



Universidade de São Paulo
Instituto de Física de São Carlos
Centro de Biotecnologia Molecular Estrutural



Estudos moleculares de duas triptofanil tRNA sintetases do parasita *Leishmania major* e de uma cisteíno protease da bactéria *Xylella fastidiosa*.

Ney Ribeiro Leite

Tese apresentada ao programa de pós-graduação do Instituto de Física de São Carlos, da Universidade de São Paulo, para obtenção do título de Doutor em Física Aplicada - área de concentração: Física Biomolecular.

Orientador: Prof. Dr. Otavio Henrique Thiemann

Julho – 2007
São Carlos – SP

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTA
TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO,
PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pelo Serviço de Biblioteca e Informação IFSC/USP

Leite, Ney Ribeiro

Estudos moleculares de duas triptofanil tRNA sintetases do parasita *Leishmania major* e de uma cisteíno protease da bactéria *Xylella fastidiosa*./Ney Ribeiro Leite; orientador Otavio Henrique Thiemann. São Carlos, 2007.

147 f.

Tese (Doutorado – Área de concentração: Física Aplicada – Opção: Física Biomolecular) – Instituto de Física de São Carlos da Universidade de São Paulo.

1. Triptofanil tRNA sintetase. 2. tRNA. 3. Cisteíno proteases. 4. *Xylella fastidiosa*. 5. Cristalografia de proteínas. I. Título.

à minha linda esposa Micheli,
e à minha querida família.

Consolo Na Praia

*Vamos, não chores.
A infância está perdida.
A mocidade está perdida.
Mas a vida não se perdeu.*

*O primeiro amor passou.
O segundo amor passou.
O terceiro amor passou.
Mas o coração continua.*

*Perdeste o melhor amigo.
Não tentaste qualquer viagem.
Não possuis carro, navio, terra.
Mas tens um cão.*

*Algumas palavras duras,
em voz mansa, te golpearam.
Nunca, nunca cicatrizam.
Mas, e o humour?*

*A injustiça não se resolve.
À sombra do mundo errado
murmuraste um protesto tímido.
Mas virão outros.*

*Tudo somado, devias
precipitar-te, de vez, nas águas.
Estás nu na areia, no vento...
Dorme, meu filho.*

(Carlos Drummond de Andrade - Antologia Poética)

AGRADECIMENTOS

Ao Prof. Dr. Otavio Henrique Thiemann pela orientação, liberdade de trabalho, amizade, respeito.

À minha Micheli, que é a parte mais importante de tudo que conquistei até hoje, “ Sem você eu não vivo ...”

Aos meus grandes e especiais amigos.

À todo o grupo de Cristalografia de Proteínas do Centro de Biotecnologia Molecular Estrutural, pelo auxílio na execução do projeto de doutoramento e pela agradável convivência.

À minha família e em especial aos parentes mais próximos.

À todo o pessoal do Instituto de Física de São Carlos/USP.

À FAPESP pelo auxílio financeiro (processo # 02/01883-9).

Sumário

LISTA DE FIGURAS	I
LISTA DE TABELAS	III
RESUMO.....	IV
ABSTRACT	V

Parte I - Caracterização molecular e estrutural de duas enzimas triptofanil tRNA sintetase de *Leishmania major*.

CAPÍTULO 1.....	1
Introdução.....	1
Leishmaniose.....	1
Editoração dos RNA transportadores mitocondriais (kinetoplastos).....	6
Aminoacil-tRNA sintetases	8
Triptofanil-tRNA sintetase.....	10
Triptofanil-tRNA sintetases em <i>L. major</i>	12
Referências do capítulo	12
CAPÍTULO 2.....	16
Objetivos	16
CAPÍTULO 3.....	17
Metodologia	17
Caracterização molecular dos genes <i>LmTrpRS1</i> e <i>LmTrpRS2</i>	17
Caracterização das regiões 5' e 3' não traduzidas (UTR) dos genes <i>LmTrpRS1</i> e <i>LmTrpRS2</i> ..	17
Caracterização do número de cópias dos genes <i>LmTrpRS1</i> e <i>LmTrpRS2</i>	20
Clonagem dos genes <i>LmTrpRS1</i> e <i>LmTrpRS2</i>	20
Sub-clonagem dos genes <i>LmTrpRS1</i> e <i>LmTrpRS2</i> em vetores de expressão e ensaios de expressão protéica.....	22
Produção de anticorpos policlonais contra as proteínas <i>LmTrpRS1</i> e <i>LmTrpRS2</i>	23
Localização celular das proteínas <i>LmTrpRS1</i> e <i>LmTrpRS2</i>	24
Expressão e Purificação da proteína <i>LmTrpRS2</i>	25
Ensaio de atividade enzimática da enzima <i>LmTrpRS2</i> recombinante.....	26
Transcrição in vitro dos tRNAs ^{Trp} editado e não editado.....	26
Purificação da transcrição in vitro dos tRNAs ^{Trp} e protocolo de re-enovelamento.....	27
Caracterização estrutural das proteínas <i>LmTrpRS1</i> e <i>LmTrpRS2</i>	27
Modelagem molecular das proteínas <i>LmTrpRS1</i> e <i>LmTrpRS2</i>	27
Referências do capítulo	28
CAPÍTULO 4.....	30
Resultados e discussões.....	30
Caracterização molecular dos genes <i>LmTrpRS1</i> e <i>LmTrpRS2</i>	30
Caracterização das regiões 5' e 3' não traduzidas (UTR) dos genes <i>LmTrpRS1</i> e <i>LmTrpRS2</i> ..	30
Caracterização do número de cópias dos genes <i>LmTrpRS1</i> e <i>LmTrpRS2</i>	36
Clonagem dos genes <i>LmTrpRS1</i> e <i>LmTrpRS2</i>	37
Sub-clonagem dos genes <i>LmTrpRS1</i> e <i>LmTrpRS2</i> em vetores de expressão e ensaios de expressão protéica.....	41
Produção de anticorpos policlonais contra as proteínas <i>LmTrpRS1</i> e <i>LmTrpRS2</i>	48
Localização celular das proteínas <i>LmTrpRS1</i> e <i>LmTrpRS2</i>	50
Expressão e Purificação da proteína <i>LmTrpRS2</i>	51
Ensaio de atividade enzimática da enzima recombinante.....	52
Transcrição in vitro do tRNA ^{Trp} editado e não editado, purificação e protocolo de re-enovelamento desses tRNAs ^{Trp}	53
Caracterização estrutural das proteínas <i>LmTrpRS1</i> e <i>LmTrpRS2</i>	56
Modelagem molecular das proteínas <i>LmTrpRS1</i> e <i>LmTrpRS2</i>	57
Considerações finais.....	70

Referências do capítulo	73
APÊNDICE I	76
Materiais utilizados em biologia molecular.	76

APÊNDICE II.....	84
Artigo publicado na revista Memórias Instituto Oswaldo Cruz.....	84

Parte II - Estudos estruturais da Xylellaína, uma cisteíno protease da bactéria *Xylella fastidiosa*.

CAPÍTULO 1.....	92
------------------------	-----------

Introdução.....	92
<i>Xylella fastidiosa</i>	92
Clorose Variegada dos Citros ou Amarelinho	93
Proteases.....	95
Cisteíno proteases.....	96
Xylellaína	97
Referências do capítulo	97

CAPÍTULO 2.....	100
------------------------	------------

Objetivos	100
------------------------	------------

CAPÍTULO 3.....	101
------------------------	------------

Metodologia	101
Estudos iniciais da Xylellaína	101
Expressão e purificação da Xylellaína.....	101
Cromatografia de exclusão molecular e espalhamento dinâmico de luz	102
Ensaio de cristalização	103
Coleta e processamento de dados de difração de raios X.....	103
Resolução, refinamento e análise da estrutura tridimensional da Xylellaína.....	104
Seqüenciamento do N-terminal da proteína Xylellaína.....	105
Referências do capítulo	105

CAPÍTULO 4.....	107
------------------------	------------

Análise Estrutural e Discussões.....	107
Expressão, purificação e cristalização da Xylellaína	107
Coleta e processamento de dados.....	112
Resolução e refinamento da estrutura da Xylellaína	115
Estrutura cristalográfica da Xylellaína	118
Interface de dimerização da Xylellaína	119
Enovelamento da Xylellaína.....	121
Sítio ativo da Xylellaína	122
Interação da região N-terminal da Xylellaína com o sulco de ligação ao substrato, mecanismo de inibição enzimática, pró-Xylellaína	123
Sítio de ligação do ribonucleotídeo uridina di-fosfato.....	127
Considerações finais.....	132
Referências do capítulo	134

APÊNDICE III	137
---------------------------	------------

Colaborações realizadas durante o período do doutorado.	137
1- Estudos Estruturais de uma Proteína de ligação a moléculas odoríficas (Odorant Binding Protein – OBP) de <i>Aedes aegypti</i>	138
2- Estudos Estruturais de uma Glyoxalase II de <i>Salmonella typhimurium</i>	142

APÊNDICE IV	146
--------------------------	------------

Artigo publicado na revista “Biochemistry”.....	146
--	------------

Lista de Figuras

Parte I - Caracterização molecular e estrutural de duas enzimas triptofanil tRNA sintetase de *Leishmania major*.

Capítulo 1

Figura I-1.1: Microscopia ótica e de varredura de macrófagos e parasitas do gênero <i>Leishmania</i> .	2
Figura I-1.2: Algumas manifestações clínicas das leishmanioses.	3
Figura I-1.3: Ciclo de vida dos parasitas do gênero <i>Leishmania</i> .	4
Figura I-1.4: Esquema do mecanismo de importação, exportação e editoração de tRNAs mitocondriais em Kinetoplastidas.	7
Figura I-1.5: Mecanismo de reação das AARSs.	9
Figura I-1.6: Representação de superfície de aminoacil-tRNA sintetases ligadas aos tRNA cognatos.	10

Capítulo 3

Figura I-3.1: Estratégia adotada para a caracterização das regiões não traduzidas dos genes <i>LmTrpRs1</i> e <i>LmTrpRs2</i> .	18
Figura I-3.3: Esquema do programa de PCR utilizado para amplificação e sub-clonagens dos genes <i>LmTrpRS1</i> e <i>LmTrpRS2</i> .	21
Figura I-3.4: Esquema representativo do cassete de transcrição do tRNA ^{Trp} .	26

Capítulo 4

Figura I-4.1: Diagrama mostrando os 100 melhores alinhamento da proteína LmTrpRS1 com seqüências de outras proteínas depositadas no banco de dados.	31
Figura I-4.2: Caracterização das regiões não traduzidas do gene <i>LmTrpRS1</i> .	32
Figura I-4.3: Diagrama mostrando composição mRNA do gene <i>LmTrpRS1</i> caracterizado por RT-PCR.	33
Figura I-4.4: Seqüência nucleotídica da porção 5' do gene <i>LmTrpRS1</i> .	34
Figura I-4.5: Caracterização das regiões não traduzidas do gene <i>LmTrpRS2</i> .	35
Figura I-4.6: Seqüência nucleotídica da porção 5' do gene <i>LmTrpRS2</i> .	35
Figura I-4.7: Diagrama mostrando composição mRNA do gene <i>LmTrpRS2</i> caracterizado por RT-PCR.	35
Figura I-4.8: Caracterização do número de cópias dos genes <i>LmTrpRS1</i> e <i>LmTrpRS2</i> .	37
Figura I-4.9: Gel em agarose a 1% com resultado da amplificação do gene <i>LmTrpRS1</i> .	38
Figura I-4.10: Esquema da região da região de sinalização mitocondrial da proteína LmTrpRS2 predita pelo programa MitoprotII.	39
Figura I-4.11: Gel em agarose a 1% com resultado da amplificação do gene <i>LmTrpRS2</i> de <i>L. major</i> .	40
Figura I-4.12: Diagrama das construções realizadas com o gene <i>LmTrpRS1</i> e <i>LmTrpRS2</i> .	43
Figura I-4.13: Digestão com endonucleases de restrição das construções realizadas com o gene <i>Lmwars1</i> nos vetores bacterianos.	44
Figura I-4.14: Resultados da expressão da proteína LmTrpRS1 em vários vetores de expressão bacterianos.	45
Figura I-4.15: Resultado da sub-clonagem do gene <i>LmTrpRS2-24A</i> no vetor pET28a e a expressão da proteína LmTrpRS2 nesse vetor.	46
Figura I-4.16: Reação de western-blotting com os anticorpos anti-LmTrpRS1 e Anti-LmTrpRS2 contra as proteínas LmTrpRSs.	50
Figura I-4.17: "Western blotting" de extratos celulares de <i>L. tarentolae</i> contra os anticorpos anti-LmTrpRSs.	51
Figura I-4.18: Expressão e purificação da proteína LmTrpRS2 sem a região de sinalização mitocondrial no vetor de expressão pET28a na coluna de cromatografia de afinidade Ni-NTA.	52
Figura I-4.19: Reação enzimática catalisada pelas AARSs.	53
Figura I-4.20: Gel de formaldeído agarose 2% mostrando a transcrição do tRNA ^{Trp} .	54
Figura I-4.21: Ensaio de aminoacilação da enzima LmTrpRS2.	56
Figura I-4.22: Alinhamento seqüencial entre as LmTrpRSs e a HsTrpRS1.	59
Figura I-4.23: Alinhamento seqüencial entre as LmTrpRSs. LmTrpRS1 e LmTrpRS2, triptofanil tRNA sintetase de <i>L. major</i> citoplasmática e mitocondrial respectivamente.	60

Figura I-4.24: Diagrama de Ramachandran dos modelos finais obtido da proteína LmTrpRS1 (A) e LmTrpRS2	62
Figura I-4.25: Representação em cartoons dos modelos obtidos para as proteínas LmTrpRS1 e LmTrpRS2.....	63
Figura I-4.26: Representação em cartoons das LmTrpRSs na forma de dímeros.	64
Figura I-4.27: Representação de cartoons da sobreposição da estrutura da HsTrpRS1 com a estrutura modelada da LmTrpRS2.....	66
Figura I-4.28: Representação em “cartoons” da sobreposição da estrutura cristalográfica da proteína HsTrpRS1 (alaranjado) complexada com tRNA ^{Trp} (azul) e da proteína LmTrpRS1 modelada (verde).....	68
Figura I-4.29: Representação em cartoon da região de interação das proteínas LmTrpRS1 e LmTrpRS2.....	70

Parte II - Estudos estruturais da Xylellaina, uma cisteína protease da bactéria *Xylella fastidiosa*.

Capítulo 1

Figura II-1.1: Microscopia eletrônica vasos do xilema mostrando a distribuição das colônias de <i>X. fastidiosa</i>	92
Figura II-1.2: Sintomas da CVC em folhas e frutos de citrus.	94

Capítulo 4

Figura II-4.1: Gel em poliacrilamida a 15% mostrando a expressão e purificação da Xylellaina em cromatografia de afinidade.	108
Figura II-4.2: Ensaio de cristalização da proteína Xylellaina.....	109
Figura II-4.3: Gel em poliacrilamida a 15% mostrando a expressão e purificação da Xylellaina selenometionina substituída em cromatografia de afinidade	110
Figura II-4.4: Ensaio de cristalização da proteína Xylellaina selenometionina substituída.....	111
Figura II-4.5: Padrão de difração dos cristais nativos da Xylellaina.....	113
Figura II-4.6: Padrão de difração dos cristais Xylellaina selenometionina derivada.	114
Figura II-4.7: Gráfico mostrando ocupância dos sítios de selênio.	116
Figura II-4.8: Diagrama de Ramachandram da estrutura refinada da Xylellaina.....	118
Figura II-4.9: Representação na forma de cartoons da estrutura refinada da Xylellaina.....	119
Figura II-4.10: Detalhe da superfície de contato entre monômeros A e B da Xylellaina.....	121
Figura II-4.11: Representação em cartoons da estrutura cristalográfica da Xylellaina.....	122
Figura II-4.12: Representação em cartoons da estrutura da Xylellaina mostrando em esferas e sticks os aminoácidos do sítio catalítico.....	123
Figura II-4.13: Representação em “cartoons” da estrutura da Xylellaina mostrando porção N-terminal encobrindo o sítio catalítico.....	124
Figura II-4.14: Sobreposição das estruturas de pró-catepsinas com a Xylellaina.....	126
Figura II-4.15: Representação da superfície da Xylellaina, mostrando interação da pró-região com a proteína.....	127
Figura II-4.16: Região de ligação ao ribonucleotídeo UDP.....	128
Figura II-4.17: Superfície da Xylellaina com sobreposição da pró-região da pró-catepsina X e da Xylellaina.....	130
Figura II-4.18: Representação de linhas dos carbonos α da Xylellaina e a sobreposição dos “loops” de oclusão das estruturas cristalográficas da catepsina B e K.....	132

Apêndice III - Colaborações realizadas durante o período do doutorado

Figura AIII-1.1: Representação em “cartoons” do monômero AegOBP1.....	140
Figura AIII-1.2: Representação em “cartoons” do dímero da AegOBP1.....	140
Figura AIII-2.1: Representação em “cartoons” da estrutura cristalográfica da glioxalase II de <i>S. tephymurium</i>	145
Figura AIII-2.2: Detalhe do sítio de ligação aos metais e da segunda esfera de coordenação.....	145

Lista de Tabelas

Parte I - Caracterização molecular e estrutural de duas enzimas triptofanil tRNA sintetase de *Leishmania major*.

