, p. 244-250, 1988.

SVERGUN, D.I. **Journal of Applied Crystallography**. v. 25, p. 495-503, 1992.

SHARMA, M.; SHARMA, N.N.; BHALLA, T.C. Hydroxynitrile lyases: at the interface of biology and chemistry. **Enzyme and Microbial Technology**, v. 37, p.279-294, 2005.

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

SIRITUNGA, D.; ARIAS-GARZON, D.; WHITE, W.; SAYRE, R.T. Over-expression of hydroxynitrile lyase in transgenic cassava roots accelerates cyanogenesis and food detoxification. **Plant Biotechnology Journal**, v. 2, p.37-43, 2004.

SMOLKA, M.B.; MARTINS, D.; WINCK, F.V.; SANTORO, C.E.; CASTELLARI, R.R.; FERRARI, F.; BRUM, I.J.; GAIEMBECK, E.; COLETTA-FILHO, H.D.; MACHADO, M. A.; MARANGONI, S.; NOVELLO, J.C. Proteome analysis of the plant pathogen *Xylella fastidiosa* reveals major cellular and extracellular proteins and a peculiar codon bias distribution. **Proteomics**, v .3, p.224-237, 2003.

SMOLKA, M.B.; MARTINS, D.; WINCK, F.V.; SANTORO, C.E.; CASTELLARI, R.R.; FERRARI, F.; GAIEMBECK, E.; COLETTA FILHO, H.D.; MACHADO, M.A.; MARANGONI, S.; NOVELLO, J.C. Proteome analysis of the plant pathogen *Xylella fastidiosa* reveals major cellular and extracellular proteins and a peculiar codon bias distribution. **Proteomics**, v. 3, p.224-237, 2003.

SOUZA, A.A.; TAKITA, M.A.; COLETTA FILHO, H.D.; CALDANA, C.; GOLDMAN, G.H.; YANAI, G.M.; MUTO, N.H.; OLIVEIRA, R.C.; NUNES, L.R.; MACHADO, M.A. Analysis of gene expression in two growth states of *Xylella fastidiosa* and its relationship with pathogenicity. **Molecular Plant-Microbe Interactions**, v. 16, n.10, p.867-875, 2003.

SPRINSON, D.B.; LEVIN, J.G. The enzymatic formation and isolation of 3-enolpyruvylshikimate 5-phosphate. **Journal of Biological Chemistry**, v. 239, n.4, p.1142-1150, 1964.

SREERAMA, N.; VENYAMINOV, S.Y.; WOODY, R.W. Estimation of the number of alpha-helical and beta-strand segments in proteins using circular dichroism spectroscopy, **Protein Science**, v. 8, p.370-80, 1999.

SREERAMA, N.; WOODY, R. W. A self-consistent method for the analysis of protein secondary structure from circular dichroism. **Analytical Biochemistry**, v. 209, p. 32-44, 1993.

SREERAMA, N.; WOODY, R. W. Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. **Analytical Biochemistry**, v. 287, p.252-260, 2000.

STRYER, L. **Biochemistry**. New York. W.H. Freeman and Company, 1988. 881p.

SWIETNICKI, W.; PETERSEN, R.; GAMBETTI, P.; SUREWICZ, W.K. pH-dependent stability and conformation of the recombinant human prion protein PrP (90-231). **The Journal of Biological Chemistry**, v. 272, n.44, p.27517-27520, 1997.

TSUMOTO, K.; EJIMA, D.; KUMAGAI, I.; ARAKAWA, T. Practical considerations in refolding proteins from inclusion bodies. **Protein Expression and Purification**, v. 28, p.1-8, 2003.

VENTURA, S. Sequence determinants of proteins aggregation: tools to increase protein solubility. **Microbial Cell Factories**, v. 4, p.11-18, 2005.

VENYAMINOV, S.Y.; YANG, J.T. In: FASMAN, G.D. **Circular dichroism and the conformational analysis of biomolecules**, New York: Plenum Press, 1996. P.69-107

VETTER, J. Plant cyanogenic glycosides. **Toxicon**, v. 38, p.11-36, 2000.

WAJANT, H.; EFFENBERGER, F. Hydroxynitrile lyases of higher plants. **Biological Chemistry**, v. 377, p.611-617, 1996.

WAJANT, H.; PFIZENMAIER, K. Identification of potential active-site residues in α -Hydroxynitrile lyase from *Manihot esculenta* by site-directed mutagenesis. **Journal of Biological Chemistry**, v. 271, n.42, p.25830-25834, 1996.

WATANABE, A.; YANO, K.; IKEBUKURO, K.; KARUBE, I. cyanide hydrolisis in a cyanide-degrading bacterium, *Pseudomonas stutzeri* AK61, by cyanidase. **Microbiology**, v. 144, p.1677-1682, 1998.

WEINER, J.H. et al. A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. **Cell**, v.93, p. 93-101, 1998.

WHITE, W.L.B.; ARIAS-GARZON, D.I.; MCMAHON, J.M.; SAYRE, R.T. Cyanogenesis in Cassava. **Plant Physiology**, v. 116, p.1219-1225, 1998.

WOOD, A.P.; KELLY, D.P.; MCDONALD, I.R.; JORDAN, S.L.; MORGAN, T.D.; KHAN, S.; MURREL, J.C.; BORODINA, E. A novel pink-pigmented facultative methyltroph, *Methylobacterium thiocyanatum* sp. nov., capable of growth on thiocyanate as sole nitrogen sources. **Archives of Microbiology**, v. 168, p. 148-158, 1998.

YAMAMOTO, K.; OISHI, K.; FUJIMATSU, I.; KOMATSU, K. Production of R(-)-mandelic acid from mandelonitrile by *Alcaligenes faecalis* ATCC 8750. **Applied Environmental Microbiology**, v. 57, n.10, p.3028-3032, 1991.

YEOW, Y.L.; LEONG, Y.K.; CHEAH, M.Y.; DANG, H.D.T.; LAW, C.K.P. Cleaving of mandelonitrile catalyzed by S-hydroxynitrile lyase from *Hevea brasiliensis* – a kinetic investigation based on the rate curve method. **Journal of Biotechnology**, v. 111, p.31-39, 2004.

ZAHA, A.; SCHRANK, A.; FERREIRA, H.B.; SCHRANK, I.S.; RODRIGUES, J.J.S.; REGNER, L.P.; PASSAGLIA, L.M.P.; ROSSETTI, M.L.R.; RAUPP, R.M.; SILVA, S.C.; GAIESKY, V.L.V. **Biologia Molecular Básica**. Porto Alegre: Mercado Aberto, 1996. 336p.

ZHONG, G.Y.; RIOV, J.; GOREN, R.; HOLLAND, D. Competitive hybridization: theory and application in isolation and quantification of differentially regulated genes. **Analytical Biochemistry**, v. 282, n.1, p.129-135, 2000.

ZHUANG, X.; HA, T.; KIM, H.D.; CENTNER, T.; LABEIT, S.; CHU, S. Fluorescence quenching: a tool for single-molecule protein-folding study. **Proceedings of the National Academy of Sciences**, v. 97, n.26, p.14241-14244, 2000.