Capítulo 3

Tabela I-3.2: Reagentes utilizados na reação de adição de deoxiadenosina (A)	19
Tabela I-3.3: Reagentes usados nas reações de PCR para amplificação e sub-clonagens dos genes <i>LmTrpRS1</i> e <i>LmTrpRS2</i>	21
Tabela I-3.4: Reagentes usados na reação de digestão de DNA com endonucleases de restrição.	22
Tabela I-3.5: Reagentes usados na reação de ligação dos genes nos vetores de expressão	23

Apêndice I

Tabela AI-1.1: Nome e seqüência dos oligonucleotídeos utilizados na reação de RT-PCR para caracterização das regiões não traduzidas dos genes <i>LmTrpRS1</i> e <i>LmTrpRS2</i> . Os números na frente dos oligonucleotídeos representam a posição de ligação aos genes de interesse (vide figura abaixo).....	77
Tabela AI-1.2: Detalhamento do oligonucleotídeos utilizados na amplificação dos genes <i>LmTrpRS1</i> e <i>LmTrpRS2</i> de <i>L. major</i> , pelo método de PCR	78
Tabela AI-1.3: Detalhamento do oligonucleotídeos utilizados nas diversas construções dos vetores de expressão com os genes <i>LmTrpRS1</i> e <i>LmTrpRS2</i> de <i>L. major</i>	79

Parte II - Estudos estruturais da Xylellaína, uma cisteíno protease da bactéria *Xylella fastidiosa*.

Capítulo 4

Tabela II-4.1: Estatísticas de coleta e processamento de dados do cristal de Xylellaína nativo.....	113
Tabela II-4.2: Estatísticas de coleta e processamento de dados do cristal de Xylellaína selenometionina derivado.	115
Tabela II-4.3: Estatísticas do refinamento da estrutura da Xylellaína.	118

Apêndice III

Tabela AIII-1.1: Estatísticas de coleta dados, processamento e refinamento.....	139
Tabela AIII-2.1: Estatísticas de coleta dados, processamento e refinamento.....	144

Resumo

As aminoacil tRNA sintetases (AaRSs) são enzimas essenciais na síntese de proteínas assegurando a correta relação entre os aminoácidos e seus tRNA cognatos. O genoma mitocondrial dos tripanossomatídeos perdeu os genes codificantes dos tRNAs, assim os tRNA mitocondriais são codificados no núcleo e importados do citoplasma. O código genético do kinetoplasto desvia do código genético pela utilização do códon de terminação UGA para a decodificação do códon do triptofano. Um único gene codificando o tRNA^{Trp}(CCA) observado no genoma de *Leishmania* é responsável pela incorporação do aminoácido triptofano durante a síntese proteica na mitocôndria. Para decodificar os dois códons do Trp (UGA e UGG) a base na posição 34 do tRNA^{Trp}(CCA) passa por um evento de editoração, convertendo o ribonucleotídeo C34 em U34, produzindo o tRNA^{Trp}(UCA) capaz de decodificar o códon UGA. Nesse trabalho foram caracterizadas duas triptofanil tRNA sintetases de *Leishmania major*. De acordo com experimentos de “western blotting” e análises “in silico” das seqüências de aminoácidos, uma enzima tem localização citoplasmática (LmTrpRS1) enquanto a outra mitocondrial (LmTrpRS2). Os mRNAs dos dois genes foram definidos por experimentos de 5’ e 3’ RT-PCR. As duas enzimas foram clonadas em diversos vetores de expressão procariotos e eucariotos. A LmTrpRS1 foi obtida somente na fração insolúvel, já a LmTrpRS2 foi obtida na fração solúvel quando clonada no vetor de expressão pET28a. Esta porém mostrou-se instável precipitando rapidamente após sua purificação. Os ensaios enzimáticos realizados com a mesma mostraram que ela é capaz de reconhecer os tRNAs^{Trp} editado e não editado. Modelagem molecular por homologia com as duas proteínas foi realizada usando a proteína citoplasmática humana como molde, para estudar a interação entre a proteína e o tRNA^{Trp}.

Xylella fastidiosa é um bactéria gram negativa limitada ao xilema, responsável por um grande número de doenças economicamente importantes, como a doença de Pierce em videiras, Clorose variegada do Citrus (CVC) e a doença da requeima das folhas em outras plantas incluindo, amendoeira, ameixeira, louro, amoreira e café. Em todos os casos a *X. fastidiosa* afeta o xilema da planta causando redução na produção de frutos. Nesse trabalho nós mostramos a estrutura da Xylellaína, uma cisteína protease desse patógeno. A estrutura foi resolvida por dispersão anômala a um único comprimento de onda, utilizando cristais de xylellaína selenometionina substituídos.

A estrutura da Xylellaína foi refinada até 1,65 Å de resolução, mostrando enovelamento similar às proteínas da família da papaína, porém algumas características interessantes como uma região N-terminal composta por 38 aminoácidos cobrindo o sulco ativo da enzima, um intrigante ribonucleotídeo encontrado fora do sítio ativo da enzima e um “loop” semelhante ao “loop” de oclusão presente na catepsina B.

Abstract

The aminoacyl tRNA synthetases (aaRSs) are essential enzymes in protein synthesis that ensure the correct match between amino acids and their cognate tRNAs. The mitochondrial (kinetoplast) genome of trypanosomatids lacks tRNA genes, and therefore nucleus-encoded tRNAs are imported from the cytoplasm, the kinetoplast genetic code deviates from the universal code in that UGA instead of UGG encodes for tryptophan. A single nucleus-encoded tRNA^{Trp}(CCA) is responsible for Trp insertion during organellar protein synthesis. To decode both Trp codons (UGA and UGG), tRNA^{Trp}(CCA) undergoes a single C to U editing event at position 34 of the anticodon yielding to versions of the tRNA in the mitochondria with anticodon CCA and UCA, permitting UGA decoding. This work has characterized two *Leishmania major* tryptophanyl-tRNA synthetase, according to western blotting experiments and “in silico” sequence analysis one of cytoplasmic localization (LmTrpRS1) and another from mitochondria localization (LmTrpRS2). The mature mRNA transcripts for both genes were defined by 5' and 3' RT-PCR. Both enzymes were cloned into several expression vectors. LmTrpRS1 was obtained as an insoluble protein and LmTrpRS2 expressed into the soluble fraction in pET28a expression system. LmTrpRS2 protein, however, is unstable precipitating shortly after purification. The enzymatic assay showed that this enzyme is able to recognize both tRNA^{Trp}. Molecular modeling for LmTrpRS1 and LmTrpRS2 were constructed using the cytoplasmic human tryptophanyl tRNA synthetase as a model, to study the interaction between proteins and tRNA^{Trp}.

Xylella fastidiosa is a xylem-limited, gram-negative bacteria responsible for a large number of economically important plant diseases, such as Pierce's disease in grapevines, citrus variegated chlorosis (CVC) in sweet oranges and leaf scorch diseases in other plants, including almond, plum, oleander, mulberry and coffee. In all cases, *X. fastidiosa* infects the plant xylem and impairs fruit production. Here, we report the crystal structure of xylellain, a cysteine protease from *X. fastidiosa*. The structure was solved by single-wavelength anomalous dispersion (SAD) using seleno-methionine containing xylellain crystals.

The final structure of Xylellain was refined against the best native data set (1.65 Å) showing R/Rfree= 17/21. Xylellain shares fold similar to Papain like Family, but contains some interesting features, like a 38 N-terminal tail covering the active site cleft; one intriguing ribonucleotide found outside the active site and one loop that resembles the occluding loop from cathepsin B.

Parte I

**Caracterização molecular e estrutural de duas enzimas
triptofanil tRNA sintetases de *Leishmania major*.**

Capítulo 1

Introdução

Nesse capítulo é feita uma descrição geral sobre as Leishmanioses; epidemiologia da doença, ciclo de vida do parasita e combate a doença. Posteriormente é realizada uma introdução sobre sistema de editoração de tRNAs mitocondriais, para contextualizar a importância da editoração $C_{34} \rightarrow U_{34}$ do anticódon tRNA^{Trp}(CCA) na decodificação do códon de terminação UGA em Kinetoplastidas, e as enzimas tRNA sintetases, em especial a triptofanil tRNA sintetase interesse particular dessa tese.

Leishmaniose.

Leishmanioses são um conjunto de doenças causadas por pelo menos 24 espécies de protozoários do gênero *Leishmania*, transmitidos por aproximadamente 30 espécies de insetos vetores pertencentes aos gêneros *Phlebotomus* e *Lutzomyia* (Herwaldt, 1999; Desjeux, 2001; Desjeux, 2004), ficando apenas atrás da malária em número de mortes causadas por doenças parasitárias. Mesmo com diferentes formas de expressão nos tecidos animais infectados, três características comuns podem ser encontradas nas infecções causadas por esses protozoários: 1) parasitismo de macrófagos do hospedeiro, 2) resposta imuno-inflamatória do hospedeiro regulando o desenvolvimento da doença e 3) infecção persistente do tecido infectado (figura I-1.1) (Murray, 2005).

A maioria das espécies de *Leishmania* causa sérias doenças crônicas, manifestando-se por lesões cutâneas ou viscerais (figura I-1.2). A cada ano, de 1,5 a 2 milhões de crianças e adultos desenvolvem doença sintomática, sendo o número de mortos estimado em 70.000 por ano (Desjeux, 2004). Aproximadamente 90% das infecções cutâneas ocorrem no Afeganistão, Paquistão, Síria, Arábia Saudita, Argélia, Irã, Brasil e Peru, enquanto que as infecções viscerais ocorrem na sua grande maioria na Índia, Sudão, Nepal, Bangladesh e Brasil (Murray, 2005). Devido

à distribuição geográfica e ao baixo potencial econômico associado a essa doença os estudos com as leishmanioses permaneceram por muito tempo negligenciados pelos países ditos desenvolvidos (Yamey, 2002). Entretanto, o aparecimento da doença em veteranos da guerra do golfo (Maggil *et al.*, 1993), pacientes imunodeprimidos (Alvar *et al.*, 1997; Davidson, 1997; Desjeux, 1998), o estudo da leishmaniose como sistema modelo de resposta celular imune contra patógenos intracelulares (Reed e Scott, 1993; Berman, 1997), e ainda a liberação dos genomas de *Leishmania major* (El-Sayed *et al.*, 2005; Ives *et al.*, 2005) e do inseto vetor *Lutzomyia longipalpis* (Dillon *et al.*, 2005) aumentaram o interesse sobre os estudos com leishmaniose nos últimos anos, principalmente no que se refere ao diagnóstico, controle do inseto vetor e na busca de terapias eficazes contra a doença (Murray, 2005).

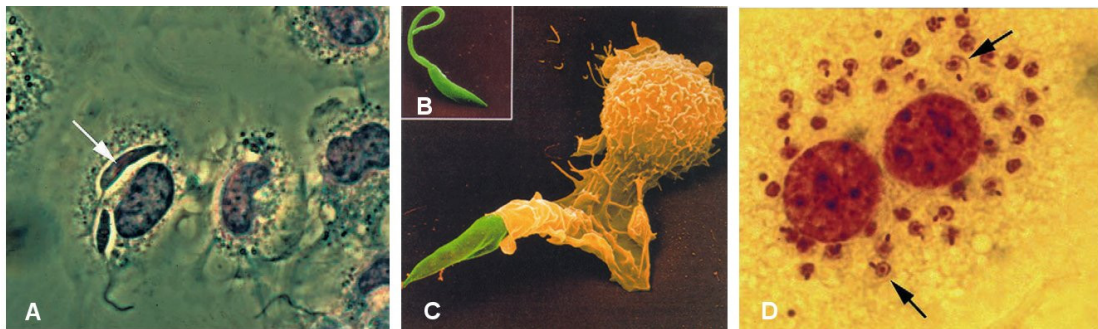


Figura I-1.1: Microscopia ótica e de varredura de macrófagos e parasitas do gênero *Leishmania*. **A)** Ingestão de duas formas promastigotas de *Leishmania donovani* por um macrófago humano. **B e C)** Microscopia eletrônica de varredura colorida artificialmente mostrando detalhe de *L. major* (verde), e uma célula de *L. major* sendo fagocitada por um macrófago (amarelo). **D)** Amostra de tecido peritoneal de camundongo, mostrando replicação intracelular de formas amastigotas (Figura adaptada de Murray, 2005 e www.icp.ucl.ac.be).

A leishmaniose cutânea (LC) manifesta-se por lesões ulcerativas da derme com feridas comumente na face, braços e pernas (figura I-1.2 A e B). A disseminação dos parasitas via corrente sanguínea ou linfática pode resultar em ulcerações nos tecidos cartilagosos (leishmaniose muco cutânea - LMC), principalmente boca e nariz, produzindo lesões mutilantes (figura I-1.2 C). Dentro do continente americano essas duas manifestações são agrupadas em uma única doença denominada leishmaniose tegumentar americana (LTA), também conhecida como úlcera de Bauru, ferida brava, entre outras denominações regionais (Herwaldt, 1999; Davies *et al.*, 2003).

A forma visceral da doença (leishmaniose visceral - LV), também conhecida como Calazar (figura I-1.2 D), acomete o sistema retículo-endotelial, causando perda

de peso, anemia e inchaço do baço e fígado podendo levar a morte dentro de 2 a 3 anos. (Herwaldt, 1999; Davies *et al.*, 2003; Murray, 2005).



Figura I-1.2: Algumas manifestações clínicas das leishmanioses. A e B) Lesão ulcerativa causada por leishmaniose cutânea (Irã, 2002). C) Aspecto clínico de lesão avançada da mucosa nasal (Etiópia, 2003). D) Menino sofrendo de leishmaniose visceral, apresenta esplenomegalia, abdômen distendido e atrofia muscular (Sudão, 1997). Fotos: WHO/TDR.

O ciclo de vida do parasita no hospedeiro vertebrado (figura I-1.3) inicia-se com a liberação das formas flageladas promastigotas dos parasitas no sangue do hospedeiro vertebrado através da picada de flebotomíneos fêmeas dos gêneros *Phlebotomus* (velho mundo) ou *Lutzomyia* (novo mundo). Dentro do hospedeiro vertebrado as formas promastigotas ligam-se aos receptores específicos CR3 e CR1 dos macrófagos sendo então fagocitadas, formando uma estrutura de vacúolos parasitóforos ou fagossomos (Handman e Bullen, 2002). Dentro dos fagossomos ocorre a mudança da forma promastigota para amastigota. Evidências apontam que a mudança de pH esta associada com a conversão de uma forma em outra (Antoine *et al.*, 1998; Burchmore e Barrett, 2001). O macrófago infectado deixa de funcionar como célula apresentadora de antígeno e passa a hospedar o parasita, fornecendo proteção contra o sistema imune. As formas amastigotas do parasita multiplicam-se por divisão binária dentro dos macrófagos, destruindo a célula hospedeira. Essas formas são então liberadas para o sangue ou linfa aonde infectam novos macrófagos. O ciclo no hospedeiro invertebrado começa com a ingestão de macrófagos contaminados com as formas amastigotas, essas transformam-se em formas

promastigotas e multiplicam-se no intestino do inseto, migrando posteriormente para as probóscides (aparelho picador-sugador dos insetos).

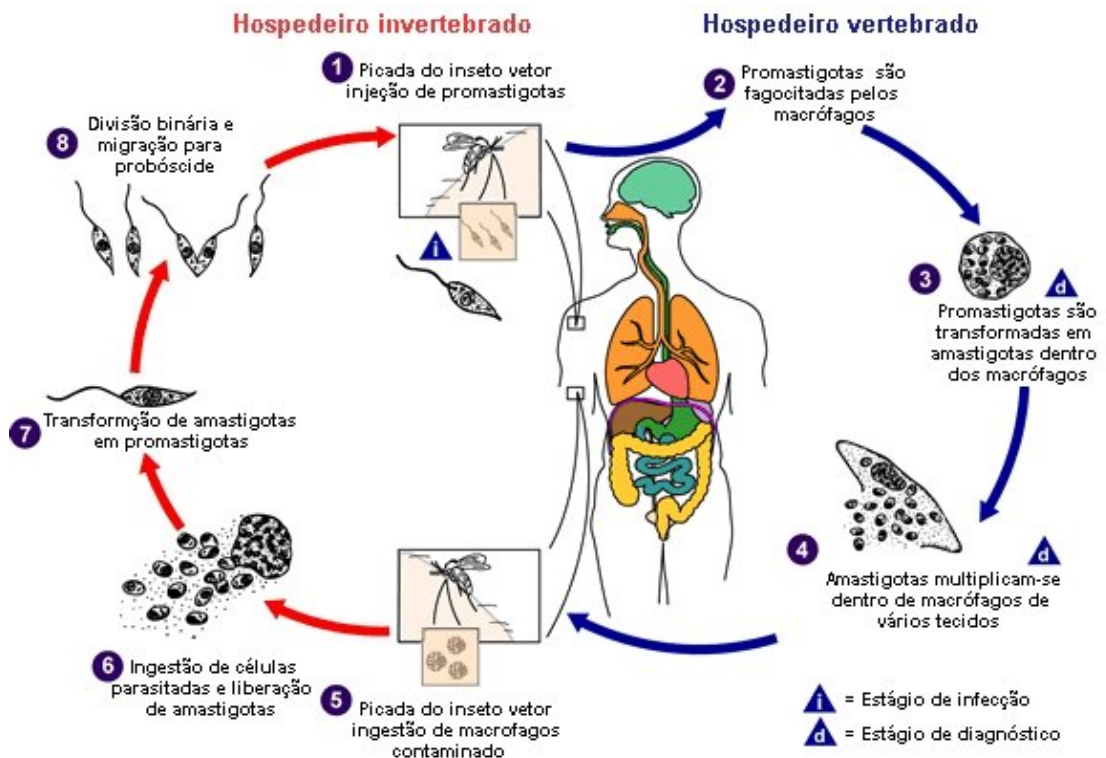


Figura I-1.3: Ciclo de vida dos parasitas do gênero *Leishmania*. (Figura adaptada de www.dpd.cdc.gov).

O diagnóstico parasitológico é a metodologia mais simples e barata, baseia-se na observação das formas amastigotas utilizando-se esfregaços de medula óssea ou aspirados de linfonodos corados pelo Giemsa. A dificuldade na obtenção dos tecidos faz com que o diagnóstico imunológico seja o mais indicado. Entretanto, variação entre as amostras, alto custo dos antígenos comercialmente disponíveis e a persistência de anticorpos após a cura são fatores limitantes a essa técnica. O método mais sensível e específico é a amplificação e detecção do DNA de *Leishmania* (PCR), este possibilita a identificação dos parasitas a partir de grande variedade de amostras (Davies *et al.*, 2003).

O tratamento da leishmaniose tem empregado desde a década de 40 os derivados de antimônios pentavalentes como medicamentos de primeira escolha (Croft e Yardley, 2002; Croft e Coombs, 2003; Croft *et al.*, 2005). Atualmente duas formulações de antimoniais pentavalentes são disponíveis no mercado: o estibogluconato de sódio (Pentostam - GlaxoSmithKline) e o antimoniato-N-metil glucamina (Glucantime - Aventis), sendo o último distribuído gratuitamente no

Brasil pelo Ministério da Saúde. Altas doses diárias, durante longos períodos (20 a 40 dias), são necessárias em virtude da curta meia-vida do composto na corrente sanguínea, apenas 2 horas. Nos últimos anos, doses progressivamente maiores dos antimoniais têm sido recomendadas devido ao aparecimento de resistência do parasita (Croft, 2001). O mecanismo de ação desses fármacos ainda não está totalmente elucidado, mas sabe-se que ele atua nas formas amastigotas do parasita inibindo sua atividade glicolítica e a via oxidativa de ácidos graxos. Diversas reações de toxicidade como a trombose e a hipersensibilidade são verificadas, podendo chegar até ao choque anafilático (Geary *et al.*, 1989). A Pentamidina (Pentacarinat - Rhodia Farma) tem sido utilizada como segunda escolha contra leishmaniose desde 1952, pois apresenta maior toxicidade ao paciente, períodos mais longos de tratamento e sérios efeitos colaterais, tais como: hipoglicemia, diabetes, taquicardia, entre outros (Croft e Coombs, 2003). Dentro da atual geração de medicamentos leishmanicidas, a Anfotericina B (Fungisone - Bristol Meyers Squibb) é o mais eficaz, atuando nas formas promastigotas e amastigotas do parasita através da ligação preferencial aos ésteres (ergosterol ou episterol) presentes na membrana plasmática do parasita. A administração do medicamento é feita lentamente via intravenosa por períodos de até 4 horas. Diversos efeitos colaterais da anfotericina B são observados, principalmente o sistema Renal, todos com estrita relação com a dose terapêutica administrada, acometendo (Croft e Coombs, 2003; Croft *et al.*, 2005). Novas formulações da Anfotericina B associadas à lipossomas, AmBisome (NeXstar Pharmaceuticals) e a dispersão coloidal Amphocil (AstraZeneca), tornaram-se disponíveis comercialmente apresentando uma menor toxicidade e maior meia vida plasmática, permanecendo mais tempo no organismo do paciente. Entretanto, o custo elevado impossibilita o uso destes novos compostos na rotina do serviço de saúde pública (Ministério da Saúde 2003; Croft e Coombs, 2003; Croft *et al.*, 2005).

No que se refere à prevenção da leishmaniose, os métodos mais comuns são: a desinsetização das áreas de risco para a erradicação dos insetos vetores, o uso de mosquiteiros impregnados de inseticidas e a aplicação de repelentes sobre a pele. Sacrifício de cães infectados também faz parte do controle da doença, uma vez que esses são reservatórios da doença. Além disso, surge com grande esperança o desenvolvimento de vacinas (Pinheiro, 2004), utilizando proteínas recombinantes, parasitas vivos atenuados e ainda as vacinas de DNA. Atualmente existe uma vacina,

denominada Leishvacin (Biobrás), produzida somente para a realização de ensaios clínicos. Estima-se que a produção em escala comercial da Leishvacin se iniciará após o ano 2020 (Toledo *et al.*, 2001).

Editoração dos RNA transportadores mitocondriais (kinetoplastos).

A importação de tRNAs codificados no núcleo ocorre em diferentes microorganismos eucariotos (protozoários, fungos, algas) e em algumas plantas e animais pertencentes aos filos Cnidária e Mollusca (Yasuhira e Simpson, 1997; Kapushoc *et al.*, 2000; Simpson *et al.*, 2000; Schneider, 2001).

Os tripanosomatídeos *Leishmania tarentolae* e *Trypanosoma brucei* não possuem no genoma mitocondrial genes codificando para tRNAs, de forma que todos os tRNAs para a tradução mitocondrial são importados para a organela do citoplasma (Yasuhira e Simpson, 1997). Em nenhum caso estudado até o presente momento o mecanismo de importação de tRNAs citoplasmáticos pela mitocôndria está plenamente compreendido. As evidências indicam um mecanismo diferente em cada organismo estudado (Lima e Simpson, 1996; Simpson *et al.*, 2000; Schneider, 2001).

Vários tipos de editoração de tRNAs já foram descritos (Janke e Paabo, 1993; Antes *et al.*, 1998; Price e Gray, 1998). Editoração por substituição e/ou modificação de nucleotídeos de tRNAs mitocondriais, que restauram o pareamento de bases em sítios conservados no braço acceptor ocorre em espécies de amebas do gênero *Acanthamoeba*, alguns fungos da família dos critidiomicetos e alguns animais. Em marsupiais, uma única modificação de um nucleotídeo na segunda posição do anticódon restaura a capacidade decodificadora do tRNA^{Asp} mitocondrial. Em *Physarum polycephalum*, vários tRNAs mitocondriais apresentam editoração por inserções de citosinas e uridinas (Hatfield, *et al.*, 1982). Além desses exemplos, a deaminação específica de adenosina para inosina ocorre em tRNAs de *Saccharomyces cerevisiae* (Auxilien *et al.*, 1996).

O genoma mitocondrial de protozoários tripanosomatídeos consiste de 20 a 50 moléculas de maxicírculos. Esses maxicírculos codificam dois RNAs ribossomais e 18 genes codificantes para proteínas. Aproximadamente 12 desses genes, dependendo da espécie, representam criptogenes; genes cujos transcritos são editorados em diferentes extensões pelo mecanismo de adição e deleção de uridinas. Em adição aos maxicírculos, o genoma mitocondrial contém entre 10.000 e 20.000

moléculas de minicírculos (Simpson e Thiemann, 1995; Estevez e Simpson, 1999). Estes codificam RNAs guia (guide RNAs - gRNAs) envolvidos na mediação do processo de editoração. Os maxicírculos e minicírculos se organizam em uma rede concatenada compacta denominada de DNA do kinetoplasto (kDNA). Nenhum gene de tRNA codificado tanto nos maxicírculos como nos minicírculos foi identificado, apesar da seqüência desse genoma ter sido completamente determinada. As evidências tanto de experimentos *in vivo* como *in vitro* indicam que os tRNAs necessários ao processo de tradução são importados do citoplasma (figura I-1.4) (Simpson e Thiemann, 1995; Lima e Simpson, 1996, Yasuhira e Simpson, 1997, Estevez e Simpson, 1999).

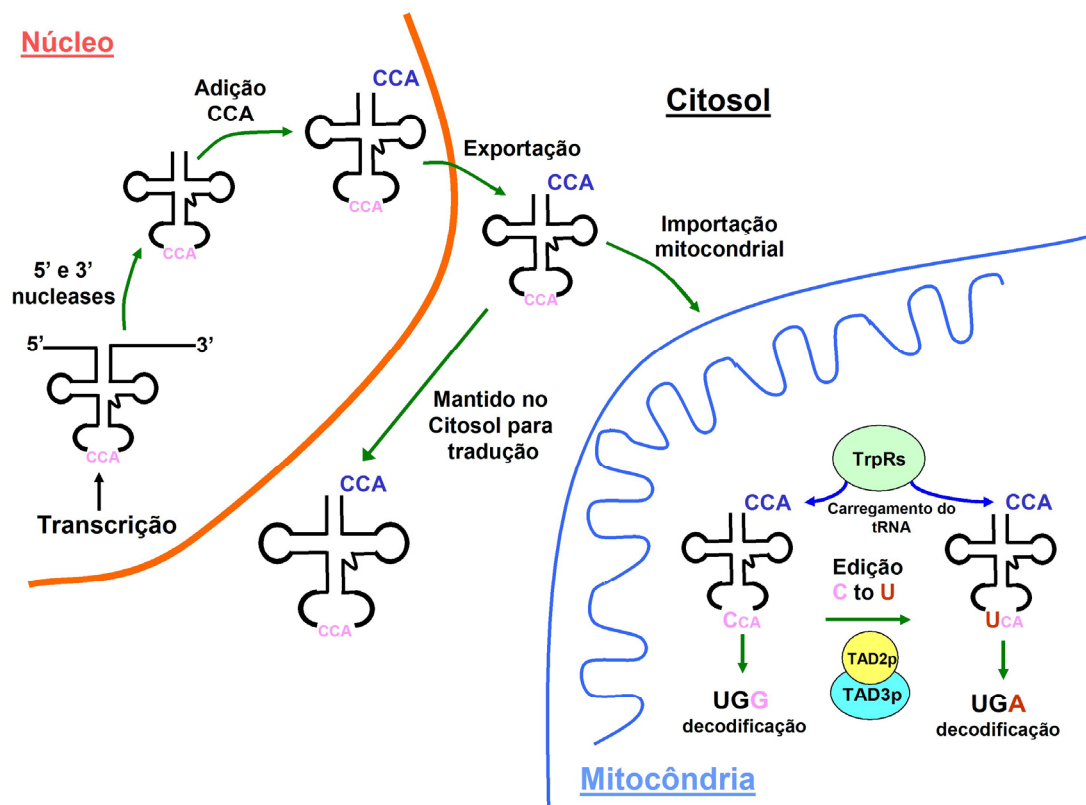


Figura I-1.4: Esquema do mecanismo de importação, exportação e editoração de tRNAs mitocondriais em Kinetoplastidas. (Figura cedida pelo prof. Dr. Juan D. Alfonso – Ohio State University)

Como no caso da maioria dos eucariotos o código genético mitocondrial (Kinetoplasto) de *Leishmania* se desvia do código universal onde o códon UGA de terminação é usado como um códon para triptofano (Estevez e Simpson, 1999; Simpson e Thiemann, 1995). Entretanto, em *L. tarentolae* um único tRNA^{Trp}

codificado no núcleo foi identificado, possuindo o anticódon CCA. Esse tRNA^{Trp} pode decodificar o códon canônico UGG para triptofano, mas não o códon UGA do kinetoplasto. Foi identificado recentemente um novo mecanismo de editoração em kinetoplasto de *L. tarentolae* pela modificação de uma citosina para uridina na posição 1 do anticódon do tRNA^{Trp} (CCA) (Alfonzo *et al.*, 1996). Até o presente momento duas enzimas, Tad2p/Tad3p, parecem estar envolvidas nesse mecanismo (Rubio, *et al.*, 2007). Este evento de editoração leva a formação do tRNA^{Trp} (UCA) que decodifica o códon mitocondrial UGA para triptofano (figura I-1.4). Em outros organismos, decodificação do códon UGA mitocondrial pelo tRNA^{Trp} (CCA) é decorrência de mutações fora do anticódon que permitem o pareamento do tipo “não Watson-Crick” (pareamento C-A) entre a primeira posição do anticódon e a terceira posição do códon ou pela modificação específica da citosina (para lisidina, por exemplo) que permite pareamento C-A.

No caso do tRNA^{Trp} em *Leishmania*, a editoração do ribonuclotídeo C₃₄ para U₃₄ no anticódon é a única modificação de seqüência observada. A presença de um tRNA^{Trp}(UCA) mitocondrial que não possui outra modificação capaz de alterar sua estrutura canônica, levou a conclusão que a editoração no tRNA^{Trp} é essencial para o apropriado funcionamento da tradução mitocondrial (Alfonzo *et al.*, 1996).

Aminoacil-tRNA sintetases

As aminoacil-tRNA sintetases (aaRSs) são uma família de enzimas que asseguram a correta ligação entre um aminoácido e seu tRNA correspondente gerando um conjunto de tRNAs aminoacilados essenciais para o processo de síntese protéica (Ibba e Soll, 2000; Lenhinger *et al.* 2000).

As aaRSs realizam a reação de aminoacilação em duas etapas: Primeiro ATP reage com o aminoácido formando aminoacil-adenilato e difosfato. Em uma segunda reação o aminoácido ativado é transferido para o tRNA para formar o aminoacil-tRNA e AMP (figura I-1.5). As aaRSs são classificadas em dois grupos (Classe I e II) segundo a reação de aminoacilação (figura I-1.5) e a suas características estruturais (figura I-1.6) (Ibba e Soll, 2000; Lenhinger *et al.* 2000).

Nestas reações o aminoácido é ligado ao tRNA por uma ligação de alta energia e assim é denominado “ativado”.

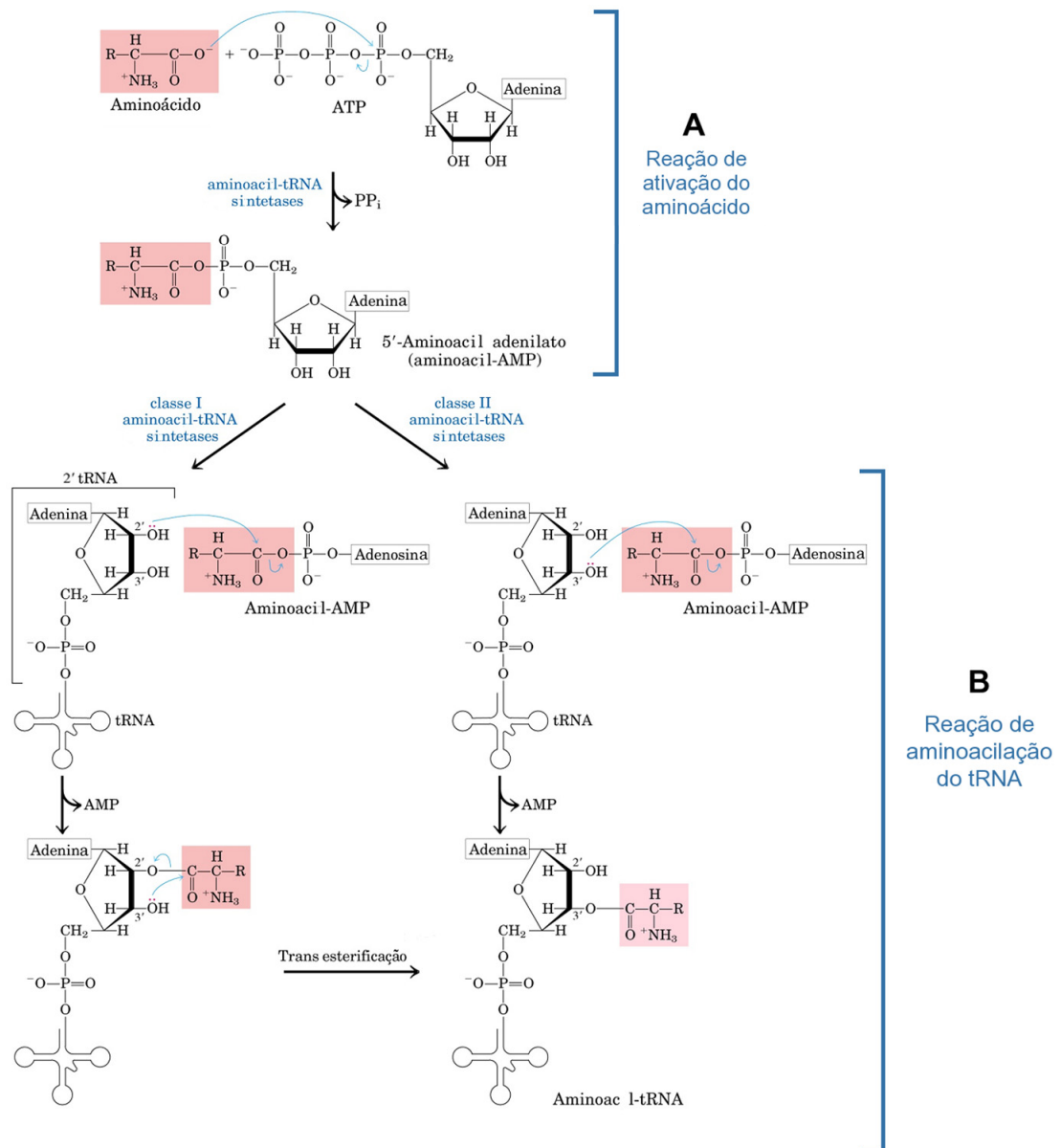


Figura I-1.5: Mecanismo de reação das AARSs. A reação de aminoacilação do tRNA é realizada em duas etapas: **A)** Reação de ativação do aminoácido. **B)** Reação de aminoacilação. Nas AARSs de classe I o aminoácido é transferido inicialmente para a hidroxila 2' e posteriormente, por uma reação de trans-esterificação passa a hidroxila 3'. Já nas AARSs de classe II a aminoacilação ocorre diretamente na hidroxila 3'. (Figura adaptada de Lehninger Principles of Biochemistry 3rd Ed. 2000)

Cada aaRSs possui cavidades precisas para o ATP e o seu aminoácido específico. A discriminação do correto tRNA é realizada em grande parte pela presença de elementos determinantes de identidade na estrutura dos tRNA, presentes principalmente nas região do “braço do aminoácido do tRNA” e no anticódon. Intensa interação ocorre entre o tRNAs e as aaRSs (figura I-1.6) de maneira a

posicionar os elementos de identidade próximos aos sítios de reconhecimento. Estruturas cristalográficas de aaRSs em complexo com tRNAs mostram que o reconhecimento é diferenciado nas classes I e II, o que pode explicar em parte a diferença de interação do aminoácido com o grupo hidroxila da adenina do braço aceptor do tRNA (Fig. I-1.5) (Ibba e Soll, 2000; Lenhinger *et al.* 2000). Outras funções têm sido atribuídas às aaRSs como: fidelidade celular, processamento de tRNA, splicing de tRNAs, transporte de tRNAs, apoptose e regulação transcricional e traducional. (Martinis *et al.*, 1999)

A utilização das aaRSs como alvo para inibidores seletivos de agentes infecciosos já está bem documentada indicando o potencial futuro deste projeto no caminho da busca de inibidores específicos contra leishmanioses (Tao e Shimmel, 2000; Schimmel e Tao, 1998; Brown *et al.*, 2002; Kim *et al.*, 2003; Ochsner *et al.*, 2007).

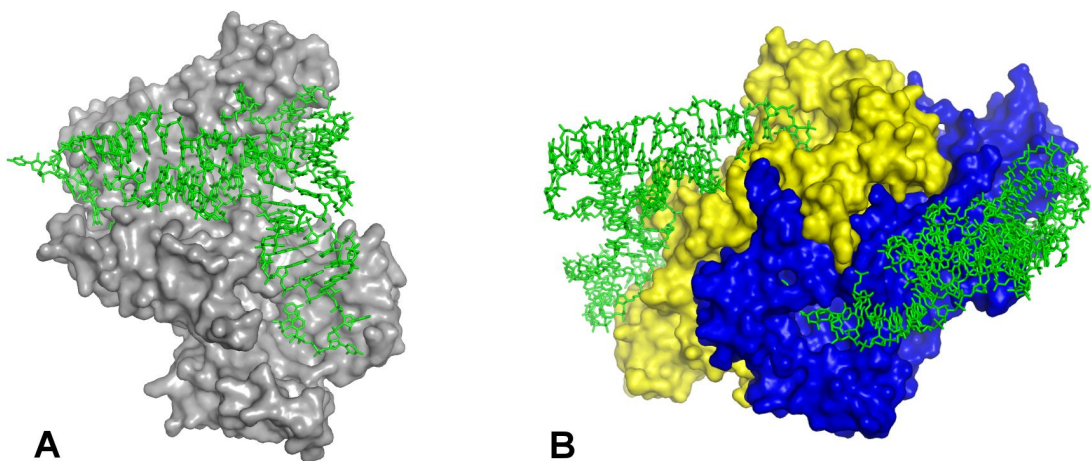


Figura I-1.6: Representação de superfície de aminoacil-tRNA sintetases ligadas aos tRNA cognatos. A) ArgRS de *Saccharomyces cerevisiae* (código PDB 1F7U), enzima monomérica da classe I. **B)** PheRS de *Thermus thermophilus* (código PDB 1E1Y) enzima dimérica de classe II.

Triptofanil-tRNA sintetase.

O reconhecimento da maioria das enzimas endereçadas à importação mitocondrial é mediado por seqüências amino-terminais de 20 a 60 resíduos de aminoácidos. As aminoacil-tRNA sintetases mitocondriais são codificadas pelo genoma nuclear e transportadas para a matriz mitocondrial (Neupert, 1997; Ibba e

Soll, 2000; Woese *et al.*, 2000). Essas seqüências são clivadas após a correta localização da enzima (Neupert, 1997).

As triptofanil-tRNA sintetases (TrpRS) pertencem à classe I das aaRSs na forma de dímeros α_2 e, ao contrário das outras aaRSs, não formam complexos de alto massa molecular, sendo as menores aaRSs identificadas. Os membros dessa classe apresentam duas seqüências consenso características, a primeira composta pelo motivo “ Φ H Φ GH” (Φ denota aminoácido hidrofóbico) e a segunda pelo motivo “KMSKS”, esse dois formam o sítio de ligação do ATP e do aminoácido triptofano. Estruturas cristalográficas recentes apontam claramente a evolução destas seqüências consenso para a discriminação entre os aminoácidos de cadeia aromática (Yang *et al.*, 2003), em especial a estrutura de um fragmento catalítico da TrpRS humana complexado com substrato não hidrolizável mostra o papel desses aminoácidos no reconhecimento e ativação do aminoácido triptofano (Yu *et al.*, 2004). Outra característica comum a essa classe de aaRSs é o domínio de ligação a nucleotídeos conhecido como “Rossman Fold” (Branden e Tooze, 1999). As seqüências consenso são trazidas próximas uma da outra por uma pronunciada curvatura de quatro fitas β paralelas, formando o sitio de ligação do nucleotídeo de adenina (Ibba e Soll, 2000; Woese *et al.*, 2000).

Dados estruturais e bioquímicos mostram que além do mecanismo de reconhecimento e carregamento do tRNA as TrpRS estão envolvidas em outros processos biológicos como: atividade angiostática (TrpRS citoplasmática humana) (Kise *et al.*, 2004), e na formação de complexos com a enzima óxido nitro sintetase em resposta ao dano do DNA (Buddha *et al.*, 2004; Buddha *et al.*, 2005a; Buddha e Crane, 2005).

Até o presente momento 264 estruturas tridimensionais de aaRSs encontram-se depositadas no “Protein Data Bank” (<http://www.rcsb.org/pdb/>). Dentre essas, 11 pertencem as TrpRS de bactérias. Recentemente duas estruturas de triptofanil tRNA sintetase citoplasmática humana complexadas com tRNA^{Trp} foram depositadas sob as coordenadas de acesso, 2AZX (Yang *et al.*, 2006), 2AKE e 2DR2 (Shen *et al.*, 2006), mostrando algumas características interessantes com relação ao modo de ligação ao tRNA.

O mecanismo de reconhecimento específico dos tRNAs^{Trp} editorado e não editorado pelas TrpRSs é ainda desconhecido. Recente trabalho com TrpRS de

Trypanosoma brucei (Charrière *et al.*, 2005) mostra que a discriminação dos dois tRNA^{Trp} é realizada pela enzima citoplasmática através de elementos presentes nas proximidades do anticódon do tRNA^{Trp}. Estudos estruturais devem ainda ser realizados e são de extrema importância para elucidar completamente esse mecanismo discriminatório.

Triptofanil-tRNA sintetases em *L. major*.

Dois genes codificantes para Triptofanil-tRNA sintetase encontram-se depositados no GeneDB, sob os números de acesso LmjF29.0060 e LmjF23.0300. Inicialmente a análise seqüencial e a predição de regiões sinalizadoras apontavam que ambas as enzimas eram destinadas à exportação mitocondrial, após caracterização das regiões não traduzidas foi observado que a janela de leitura depositada do gene LmjF29.0060 estava errada. A nova janela de leitura não mostrou seqüência de sinalização mitocondrial, caracterizando essa proteína como citoplasmática. Seguindo a notação amplamente utilizada para proteínas que possuem formas distintas para mitocôndrial e citoplasmática a forma citoplasmática foi então denominada LmTrpRS1 (Lm de *L. major*, TrpRS de enzima tRNA sintetase que reconhece o aminoácido triptofano e o número 1 denota enzima citoplasmática) e a forma mitocondrial foi denominada LmTrpRS2 (número 2 denota enzima mitocondrial).

Referências do capítulo

- Alfonzo, J. D.; Blanc, V.; Estevez, A. M.; Rubio, M. A. T. e Simpson, L. (1999). *EMBO J.*, **18**: 7056-7062.
- Alvar. J.; Canavate, C.; Gutierrez-Solar, B.; Jimenez, M.; Laguna. F.; Lopez-Velez, R.; Molina, R, e Moreno, J. (1997). *Clin. Microbiol. Rev.*, **10**: 298–319.
- Antes, T.; Costandy, H.; Mahendran, R.; Spottswood, M. e Miller, D. (1998). *Mol. Cell. Biol.*, **18**: 7521–7527.
- Antoine, J.; Prina, E.; Lang, T. e Courret, N. (1998). *Trends Microb.*, **7**: 392-400.

- Auxilien, S.; Crain, P. F.; Trewyn, R. W. e Grosjean, H. (1996). *J. Mol. Biol.*, **262**: 437-458.
- Berman, J. D. (1997). *Clin. Infect. Dis.*, **24**: 684–703.
- Branden, C. e Tooze, J. (1999) In *Introduction to Protein Structure* 2^{ed}.
- Brown, M. J.; Carter, P. S.; Fenwick, A. S.; Fosberry, A. P.; Hamprecht, D. W.; Hibbs, M. J.; Jarvest, R. L.; Mensah, L.; Milner, P. H.; O’Hanlon, P. J.; Pope, A. J.; Richardson, C. M.; West, A. e Witty, D. R. (2002). *Bioorg. Med. Chem. Lett.*, **12**: 3171-3174.
- Buddha, M. R.; Kerry, K. M. e Crane, B. R. (2004). *Proc. Natl. Acad. Sci. U. S. A.*, **101**: 15881-15886.
- Buddha, M. R. e Crane, B. R. (2005a). *Nat. Struct. Mol. Biol.*, **12**: 274-275.
- Buddha, M. R. e Crane, B. R. (2005b). *J. Biol. Chem.*, **280**: 31965-31973.
- Burchmore, R. J. S. e Barrett, M. P. (2001). *Int. J. Parasitol.*, **31**: 1311–1320.
- Charriere, F.; Helgadottir, S.; Horn, E. K.; Soll, D. e Schneider, A. (2006). *Proc. Natl. Acad. Sci. U.S. A.*, **103**: 6847-6852.
- Croft, S. L. (2001). *Tropical, Medical and International Health*, **6**: 899-905.
- Croft, S. L. e Yardley, V. (2002). *Curr. Pharm. Des.*, **8**: 319-342.
- Croft, S. L. e Coombs, G. H. (2003). *Trends Parasitol.*, **19**: 502-508.
- Croft, S. L.; Barrett, M. P. e Urbina, J. (2005). *Trends Parasitol.*, **21**: 508-512.
- Davidson, R. N. (1997). *Genitourin Med.*; **73**: 237–239.
- Davies, C.R.; Kaye, P.; Croft, S.L. e Sundar, S. (2003). *BMJ*, **326**: 377-382.
- Desjeux P e UNAIDS. (1998) in *Gridlock*. Geneva: WHO and the UN AIDS WHO/CTD/LEISH/98.9 Add. 1 and UNAIDS/98.23.
- Desjeux, P. (2001). *Trans. R. Soc. Trop. Med. Hyg.*, **95**: 239–243.
- Desjeux, P. (2004). *Comp. Immunol. Microbiol. Infect. Dis.*, **27**: 305–218.
- Dillon, R. J.; Ivens, A. C.; Soares, M. B.; Lehane, M. J.; Bates PA. (2005), *Arch. Instit. Pasteur Tunis*; **82**: 39.
- El-Sayed, N. M.; Myler, P. J.; Blandin, G.; *et al.* (2005). *Science*, **309**: 404–09.
- Estevez, A. M. e Simpson, L. (1999). *Gene* **240**: 247-260.

- Geary, T. G.; Edgar, A. e Jensen, J. B. (1989) In Campbell, W.C. e Rew. R.S. (eds.), *Chemoterapy of Parasitic Diseases*. Plenum Press, New York, pp. 209-238.
- Hatfield, D.; Rice, M.; Hession, C. A. e Melera, P. W. (1982), *J. Bacteriol.*, **151**: 1013-1021.
- Handman, E. e Bullen, D. (2002), *Trends Parasitol.*, **18**: 332-334.
- Herwaldt, B.L. (1999), *Lancet*, **354**: 1191–1199.
- Ibba, M. e Soll, D. (2000), *Annu. Rev. Biochem.*, **69**: 617-650.
- Ilyin, V. A.; Temple, B.; Hu, M.; Li, G.; Yin, Y.; Vachette, P. e Carter C.W. Jr. (2000), *Protein Sci.* **9**: 218-231.
- Ivens, A. C.; Peacock, C. S.; Worthey, E. A.; *et al.* (2005) *Science*, **309**: 436–442.
- Janke, A. e Paabo, S., (1993), *Nucleic Acids Res*, **21**: 1523–1525.
- Kapushoc, S. T.; Alfonzo, J. D.; Rubio, M. A. e Simpson, L. (2000), *J. Biol. Chem.*, **275**: 37907-37914
- Kim, S.; Lee, S. W.; Choi, E. C e Choi, S. Y. (2003), *Appl. Microbiol. Biotechnol.*, **61**: 278–288
- Kise, Y.; Lee, S. W.; Park, S. G.; Fukai, S.; Sengoku, T.; Ishii, R.; Yokoyama, S.; Kim, S. e Nureki, O., (2004), *Nat. Struct. Mol. Biol.*, **11**:149-156
- Lenhinger, A. L.; Nelson, D.N. e Michael M.C. (2000) In *Lehninger Principles of Biochemistry 3rd.*
- Lima, B. D.e Simpson, L. (1996), *RNA*, **2**: 429-440.
- Magill, A. J; Grogl, M.; Gasser, R. A.; Sun, W. e Oster, C. N. (1993). *N. Engl. J. Med.*, **328**: 1383–1387.
- Martinis, S. A.; Plateau, P.; Cavarelli, J.; Florentz, C. (1999) *EMBO J.* **18**: 4591-4596.
- Murray, H. W.; Berman, J. D.; Davies, C. R. e Saravia, N. G. (2005). *Lancet*, **366**: 1561-1577.
- Neupert, W. (1997). *Annu. Rev. Biochem.* **66**: 863-917.
- Ochsner, U. A.; Sun, X.; Jarvis, T.; Critchley, I. e Janjic, N. (2007), *Expert. Opin. Investig. Drugs.*, **16**: 573-593.
- Pinheiro, R. O. (2004). *Infarma*, **16**: 79-82.

- Price, D. H. e Gray, M. W. (1998). In Grosjean, H. e Benne, R. (eds), *Modification e Editing of RNA*. ASM Press, Washington, DC, pp. 289–306.
- Reed, S. G. e Scott, P. (1993). *Curr Opin Immunol*; **5**: 524–31.
- Rubio, M. A.; Pastar, I.; Gaston, K. W.; Ragone, F. L.; Janzen, C. J.; Cross, G. A.; Papavasiliou, F. N. e Alfonzo, J. D. (2007), *Proc. Natl. Acad. Sci. USA*, **104**: 7821-7826.
- Schimmel, P. e Tao, J. (2000). *Exp. Opin. Invest. Drugs*, **9**: 1767-1775.
- Schneider A. (2001). *Int. J. Parasit.*, **31**: 1403-1415.
- Schumacher, M.A.; Carter, D.; Ross, D.S.; Ullman, B. e Brennan, R.G. (1996). *Nat. Struct. Biol.*, **3**: 881-887.
- Shen, N.; Guo, L.; Yang, B.; Jin, Y. e Ding. (2006). *J. Nucleic Acids Res.* **34**: 3246-3258
- Simpson, L. e Thiemann, O. H. (1995). *Cell*, **81**: 837-840.
- Simpson, L.; Thiemann, O. H.; Savill, N. J.; Alfonzo, J. D. e Maslov, D. A. (2000). *P. Natl. Acad. Sci. U. S. A.* **97**: 6986-6993.
- Tao, J.S. e Schimmel, P. (2000). *FASEB J.*, **9**: 1767-1775.
- Toledo, V. P. C. P.; Mayrink, W.; Gollob, K. J.; Oliveira, M. A. P.; Costa, C. A.; Genaro, O.; Pinto, J. A. e Afonso, L. C. C. (2001). *Mem. Inst. Oswaldo Cruz*, **96**: 89-98.
- Woese, C. R.; Olsen, G. J.; Ibba, M. e Soll, D. (2000). *Microbiol. Mol. Biol. R.*, **64**: 202-235.
- Yamey, G. (2002). *BMJ*, **325**: 176–77.
- Yang, X.L.; Otero, F.J.; Ewalt, K.L.; Liu, J.; Swairjo, M.A.; Kohrer, C.; Rajbhandary, U.L.; Skene, R.J.; McRee, D.E e Schimmel, P. (2006). *EMBO J.* **25**(12): 2919-29
- Yang, X.; Otero, F. J.; Skene, R. J.; Mcree, D. E.; Schimmel, P. e Pouplana, L. R. (2003), *P. Natl. Acad. Sci. U. S. A.* **100**: 15376–15380
- Yasuhira, S. e Simpson, L. (1997). *J. Mol. Evol.*, **44**: 341-347.
- Yu, Y.; Liu, Y.; Shen, N.; Xu, X.; Xu, F.; Jia, J.; Jin, Y.; Arnold, E. e Ding, J. (2004). *J. Biol. Chem.*, **279**: 8378-8388.

Capítulo 2

Objetivos

O objetivo principal desse trabalho foi a caracterização bioquímica e estrutural de duas enzimas triptofanil tRNA sintetases do parasita *L. major*.

Objetivos específicos:

- Caracterização das seqüências nucleotídicas dos genes *LmTrpRS1* e *LmTrpRS2*.
- Clonar os genes codificantes das proteínas LmTrpRS1 e LmTrpRS2 em vetores de expressão apropriados para produção de proteína em larga escala.
- Desenvolver protocolos de expressão e purificação das proteínas recombinantes obtidas.
- Localização celular das enzimas no parasita.
- Caracterização estrutural das proteínas

Capítulo 3

Metodologia

Nesse capítulo é feita uma descrição geral sobre os procedimentos realizados desde a caracterização das regiões não traduzidas até as comparações estruturais das enzimas em estudo com outras TrpRSs. Resumidamente a caracterização molecular dos genes envolveu clonagem dos genes em estudo, determinação das regiões não traduzidas (UTRs) e do número de cópias no genoma de *L. major*. As proteínas em estudo foram sub-clonadas em vetores de expressão específicos. A proteína LmTrpRS2 obtida na fração solúvel foi purificada em cromatografia de afinidade, sendo então realizados os ensaios enzimáticos. Modelos estruturais das duas enzimas foram construídos a partir de modelagem molecular por homologia e comparados com estruturas de outras TrpRSs, em especial a proteína humana complexada com tRNA.

Caracterização molecular dos genes *LmTrpRS1* e *LmTrpRS2*

Caracterização das regiões 5' e 3' não traduzidas (UTR) dos genes *LmTrpRS 1* e *LmTrpRS2*

Entender os mecanismos básicos de crescimento, diferenciação e resposta celular aos estímulos ambientais, ou seja, compreender a organização espacial e temporal dos eventos moleculares vem sendo um desafio para a biologia molecular. Na verdade, embora muitos dos elementos regulatórios situem-se em regiões não codificadoras do genoma, os dados disponíveis de seqüências nucleotídicas são referentes, na sua grande maioria, às seqüências expressas correspondentes às porções gênicas codificadoras de proteínas. Dentre as porções não codificadoras, as regiões 5' e 3' não traduzidas dos mRNAs, 5'UTR e 3'UTR, têm se mostrado

interessantes por conter elementos de regulação da expressão gênica (Pesole *et al.*, 2002).

A caracterização das regiões não traduzidas foi realizada através da técnica de RT-PCR (Sambrook e Russel, 2000) adaptada para a caracterização de mRNA de *Leishmania* (Thiemann *et al.*, 1994). Esta consiste na síntese de um cDNA a partir de mRNA (transcrição reversa - RT) seguida de amplificação do cDNA específico por reação em cadeia de polimerase (PCR). Oligonucleotídeos específicos foram desenhados para as regiões que flanqueiam as extremidades dos genes de interesse (figura I-3.1 e tabela AI-1.1 no apêndice I, pág. 77). Para amplificação da região 5'UTR foi utilizado um oligonucleotídeo iniciador complementar a região conservada do “Spliced Leader” (SL) (Boothroyd e Cross, 1982; Maden e Hughes, 1997), já para região 3'UTR dos genes, foram empregados o oligonucleotídeos complementares a região poli-A (figura I-3.1 e tabela AI-1.1 no apêndice I).



Figura I-3.1: Estratégia adotada para a caracterização das regiões não traduzidas dos genes *LmTrpRs1* e *LmTrpRs2*. As setas indicam os oligonucleotídeos utilizados nas reações, cujas seqüências encontram-se descritos na tabela AI-1.1 do apêndice I (pág. 77).

A extração do RNA total foi realizada a partir de cultura celular de formas promastigotas de *L. major* crescidas em meio de cultura M199 a 26° C (Kapler *et al.*, 1990), utilizando-se o reagente “Trizol[®]” (Invitrogen), segundo orientações do fabricante. As reações de RT-PCR foram realizadas utilizando-se o conjunto “OneStep RT-PCR Kit[®]” (Qiagen), segundo condições descritas na tabela I-3.1 e o ciclo de RT-PCR mostrado na figura I-3.2, em termociclador PTC-100[™] Programmable Thermal Controller (MJ Research Inc.). A fim de restringir a seleção dos genes, os produtos das reações de RT-PCR obtidos foram submetidos a novos ciclos de PCR utilizando-se oligonucleotídeos mais externos aos genes (figura I-3.1 e tabela AI-1.1 no apêndice I). Os fragmentos de DNA obtidos nesta reação de PCR foram separados por eletroforese em gel de agarose, purificados com conjunto “Wizard[®] SV Gel and PCR Clean-Up System” (Promega). Esses foram então submetidos à adição de dATP (deoxi-adenosina), segundo a reação descrita na tabela I-3.2, por 30 minutos a 70° C e ligados ao vetor de clonagem pGEM-T[®] (Promega), conforme especificações do fabricante.

Células competentes de *Escherichia coli* DH5 α foram transformadas, por choque térmico (Sambrook e Russel, 2000), com os plasmídeos provenientes das reações de ligação. As colônias cresceram em placas contendo ampicilina (100 $\mu\text{g}/\text{mL}$), IPTG (0,5 mM) e X-Gal (80 $\mu\text{g}/\text{mL}$). A confirmação dos plasmídeos recombinantes foi realizada pela técnica de PCR de colônia utilizando-se os primers de seqüenciamento de vetor pGEM-T[®] (T73G e SP6). Os plasmídeos recombinantes foram extraídos pelo protocolo de lise alcalina (Sambrook e Russel, 2000) e submetidos ao seqüenciamento de DNA, pelo método de terminação de cadeia por di-deoxinucleotídeos (Sanger *et al.* 1977). As seqüências nucleotídicas obtidas foram analisadas e submetidas à comparação com as seqüências depositadas no banco de dados NCBI, através do programa “BLAST Sequences” (BLASTn).

Tabela I-3.1: Reagentes usados na reação de RT-PCR para amplificação das regiões 3’UTR e 5’UTR

Mistura da reação de RT-PCR	Volume (μL)	Concentração final
Tampão RT-PCR [5X]	10,0	1X
dNTPs 10 mM (A, G, C, T)	2,0	0,4 mM cada
Oligonucleotídeos 100 pmol/uL	0,5 cada	1 μM cada
RNA molde		1,0 μg
Enzima MIX (Reverse Transcriptase+ Hot Start Taq DNA Polymerase)	2,0	
RNAse out (40U/ μL)	0,25	10U
H ₂ O milli-Q autoclavada q.s.p.	50,0	

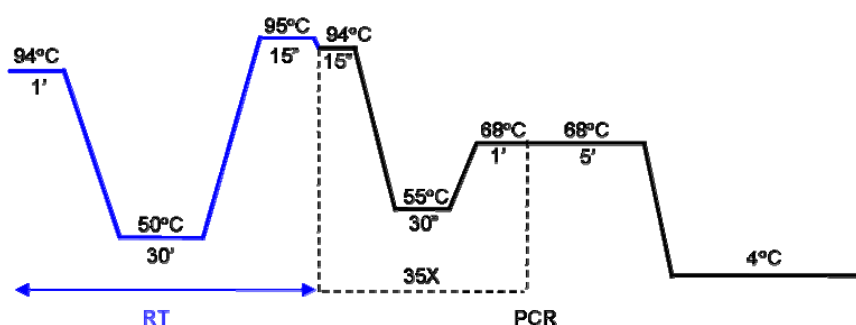


Figura I-3.2: Ciclo da reação de RT-PCR utilizado para amplificação das regiões 3’UTR e 5’UTR do gene *LmTrpRS1* e *LmTrpRS2*. Em azul a reação de transcrição reversa, em preto o ciclo de PCR.

Tabela I-3.2: Reagentes utilizados na reação de adição de deoxiadenosina (A)

Reagentes para a reação de adição de deoxiadenosina.	Volume (μL)
DNA purificado	7,3
<i>Taq</i> DNA polimerase recombinante (Invitrogen) (1U)	1,0
dATP 5 mM	0,4
Tampão PCR <i>Taq</i> DNA polimerase recombinante [10x] (Invitrogen)	1,0
MgCl ₂ 50 mM	0,3

Caracterização do número de cópias dos genes *LmTrpRS1* e *LmTrpRS2*

Com o objetivo de caracterizar os genes *LmTrpRS1* e *LmTrpRS2* em relação ao número de cópias presentes no organismo do parasita *L. major*, foi utilizada a análise com endonucleases de restrição seguida de “southern blotting” (Southern, 1975).

Amostras de aproximadamente 1,0 µg de DNA genômico de *L. major*, extraído de culturas celulares pelo protocolo de fenol/clorofórmio (Sambrook e Russel, 2000), foram submetidos à digestão com as endonucleases de restrição: *NdeI*, *NotI*, *XbaI*, *XhoI*, *BglI*, *Bam HI*, *Hind III*, *PstI* e *Eco RI* (New England Biolabs). Essas amostras foram separadas de acordo com o tamanho, por eletroforese em gel de agarose 0,8%, seguida de desnaturação do DNA e transferência por capilaridade para uma membrana de nitrocelulose Hybond™ N⁺ (GE healthcare) (Sambrook e Russel, 2000). O DNA foi fixado na membrana através de exposição à radiação UV “*Cross-linking*” em Stratalinker® UV Crosslinker 1800 (stratagene).

Uma sonda radioativa foi utilizada para os experimentos de hibridização, esta sonda consistiu no produto de amplificação dos genes de interesse através da técnica de “Random Priming”, utilizando-se α-dAT³²P (Feinberg e Vogelstein, 1983) e o conjunto “Megaprime™ DNA Labeling System” (GE Healthcare). Os procedimentos de hibridização e lavagens foram realizados de acordo com protocolo padrão (Sambrook e Russel, 2000) e manual de instruções da membrana.

A visualização das hibridizações foi realizada com exposição da membrana em filme radiográfico (Kodak Biomax®), que foi incubado em cassete adequado a -80° C por 2 horas para intensificação da exposição. Os filmes foram revelados com solução reveladora e fixador (Kodak Biomax® developer).

Clonagem dos genes *LmTrpRS1* e *LmTrpRS2*

A amplificação dos genes *LmTrpRS1* e *LmTrpRS2* de *L. major* foi realizada pelo método de PCR, utilizando como molde DNA total de *L. major* e oligonucleotídeos específicos às extremidades dos genes de interesse (tabela AI-1.2 no apêndice I). Os oligonucleotídeos foram construídos considerando-se: 1) As seqüências codificantes das proteínas *LmTrpRS1* e *LmTrpRS2* depositadas no banco

de dados do NCBI (<http://www.ncbi.nlm.nih.gov>); 2) A predição das regiões destinadas à exportação celular destas enzimas, realizada pelo programa MitoProt II (Claros e Vincens, 1996); 3) Os sítios de restrição adequados à inserção dos genes nos vetores de expressão utilizados. Duas formas da enzima mitocondrial (LmTrpRS2) foram clonadas, uma com a seqüência N-terminal de sinalização mitocondrial e outra com a remoção dessa região predita pelo programa MitoProt II.

As reações de PCR foram realizadas segundo condições descritas na tabela I-3.3 e programa de PCR mostrado na figura I-3.3 (oligonucleotídeos - tabela AI-1.2 no apêndice I), com pequenas adaptações entre os ciclos de PCR para adaptá-lo às diversas construções realizadas. Durante todas as amplificações foi utilizada a enzima DNA polimerase (Invitrogen, Promega ou Eppendorf) com atividade corretiva para evitar a inserção de mutações.

Tabela I-3.3: Reagentes usados nas reações de PCR para amplificação e sub-clonagens dos genes *LmTrpRS1* e *LmTrpRS2*.

Mistura da reação de amplificação	Volume (µL)	Concentração final
Tampão PCR [10X] sem Mg	5,0	1X
dNTPs 10 mM (A, G, C, T)	1,0	0,2 mM cada
MgSO ₄ 50 mM	1,0	2,0 mM
Oligonucleotídeo <i>sense</i> 100 pmol/ul	1,0	2 µM
Oligonucleotídeo <i>antisense</i> 100 pmol/ul	1,0	2 µM
DNA molde	1,0	100 ng
Platinum <i>Taq</i> DNA polimerase (5 U/µL)	2,0	1,0 U
H ₂ O milli-Q autoclavada	38,0	

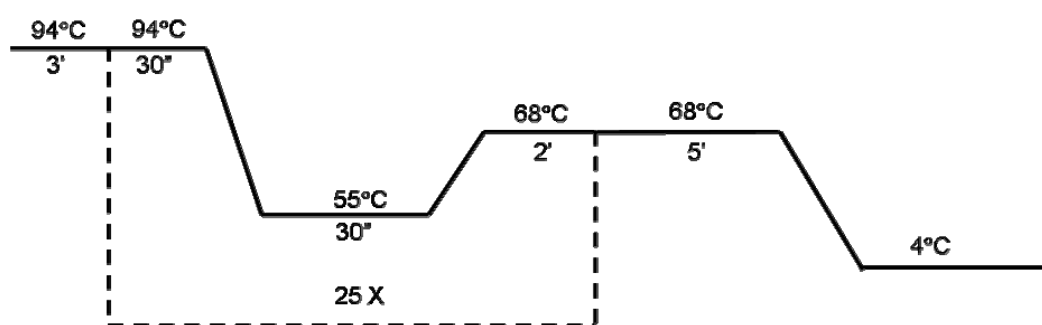


Figura I-3.3: Esquema do programa de PCR utilizado para amplificação e sub-clonagens dos genes *LmTrpRS1* e *LmTrpRS2*.

Os genes obtidos nesta reação foram inseridos no vetor de clonagem pGEM-T[®] (Promega). Os produtos de ligação foram inseridos em bactérias *E. coli* competentes, e a presença do inserto foi confirmada por extração plasmidial das

colônias recombinantes selecionadas, seguida de análise de digestão com endonucleases de restrição compatíveis com os oligonucleotídeos construídos, ou então reação de PCR com oligonucleotídeos específicos aos vetores utilizados (tabela AI-1.2 no apêndice I). A reação geral de digestão (tabela I-3.4), foi realizada por 4 horas, seguindo condições de reação sugeridas pelo fabricante.

Tabela I-3.4: Reagentes usados na reação de digestão de DNA com endonucleases de restrição.

Reagentes para a reação de digestão para cada vetor	Volume (µl)
DNA - 300 ng/µl	20
Enzima 1 (5 U/µl) - compatível com extremidade sense do oligonucleotídeo	1,5
Enzima 2 (5 U/µl) - compatível com extremidade anti-sense do oligonucleotídeo	1,5
Tampão compatível com enzimas utilizadas	3
Água milli-Q	4

As seqüências nucleotídicas dos genes de interesse foram confirmadas pelo seqüenciamento do DNA dos plasmídeos recombinantes, seguido de análise com as seqüências depositadas no banco de dados NCBI, através do programa BLASTn.

Sub-clonagem dos genes *LmTrpRS1* e *LmTrpRS2* em vetores de expressão e ensaios de expressão protéica

Diversas construções envolvendo os genes de interesse, tanto em vetores de expressão procarióticos quanto eucarióticos foram realizadas para contornar os problemas de solubilidade das proteínas em estudo. Para tanto, oligonucleotídeos específicos para essas construções foram construídos, levando-se em consideração os sítios de restrição adequados aos vetores empregados e a correta janela de leitura (tabela AI-1.3 no apêndice I). As reações e os ciclos de PCR foram realizadas como mostrado na tabela I-3.3 e figura I-3.3 respectivamente, utilizando os genes *LmTrpRS1* e *LmTrpRS2* clonados no vetor pGEM-T, com pequenas alterações para observar as temperaturas de anelamento e extensão adequadas a cada oligonucleotídeo. Os fragmentos com tamanhos esperados aos genes *LmTrpRS1* e *LmTrpRS2* foram separados em gel de agarose 1% e purificados como descrito anteriormente.

Para a inserção dos genes de interesse nos plasmídeos específicos, os produtos de PCR purificado e oDNA plasmidial (vetores de expressão) foram

digeridos por 4 horas em condições de reação sugeridas pelo fabricante das endonucleases de restrição, reação geral sumarizada na tabela I-3.4. Após digestão, as amostras foram separadas em gel de agarose 1% e purificadas. Os produtos de digestão purificados foram inseridos nos plasmídeos de interesse através de reação de ligação utilizando a enzima T4 DNA ligase (Sambrook e Russel, 2000), segundo reação mostrada na tabela I-3.5. Esta foi realizada a 4° C por 12 horas.

Os produtos de ligação foram inseridos em bactérias *E. coli* competentes. As bactérias recombinantes foram selecionadas em meio de cultura sólido com antibiótico adequado. As colônias recombinantes tiveram o DNA plasmidial extraído, e a conferência da incorporação do inserto foi realizada por análise de restrição ou PCR. Após seqüenciamento, os clones sem mutações foram inseridos em cepas de *E. coli* competentes para o início dos ensaios de expressão.

Tabela I-3.5: Reagentes usados na reação de ligação dos genes nos vetores de expressão.

Reagentes para a reação de ligação para cada vetor	Volume (µL)
Produto de PCR digerido com as enzimas de restrição (100 ng/µL)	6
Plasmídeo digerido com as enzimas de restrição (100 ng/µL)	2
Buffer T4 DNA ligase (Promega)	1
T4 DNA ligase (Promega) (1 U/µl)	1

Ensaio de expressão protéica foram realizados em meios de cultura bacterianos (LB, 2XYT, meio mínimo) (Sambrook e Russel, 2000), variando-se as condições de indução a fim de aperfeiçoar a expressão solúvel da proteína. Após lise celular realizada por ultra-som, as frações solúveis e insolúveis foram separadas por centrifugação e analisadas em eletroforese desnaturante em gel de poliacrilamida a 15% (Laemmli, 1970).

Produção de anticorpos policlonais contra as proteínas LmTrpRS1 e LmTrpRS2

A produção de anticorpos policlonais contra as proteínas LmTrpRS1 e LmTrpRS2 foi realizada através da inoculação das proteínas de interesse em camundongos (Frederick *et al.*, 1997; Sambrook e Russel, 2000), em colaboração com laboratório da Prof^a. Dra. Heloísa Selistre Araújo - UFSCar. As proteínas de interesse, expressas na forma de corpos de inclusão no vetor pET29a, foram separadas por eletroforese desnaturante em gel de poliacrilamida e as bandas

correspondentes às proteínas de interesse foram recortadas do gel de poliacrilamida e maceradas em tampão “PBS” (Sambrook e Russel, 2000). Após dosagem protéica (Bradford, 1976) 300 µl de proteína a 100 µg/mL foram aplicadas em cada camundongo, processo repetido após 45 dias da injeção inicial. Dez dias após a segunda injeção os animais foram sacrificados e os soros contendo os anticorpos foram removidos. A produção de anticorpos foi conferida pela técnica de “western blotting” (Burnette, 1981; Sambrook e Russel, 2000).

Localização celular das proteínas LmTrpRS1 e LmTrpRS2.

A localização celular foi realizada através do fracionamento do conteúdo celular de células de *L. tarentolae*, seguido de “western botting” com os anticorpos policlonais produzidos. O fracionamento celular foi realizado em colaboração com Prof. Dr. Juan D. Alfonzo - Ohio State University. Para o fracionamento, culturas de células foram crescidas em meio M199 até fase logarítmica, 10^6 células/mL, essas foram ressuspensas em tampão DTE (1 mM Tris-HCl, 1 mM EDTA, pH 7.9) gelado até a concentração de 5×10^8 células/mL. Após 10 minutos, as células são passadas através de uma agulha 0,26mm a 110lbs/in², aonde são lisadas e o kinetoplasto permanece íntegro. Esse lisado foi transferido para uma solução de sacarose a 60%, centrifugado por 15 minutos a 9000 rpm. A fração precipitada foi ressuspensa em tampão STM (20 mM Tris-HCl, pH 7.9, 2 mM EDTA ou 2 mM MgCl₂) e DNase, na proporção de 1 mL pra cada 6 mL de lisado. Este foi incubado 1 hora em gelo. Após esse tempo adicionou-se o mesmo volume de tampão STE (250 mM sacarose, 10 mM g Tris, 2 mM EDTA, pH 7,9) e a amostra foi centrifugada a 9000 rotações por minuto (rpm) por 15 minutos. A fração precipitada foi utilizada para os gradientes de Percoll (Sigma-Aldrich). A amostra foi centrifugada a 24000 rpm por uma hora. A banda contendo a fração mitocondrial foi então centrifugada em excesso de STE por 15 minutos a 11500 rpm e foi lavada duas vezes. (Simpson *et al.*, 1993; Simpson *et al.*, 1996).

Os extratos mitocondrial e citoplasmático obtidos após lise e fracionamento celular, foram separados por eletroforese em gel de poliacrilamida, sendo então prontamente transferido para uma membrana de nitrocelulose. Os anticorpos policlonais anti-LmTrpRSs foram incubados com a membrana por 2 horas. A membrana foi lavada com solução de TBS, e um anticorpo secundário anti-igG de

camundongo conjugado com peroxidase foi colocado em contato com a membrana. Após a incubação a membrana foi revelada com solução contendo um composto químico luminescente (Amersham ECL™ Advance Western Blotting Detection Kit - Ge Healthcare), essa membrana foi exposta a um filme fotográfico por 15 minutos (KODAK BioMax® - Kodak) e depois revelada. Os controles positivos da reação foram amostras das proteínas expressas em *E. coli*.

Expressão e Purificação da proteína LmTrpRS2.

A expressão da enzima LmTrpRS2 foi realizada em bactéria *E. coli* BL21 (DE3) contendo o plasmídeo recombinante *LmTrpRS2/pET28a* (proteína LmTrpRS2 sem sinalização mitocondrial, clonada no vetor pET28a- Novagem). A proteína foi expressa em meio LB líquido, contendo Kanamicina (30 µg/ml), estas culturas foram crescidas a 37 °C, sob agitação constante de 250 rpm, até as culturas atingirem uma densidade ótica a 600nm (D.O._{600 nm}) de aproximadamente 0,8. Neste instante a expressão da proteína recombinante foi induzida com IPTG (Isopropil β-galactopiranosídeo) na concentração de 10 µM, por 24 horas à temperatura de 20° C. Após indução, as células foram centrifugadas a 5.000 xg, por 15 minutos, a 4° C. O precipitado foi ressuspensionado em tampão de lise (PBS e 10 µg de lisozima por ml de células ressuspensionadas) durante 30 minutos a 4° C. As células foram lisadas através de 8 pulsos de ultra-som com duração de 15 segundos e intervalos de 30 segundos entre cada pulso. As amostras lisadas foram centrifugadas a 15.000 x g por 30 minutos, a 4° C, separando a fração solúvel da insolúvel.

A proteína foi purificada por cromatografia de afinidade utilizando-se a resina “Ni-NTA Superflow” (Qiagen). A proteína recombinante foi eluída da coluna com a aplicação de tampão PBS e gradiente de concentração de imidazol (três concentrações de imidazol, 50, 100 e 150 mM foram aplicadas). As frações eluídas da coluna de níquel foram analisadas em gel de poliacrilamida a 15% .

Ensaio de atividade enzimática da enzima LmTrpRS2 recombinante.

Transcrição *in vitro* dos tRNAs^{Trp} editado e não editado.

As reações de transcrição *in vitro* dos tRNA^{Trp}, foram realizadas segundo Amberg *et al.*, 1996. Os DNAs plasmidiais contendo os genes codificantes para os tRNA^{Trp} (UCA) e tRNA^{Trp}(CCA), inseridos no vetor de propagação pUC19 foram gentilmente cedidos pelo prof. Dr. Juan D. Alfonso (figura I-3.4). O plasmídio foi propagado em *E. coli* DH5α e após extração plasmidial o mesmo foi digerido com a enzima de restrição *Bst* NI a 60° C por 4 horas, para proporcionar a terminação CCA-3' encontrada nos tRNAs. A reação de digestão está sumarizada na tabela I-3.6. Após a digestão, o DNA plasmidial foi precipitado seguindo protocolo de acetato de amônio (Sambrook e Russel, 2000). O DNA plasmidial foi ressuspenso em água livre de RNase e utilizado como molde para a enzima T7 RNA polimerase. A reação de transcrição, sumarizada na tabela I-3.7, foi incubada a 37° C por 24 horas.

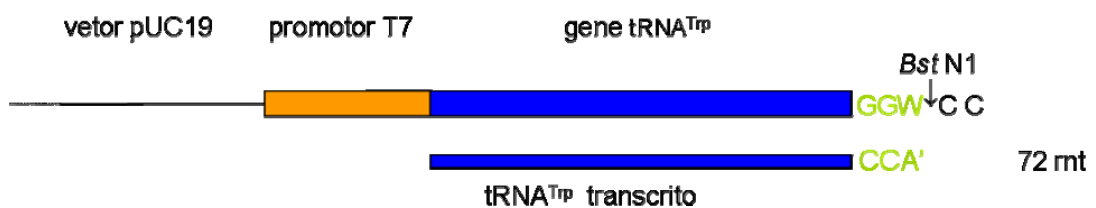


Figura I-3.4: Esquema representativo do cassete de transcrição do tRNA^{Trp}. Na parte superior um esquema do vetor pUC19 com a sequência do promotor para enzima T7 Rna polimerase (alaranjado), seguido pelo gene tRNA^{Trp} (azul) e sítio para enzima *Bst* NI na porção final do cassete (verde). Na parte inferior (azul) um esquema do tRNA^{Trp} transcrito *in vitro*.

Tabela I-3.6: Reagentes usados nas reações de digestão do DNA plasmidial tRNA^{Trp}.

<i>Reagentes para a reação de digestão</i>	<i>Volume (μL)</i>
DNA plasmidial (~250 ng/μL)	20
Enzima <i>Bst</i> NI (NEB) (10 U/μL)	2,5
NEbuffer 2	5
BSA 10 mg/mL	1
Água milli-Q	21,5

Tabela I-3.7: Reagentes usados nas reações de transcrição *in vitro* do tRNA^{Trp}.

<i>Reagentes para a reação de digestão</i>	<i>Volume (μL)</i>
DNA plasmidial (~200 ng/μL)	30
mix de rNTP (25 mM)	20
Inibidor de RNase (39,8 U/μL)	1
T7 RNA polimerase (100U/μL)	1
Tampão [10X] (tris-HCl 0,4 M; NaCl 0,05 M; MgCl ₂ 0,22 M; DTT 0,1 M; Espermidina 0,02 M)	10
H ₂ O Milli Q livre de RNase	38

Purificação da transcrição in vitro dos tRNAs^{Trp} e protocolo de re-novelamento.

Os tRNAs transcritos foram purificados utilizando-se gel desnaturante de poliacrilamida a 8% em 8 M de uréia (Curnow *et al.*, 1993). As amostras de tRNAs foram preparadas com tampão desnaturante (20 mM MOPS, pH 7,0, 1 mM EDTA, 6% de formaldeído, 31% formamida, 0,025% azul de bromofenol e 0,025% de xileno cianol), aquecidas a 70° C por 5 minutos e resfriadas em gelo por 3 minutos. O gel foi corado em solução de brometo de etídio a 1 µg/mL e as bandas com tamanho molecular semelhante aos tRNAs foram recortadas. Os tRNAs foram eluídos após incubação com 0,3 M de acetato de sódio durante 12 horas a temperatura ambiente e precipitados com etanol (Sambrook e Russel, 2000). Para o re-novelamento os tRNAs foram ressuspensos em água livre de RNase e desnaturados a 70° C por 3 minutos, e resfriados lentamente até temperatura ambiente.

Os ensaios de aminoacilação foram realizados a 37° C em um volume final de 50 µl, com 500 nM da enzima recombinante LmTrpRS2 em um meio de reação contendo: 50 mM Hepes; 10 mM acetato de magnésio; 4 mM DTT; 2 mM ATP; 0,05% de albumina bovina (BSA) e uma mistura de 2 µM de L- [5-³H]-triptofano (32 Ci/mmol - Amersham Life Science) e 38 µM de triptofano frio (Sever *et al.*, 1996; Charriere *et al.*, 2006). Aliquotas de 10 µl de cada reação foram aplicadas em papel de filtro Watman 3 MM (Watman), pré-equilibrado com 10% de ácido tricloroacético (TCA). Os papéis foram lavados duas vezes em 5% TCA por 10 minutos, seguido de lavagem com etanol por um minuto. Os filtros foram incubados em estufa à 80° C (para a remoção de água e etanol) e a radioatividade foi determinada em um contador de cintilação líquida (Beckman LS 6500) utilizando-se 2 mL de líquido de cintilação (0,02% POP e 0,6% POPOP em tolueno).

Caracterização estrutural das proteínas LmTrpRS1 e LmTrpRS2.

Modelagem molecular das proteínas LmTrpRS1 e LmTrpRS2.

A seleção das estruturas homólogas às proteínas LmTrpRS1 e LmTrpRS2 foram realizadas através de alinhamento seqüencial das proteínas de interesse com as

seqüências protéicas depositadas no “*Protein Data Bank*”, PDB, usando a ferramenta Blastp. As seqüências de aminoácidos das proteínas com maior similaridade foram analisadas uma a uma com o programa Clustal W (Chenna *et al.*, 2003). A modelagem molecular por homologia foi realizada utilizando-se o programa MODELLER 6a (Sali e Blundell, 1993), seguido da validação dos modelos com programas, PROCHECK (Laskowski *et al.*, 1998), WHATIF (Vriend, 1990) e VERIFY 3D (Luthy *et al.*, 1992).

Referências do capítulo

- Amberg, R.; Mizutani, T.; Wu, X. Q. e Gross, H. J. (1996). *J. Mol. Biol.*, **263**: 8-19.
- Boothroyd, J. C. e Cross, G. A. M. (1982). *Gene*, **20**: 281–289
- Bradford, M. M. (1976). *Anal. Biochem.*, **72**: 248-254.
- Burnette, W. N. (1981). *Anal. Biochem.*, **112**: 195-203.
- Charriere, F.; Helgadottir, S.; Horn, E. K.; Soll, D. e Schneider, A. (2006). *Proc. Natl. Acad. Sci. U. S. A.*, **103**: 6847-6852.
- Chenna, R.; Sugawara, H.; Koike, T.; Lopez, R.; Gibson, T. J.; Higgins, D.G. e Thompson, J.D. (2003). *Nucleic Acids Res.*, **31**: 3497-500.
- Claros, M. G. e Vincens, P. (1996). *Eur. J. Biochem.* **241**: 779-786
- Curnow, A.W.; Kung, F. L; Koch, K. A. e Garcia, G. A. (1993). *Biochemistry*, **32**: 5239-5234
- Feinberg, A. P. e Vogelstein, B. (1983). *Anal. Biochem.*, **132**: 6-13.
- Frederick, M. A. et al (1997). *Short Protocols in Molecular Biology*, 3th Edition - Edited by Frederick M. Ausubel *et al.* -Wiley John & Sons Inc.
- Ibba, M.; Hong, K. W.; Sherman, J. M.; Sever, S. e Soll D. (1996), *Proc Natl Acad Sci U. S. A.*, **93**: 6953-6958.
- Kapler, G. M.; Coburn, C. M. e Beverley, S. M. (1990). *Mol. Cell. Biol.*, **10**: 1084–1094.
- Laemmli, U. K. (1970). *Nature*, **227**: 680-685.

- Laskowski, R. A.; Macarthur, M. W. e Thornton, J. M. (1998). *Curr. Opin. Struct. Biol.*, **8**: 631-639.
- Luthy, R.; Bowie, J. U. e Eisenberg, D. (1992), *Nature*, **356**: 83-85.
- Maden, B. E. H. e Hughes, J. M. X. (1997), *Chromosoma* **105**: 391–400.
- Pesole, G.; Liuni, S.; Grillo, G.; Licciulli, F.; Mignone, F.; Gissi, C. e Saccone, C. (2002). *J. Nucleic Acids Res.*, **30**: 335-340.
- Rubio MA, Pastar I, Gaston KW, Ragone FL, Janzen CJ, Cross GA, Papavasiliou FN, Alfonzo JD. (2007), *Proc Natl Acad Sci U S A*, **104**:7821-7826
- Sali, A. e Blundell, T. L. (1993), *J. Mol. Biol.*, **234**: 779-815.
- Sambrook, J. e Russel, D. W. (2000). *Molecular cloning: a Laboratory Manual*. 3^a ed. Cold Spring Harbor Laboratory Press.
- Sanger, F.; Nicklen, S. e Coulsen, A. R. (1977). *Proc. Natl. Acad. Sci. USA*, **74**: 5463-5467
- Sever, S.; Rogers, K.; Rogers, M. J.; Carter, C. Jr. e Soll, D. (1996). *Biochemistry*, **35**: 32-40.
- Simpson, L., Frech, G. C. e Maslov, D. (1996). *Methods In Enzymol.*, **264**: 99-121.
- Simpson, L.; Simpson, A. M. e Blum, B. (1993). In *RNA Processing - A Practical Approach*. (Hames D. e Higgins, S., Eds.) IRL Press.
- Southern, E. M. (1975). *J. Mol. Biol.*, **98**: 503-517.
- Thiemann, O. H.; Maslov, D.A. e Simpson, L. (1994). *EMBO J.*, **13**: 5689-5700.
- Vriend, G. (1990). *J. Mol. Graph.*, **8**: 52-56.

Capítulo 4

Resultados e discussões*

Neste capítulo são abordados os resultados obtidos com a caracterização molecular e a modelagem por homologia das proteínas LmTrpRS1 e LmTrpRS2. Os experimentos de caracterização das regiões não traduzidas 3' e 5' mostraram que o mRNA do gene *LmTrpRS1* têm 1672 nucleotídeos, e que a janela aberta de leitura depositada no Gene DB (LmjF29.0060) começa na adenina 253 do códon ATG, correspondente a Metionina 85 da proteína depositada no banco de dados, enquanto o mRNA do gene *LmTrpRS2* possui 2258 nucleotídeos. Os dois genes são cópias simples como mostrado nos experimentos de “southern blotting”, localizados nos cromossomos 29 (*LmTrpRS1*) e 23 (*LmTrpRS2*). Ambos os genes foram clonados através de PCR e subclonados em vetores de expressão. Os ensaios de expressão mostraram que a insolubilidade das proteínas foi uma constante na maioria dos vetores utilizados. “western blotting” com extratos celulares de *Leishmania* confirmaram a localização celular das enzimas. Uma forma solúvel e estável da forma madura da proteína LmTrpRS2 expressa no vetor pET28a foi purificada por afinidade, os ensaios de aminoacilação mostraram que esta é capaz de reconhecer os tRNA^{Trp} editado e não editado. A análise estrutural das proteínas foi realizada em modelos gerados por modelagem molecular por homologia, estes mostram que os domínios de ligação aos substratos são conservados com algumas diferenças no mecanismo de reconhecimento do tRNA^{Trp} pela enzima LmTrpRS2.

Caracterização molecular dos genes *LmTrpRS1* e *LmTrpRS2*.

Caracterização das regiões 5' e 3' não traduzidas (UTR) dos genes *LmTrpRS1* e *LmTrpRS2*.

As reações de RT-PCR iniciais que visavam à amplificação da região 5'UTR do gene *LmTrpRS1*, partindo do RNA total de *L. major*, resultaram na amplificação de uma banda não definida, cobrindo as regiões de 100 até 500 pb. A re-amplificação dessa região utilizando-se oligonucleotídeos internos ao gene mostrou 3 bandas

* A divisão dos tópicos descritos nessa sessão não reflete a ordem pelos quais os mesmos foram executados durante o trabalho de doutorado. Essa divisão foi realizada apenas para facilitar a compreensão do trabalho como um todo.

difusas, com tamanhos aproximados de 90, 190 e 300 pares de base (pb) cada uma, calculados com programa Kodak Digital Science 1D. Experimentos realizados com variações na reação de RT-PCR continuaram mostrando as mesmas bandas, mesmo quando as concentrações de RNA total eram aumentadas. Esses fragmentos foram recortados de gel de agarose 1%, purificados com conjunto comercial de purificação de DNA. Após reação a adição de adeninas nas extremidades dos fragmentos obtidos (reação de adenilação), esses foram clonados no vetor pGEM-T[®] (Promega). As seqüências nucleotídicas desses fragmentos obtidas após seqüenciamento do DNA não mostraram similaridade com qualquer seqüência genômica depositada no GeneDB.

Alinhamento seqüencial da proteína LmTrpRS1 com outras TrpRSs (figura I-4.1) mostrou que essa proteína possui uma região N-terminal, compreendida da metionina 1 à valina 79, sem alinhamento algum com as outras TrpRSs. A análise inicial dessa região mostrou que os 30 aminoácidos iniciais seriam destinados a sinalização mitocondrial.

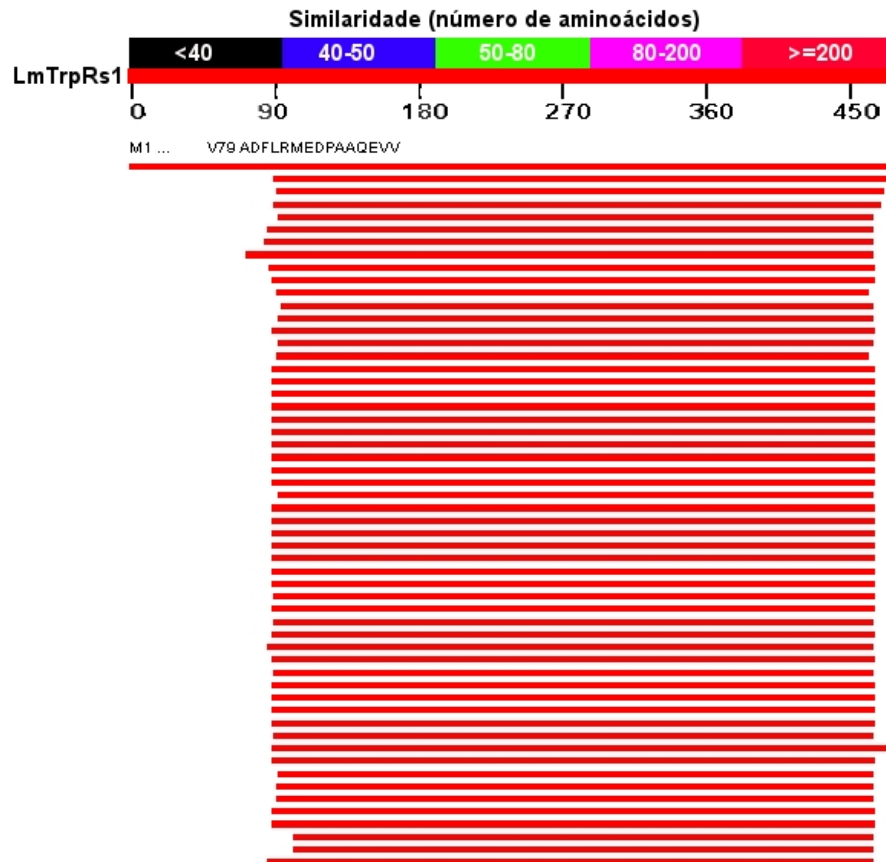


Figura I-4.1: Diagrama mostrando os 100 melhores alinhamento da proteína LmTrpRS1 com seqüências de outras proteínas depositadas no banco de dados. Alinhamentos múltiplos são conectados por uma barra.

Baseando-se nas regiões conservadas dos alinhamentos obtidos, oligonucleotídeos mais internos ao gene *LmTrpRS1* foram construídos para caracterizar completamente essa região.

Utilizando-se esses oligonucleotídeos, uma nova reação de RT-PCR para a região 5'UTR do gene *LmTrpRS1* foi realizada. Como resultados foram obtidas três bandas difusas, com tamanhos entre 400 e 100 pb (figura I-4.2 A). O produto de RT-PCR foi utilizado como molde para um novo ciclo de PCR com oligonucleotídeos mais externos ao gene (figura I-4.1, oligonucleotídeo 2 – Tabela A1), a fim de restringir a reação ao gene de interesse. Uma banda única e bem definida com aproximadamente 170 pb foi observada após o PCR, (figura I-4.2 B). Essa banda foi purificada em gel de agarose, clonada no vetor pGEM-T[®] e o DNA recombinante obtido foi submetido ao seqüenciamento.

A reação de RT-PCR para amplificar a região 3' não traduzida (UTR) do gene *LmTrpRS1* resultou duas bandas bem definidas, uma fraca próxima a 300 pb e outra muito intensa, cobrindo as regiões de 250 até 70 pb, provavelmente esta última seja devido ao excesso de RNA total utilizado na reação (figura I-4.2 C). Essa reação foi re-amplificada, por reação de PCR, utilizando-se oligonucleotídeos mais externos ao gene e também um oligonucleotídeo complementar ao primeiro oligonucleotídeo utilizado no ciclo de RT-PCR, aumentando assim a seletividade da reação (figura I-4.1, oligonucleotídeo 5 e 6 – Tabela A1).

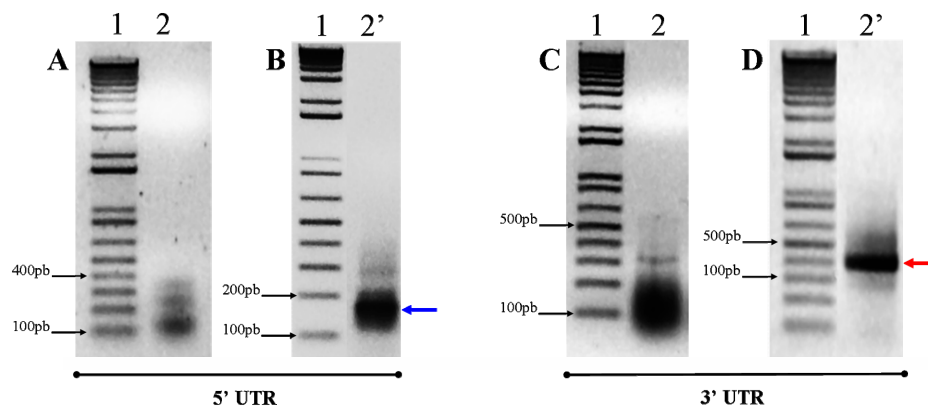


Figura I-4.2: Caracterização das regiões não traduzidas do gene *LmTrpRS1*. Gel de agarose com amostras das reações de RT-PCR (A e C) e a re-amplificação dessas regiões por PCR (B e D). **A e B)** Reação 5'UTR, **C e D)** Reação 3'UTR. **1** - 1 Kb plus DNA ladder (Invitrogen[®]); **2** - Reação de RT-PCR e **2'** - Reações de re-amplificação dos RT-PCR, utilizando oligonucleotídeos mais internos ao gene. As setas indicam as bandas de amplificação das regiões 5'UTR e 3'UTR do gene *LmTrpRS1* (Azul e vermelho respectivamente)

Após o ciclo de PCR, um fragmento de aproximadamente 450 pb foi obtido para a região 3'UTR do gene *LmTrpRS1* (figura I-4.2 D), este fragmento foi purificado em gel de agarose, clonado no vetor pGEM-T[®] e o DNA recombinante obtido foi submetido ao seqüenciamento.

Os mRNA de kinetoplastidas são traduzidos com longos pré-RNAs policistrônicos, de 60 Kb ou mais, que são processados por um mecanismo conhecido com “Trans Splicing” para a formação da extremidade 5'UTR (Ullu *et al.*, 1993; Lebowitz, *et al.*, 1993). Esse mecanismo de “trans splicing” tipicamente ocorre pela inserção de um pequeno fragmento de RNA com 39 ribonucleotídeos, conhecido com “Spliced Leader”, nas regiões intergênicas (Matthews, *et al.*, 1994). Uma pequena região “consensus”, rica em ribonucleotídeos pirimidínicos que é flanqueada por uma base Guanina, atua como sinalizador da correta localização da região de “trans splicing” (Matthews, *et al.*, 1994). A região 3'UTR é formada por uma região que geralmente costuma ser maior que a região 5'UTR. Nessa região 5'UTR ocorre a inserção de uma extremidade formada por vários ribonucleotídeos adenina (cauda poli-A), comumente um ou mais sítios para inserção cauda poli-A podem ser observados nos mRNA de kinetoplastídeos. (Ullu *et al.*, 1993; Lebowitz *et al.*, 1993)

O mRNA do gene *LmTrpRS1* apresenta a organização clássica descrita para os kinetoplastídeos (figura I-4.3). A análise da seqüência 5'UTR mostrou que ela é composta por 100 ribonucleotídeos, idênticos aos dados depositados no genoma de *L. major* (GeneDB). Uma curta região com predominância de bases pirimidínicas é observada precedendo a região 5'UTR, seguindo-se a essa região um ribonucleotídeo guanina sinaliza o local de inserção do “spliced leader” (figura I-4.3).

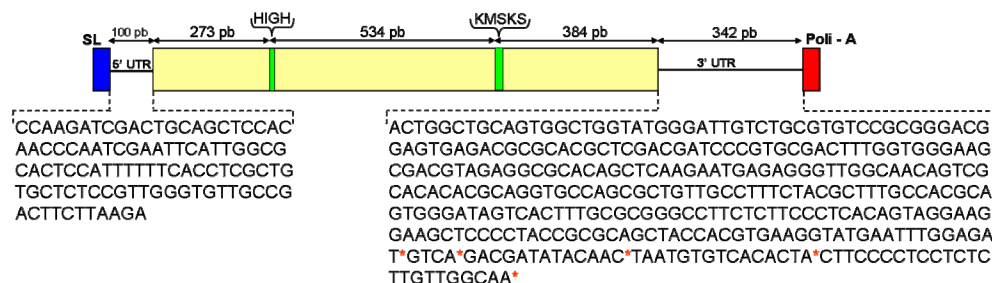


Figura I-4.3: Diagrama mostrando composição mRNA do gene *LmTrpRS1* caracterizado por RT-PCR. Em azul “spliced leader” (SL), em amarelo a janela aberta de leitura do gene *LmTrpRS1* (1191pb), em verde destacando as posições das regiões “consensus” conservadas nas TrpRSs. Em vermelho região poli-A. As seqüências abaixo destacada mostram os 100 nucleotídeos da região 5'UTR (esquerda), e os 342 nucleotídeos da seqüência 3'UTR (direita). Os asteriscos em vermelho apontam locais de inserção da cauda Poli-A do RNA mensageiro.

Já a análise da região 3'UTR mostrou que essa é composta por um máximo 342 ribonucleotídeos, com pelo menos 5 sítios para inserção da cauda poli-A, nos 25 clones analisados. Essa região 3'UTR mostram completa identidade com os dados depositados no GeneDB (figura I-4.3).

O resultado da caracterização da região 5'UTR mostra que a janela aberta de leitura predita do gene *LmTrpRS1* depositada no GeneDB (figura I-4.4) é 252 nucleotídeos maior do que a caracterizada nesse trabalho. Com isso a proteína *LmTrpRS1* é iniciada na Adenina 253 do códon ATG, correspondendo à Metionina 85 da seqüência previamente predita. Essa nova janela aberta de leitura da proteína é condizente com os alinhamentos de outras *TrpRSs* (figura I-4.1) e a predição de sinalização molecular dessa nova janela aberta de leitura não aponta probabilidade de sinalização mitocondrial, como outrora mostrada para a proteína depositada no banco de dados.

```

5' atgatgtccccctttccccgctgctgtctatctctctggcctgctcacccttc
atgttcctactcgctgccttcacttccttgcttacaccaatgaaaaccgggttg
gactgcafcgctgctgcgcatcttacctgcttgctcagccaagatcgactgcagc
tccacaacccaatcgaattcattggcgactccattttttcacctcgctgtgcttc
cgttgggtgttgccgacttcttaagaATGgagg...3'
    
```

Figura I-4.4: Seqüência nucleotídica da porção 5' do gene *LmTrpRS1*. Em amarelo mostra a metionina iniciadora da proteína predita e depositada no banco de dados em decorrência do projeto genoma de *L. major*. Em vermelho são mostrados os 100 nucleotídeos da região 5'UTR. Em cinza a região poli-pirimidínica, que é flanqueada por uma Guanina, em azul. Em itálico a seqüência do gene *LmTrpRS1*, com a metionina iniciadora correta em verde.

A caracterização do mRNA do gene *LmTrpRS2* foi realizado de maneira idêntica ao procedimento descrito para o gene *LmTrpRS1*. O produto da reação de RT-PCR da região 5'UTR do gene *LmTrpRS2*, resultou uma banda bem fraca e não definida, cobrindo as regiões de 100 até 400 pb (figura I-4.5 A-B). Já reação de RT-PCR da região 3'UTR resultou duas bandas, uma não definida cobrindo as regiões de 100 até 400 pb e outra mais fraca na região de 700 pb. (figura I-4.5 C). Após a re-amplificação desses produtos, por reações de PCR, utilizando-se oligonucleotídeos específicos, um fragmento de aproximadamente 190 pb foi obtido para a região 5'UTR (figura I-4.5 B) e outro de aproximadamente 630 pb para a região 3'UTR. Estes fragmentos foram clonados no vetor de pGEM-T[®], seqüenciados e as seqüências obtidas foram alinhadas com as seqüências nucleotídicas de *L. major*.

A análise da sequência da região 5'UTR mostrou que essa é composta por 134 pb, uma ampla região poli-pirimidínica flanqueada por uma base Guanina, que sinaliza para a inserção “spliced leader” no mRNA pode ser observado (figura I-4.6). Já a sequência da região 3'UTR é composta por 612pb.

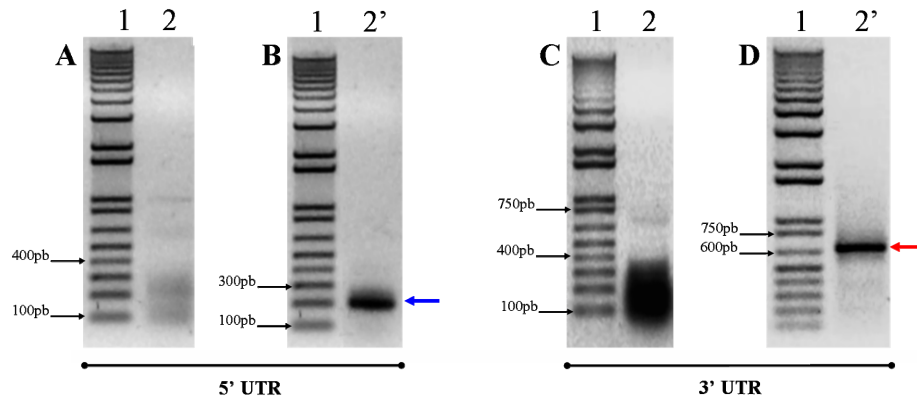


Figura I-4.5: Caracterização das regiões não traduzidas do gene *LmTrpRS2*. Gel de agarose com amostras das reações de RT-PCR (A e C) e a re-amplificação dessas regiões por PCR (B e D). **A e B)** Reação 5'UTR, **C e D)** Reação 3'UTR. **1 -** 1 Kb plus DNA ladder (Invitrogen®); **2 -** Reação de RT-PCR e **2' -** Reações de re-amplificação dos RT-PCR, utilizando oligonucleotídeos mais internos ao gene. As setas indicam as bandas de amplificação das regiões 5'UTR e 3'UTR do gene *LmTrpRS2* (Azul e vermelho respectivamente)

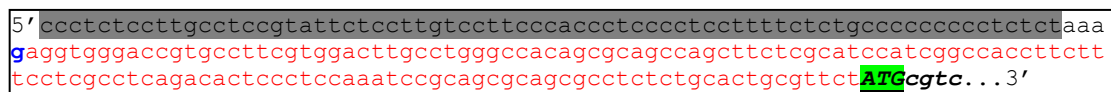


Figura I-4.6: Sequência nucleotídica da porção 5' do gene *LmTrpRS2*. Em vermelho são mostrados os 134 nucleotídeos da região 5'UTR. Em cinza a região poli-pirimidínica, que é flanqueada por uma guanina, em azul. Em itálico a sequência do gene *LmTrpRS2*, com a metionina iniciadora em verde.

O mRNA do gene *LmTrpRS2* foi definido como uma sequência de 2333 nucleotídeos (figura I-4.7).

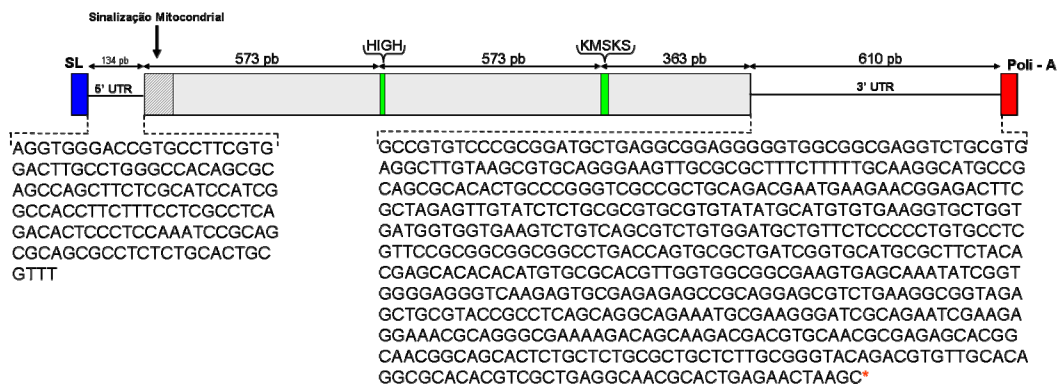


Figura I-4.7: Diagrama mostrando composição mRNA do gene *LmTrpRS2* caracterizado por RT-PCR. Em azul “Spliced Leader” (SL), em cinza a janela aberta de leitura do gene *LmTrpRS2* (1509pb), em verde destacando as posições das regiões “consensus” conservadas na TrpRSs. Em vermelho região poli-A. As sequências abaixo destacada mostram os 134 nucleotídeos da região 5'UTR (esquerda), e os 610 nucleotídeos da sequência 3'UTR (direita). O asterisco em vermelho aponta o local de inserção da cauda poli-A dos mRNAs.

Caracterização do número de cópias dos genes *LmTrpRS1* e *LmTrpRS2*.

A análise dos sítios de clivagem dos genes de interesse para as endonucleases de restrição foi realizada com o programa WebCutter 2.0 (<http://rna.lundberg.gu.se/cutter2/>). Combinações de endonucleases de restrição foram realizadas de maneira que os produtos das digestões mantivessem os genes íntegros ou então os clivassem em pelo menos um ponto.

Os experimentos de “southern blotting”, provenientes das digestões do DNA genômico de *L. major*, revelaram que as enzimas de restrição que não clivam dentro da seqüência do gene *LmTrpRS1*, *Nde* I, *Xba* I, *Hind* III e *Bgl* II, (figura I-4.8 A), gerando um único fragmento de DNA homólogo à sonda (seqüência nucleotídica completa do gene marcado). Já as enzimas de restrição *Not* I, *Xho* I, e *EcoR* I, que apresentam sítio de clivagem único (figura I-4.8 A) interno à fase aberta de leitura do gene *LmTrpRS1*, resultaram no padrão esperado de duas bandas hibridizadas. A endonuclease de restrição *Pst* I que apresenta dois sítios de clivagem internos à fase aberta de leitura deveria produzir três bandas de hibridização, porém somente duas bandas foram observadas (figura I-4.8 A). Isso ocorreu provavelmente pelo pequeno tamanho de um dos fragmentos, 160pb aproximadamente, que produz uma pequena região de hibridização, que podem ter sido removidas durante as lavagens de estringência.

A sonda radioativa *LmTrpRS1* foi removida da membrana com lavagens em solução SDS 2% a 75° C (Sambrook e Russel, 2000). A mesma membrana foi então hibridizada com a sonda *LmTrpRS2* (seqüência nucleotídica completa do gene), esse experimento revelou que as enzimas de restrição *Nde* I, *Not* I, *Xho* I, *Xba* I, *EcoR* I e *Bgl* II as quais não clivam dentro da seqüência do gene *LmTrpRS2*, geraram um único fragmento de DNA homólogo à sonda (figura I-4.8 B). Já as enzimas *Hind* III e *Pst* I e, que apresentam sítio de clivagem único, interno à fase aberta de leitura do gene *LmTrpRS1*, resultaram no padrão esperado de duas bandas hibridizadas (figura I-4.8 B).

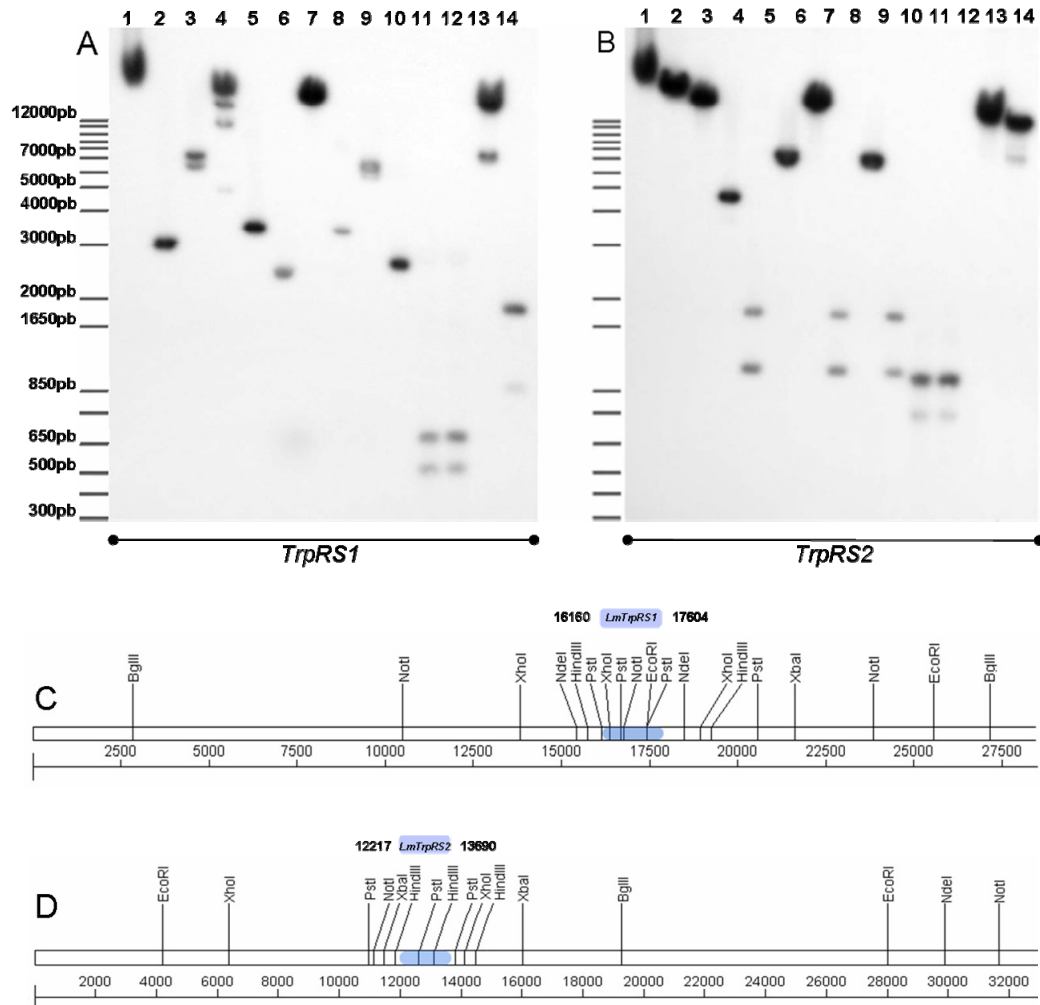


Figura I-4.8: Caracterização do número de cópias dos genes *LmTrpRS1* e *LmTrpRS2*. A e B) Análise de “southern blotting” proveniente da digestão do DNA genômico de *L. major* Friedlin, com diferentes endonucleases de restrição, hibridizadas com as sondas *LmTrpRS1* (A) e *LmTrpRS2* (B). 1 - DNA Genômico de *L. Major* não digerido. 2 - *Nde* I, 3 - *Not* I, 4 - *Xba* I, 5 - *Hind* II, 6 - *Xho* I, 7 - *Bgl* II, 8 - *Hind* III/*Xba* I, 9 - *Not* I/*Bgl*, 10 - *Hind* III/*Nde*, 11 - *Pst* I, 12 - *Pst* I/*Bgl* II, 13 - *Eco* RI e 14 - *Eco* RI/*Nde* I. C e D) Diagrama da região gênica codificada no DNA do parasita *L. major* para os genes *LmTrpRS1* e *LmTrpRS2*.

Estes resultados mostram que os genes *LmTrpRS1* e *LmTrpRS2* são genes de cópia simples no genoma do parasita *L. major*. Este dado foi confirmado após a completa deposição do genoma de *L. major*. Uma análise dessa região gênica codificada no DNA do parasita é mostrada na figura I-4.8 C e D.

Clonagem dos genes *LmTrpRS1* e *LmTrpRS2*.

A janela aberta de leitura do gene *LmTrpRS1* depositada no GeneDB, foi amplificada com sucesso pela técnica de PCR, utilizando-se DNA total extraído de culturas celulares de *L. major*, uma banda única com 1440 pb foi observada (figura I-4.9 A), esta banda foi purificada e clonada diretamente no vetor pGEMT[®]. O DNA

recombinante das colônias bacterianas transformantes foi seqüenciado, e sua seqüência nucleotídica foi alinhada contra o banco de dados GeneDB, mostrando que nenhuma mutação foi inserida no gene durante o processo de amplificação. Esse clone sem mutações, denominado *LmTrpRS1-full/pgemT*, foi propagado em bactérias *E. coli* DH5 α . Após extração plasmidial o DNA deste clone foi armazenado a -20° C para os demais procedimentos de sub-clonagem.

A caracterização da região 5'UTR do gene *LmTrpRS1* mostrou que a janela aberta de leitura do gene depositada estava superestimada em 252 nucleotídeos, então uma nova reação de amplificação foi utilizada para a remoção dessa região do gene *LmTrpRS1*. Um ciclo de PCR utilizando oligonucleotídeo compatível com a região caracterizada foi realizado utilizando-se como molde o gene previamente clonado. Um fragmento de 1191 pb (figura I-4.9 B) foi obtido e após purificação e clonagem no vetor pGEM[®]-T, esse fragmento foi seqüenciado mostrando 100% de identidade com a seqüência nucleotídica depositada no banco de dados. Esse novo clone sem mutações e com a janela aberta de leitura correta foi denominado *LmTrpRS1/pgemT*. O mesmo foi propagado em bactérias *E. coli* DH5 α e após extração plasmidial o seu DNA plasmidial foi armazenado a -20° C para os demais procedimentos de sub-clonagem.

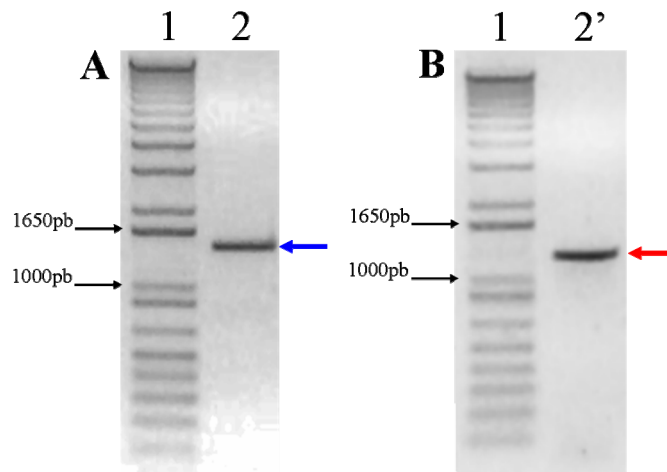


Figura I-4.9: Gel em agarose a 1% com resultado da amplificação do gene *LmTrpRS1*. **A)** Gene *LmTrpRS1* com a janela aberta de leitura de leitura depositada no GeneDB, a seta azul indica a posição do gene (1440pb). **B)** Gene *LmTrpRS1* com a janela de leitura caracterizada pelo experimento de 5'UTR, a seta vermelha indica o gene *LmTrpRS1* (1191pb). **1** – 1 kb Plus DNA Ladder, **2 e 2'** – Amostras dos genes amplificados.

A análise teórica da seqüência de aminoácidos das proteínas LmTrpRS1 e LmTrpRS2 com o programa MitoprotII identificou uma região com características de

sinalização mitocondrial na proteína LmTrpRS2. Essa região composta por uma região N-terminal de 23 aminoácidos, apresentou 93% de probabilidade para exportação mitocondrial, segundo o Mitoprot II (figura I-4.10). Segundo esse programa o sítio de clivagem está localizado entre a Asparagina 23 e a Alanina 24.

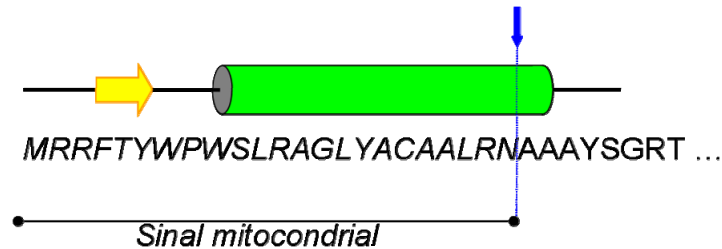


Figura I-4.10: Esquema da região da região de sinalização mitocondrial da proteína LmTrpRS2 predita pelo programa MitoprotII. Seqüência de aminoácidos dos 32 aminoácidos N-terminal da proteína LmTrpRS2, em itálico é mostrado o sinal de localização mitocondrial. Acima da seqüência a predição de estrutura secundária assinalada pelo programa PSIPRED (Jones, 1999), fita- β (amarelo) e hélices- α (verde), a seta em azul mostra o local da clivagem do N-terminal.

As regiões destinadas à localização mitocondrial caracterizam-se em grande parte por pequenas seqüências amino ou carboxi-terminais com 20 a 60 resíduos de aminoácidos, com grande predominância de aminoácidos positivos e resíduos hidroxilados, e raros ou inexistentes aminoácidos negativos (Lemire *et al.*, 1989; Neupert, 1997). Essas seqüências comumente formam hélices- α anfipáticas que são elementos importantes para o reconhecimento específico da maquinaria de importação mitocondrial. (Lemire *et al.*, 1989; Von Heijne *et al.*, 1990). Uma vez as proteínas internalizadas na mitocôndria, essas seqüências são prontamente removidas por proteases específicas, assegurando a manutenção da proteína importada na matriz mitocondrial (Pfanmer e Neupert, 1990; Braun e Schmitz, 1994; Neupert, 1997). Essas características estruturais e seqüenciais podem ser observadas na porção N-terminal da proteína LmTrpRS2, apontando a possível localização mitocondrial dessa enzima (figura I-4.10).

Duas abordagens para clonar o gene *LmTrpRS2* foram empregadas: clonagem do gene completo e do gene sem a região de sinalização. A primeira é justificada pela importância que regiões de sinalização podem apresentar no direcionamento do correto enovelamento de certas proteínas (Pfanmer e Neupert, 1990, Neupert, 1997), já a segunda é justificada pelo fato de que a proteína madura no interior da mitocôndria não necessita desta região N-terminal para sua completa atividade

(Neupert, 1997; Felter *et al.*, 1981). Além disso, essa região poderia também influenciar nos ensaios de cristalização, uma vez que essas regiões sinalizadoras apresentam uma grande flexibilidade.

As duas construções para a proteína LmTrpRS2 foram amplificadas com sucesso através de PCR (figura I-4.11), utilizando-se oligonucleotídeos específicos com sítios de restrição para as endonucleases de restrição *NdeI* e *XhoI* (para maiores detalhes consultar a tabela AI-1.2 no apêndice I). Essa estratégia de clonagem permite que o gene de interesse seja clonado diretamente em um vetor de expressão através da digestão do produto de PCR, ou então seja subclonado inicialmente em um vetor de propagação.

A reação de PCR para o gene completo, denominada *LmTrpRS2-Full*, produziu um fragmento único com tamanho condizente ao esperado para gene completo. A construção do gene sem a região de sinalização mitocondrial, *LmTrpRS2-A24* (proteína LmTrpRS2 iniciando no aminoácido Alanina 24) produziu uma banda única com tamanho aproximado de 1400 pb. Essas duas bandas foram separadas em gel de agarose a 1% e purificadas. Após reação de adenilação as construções purificadas foram clonadas diretamente no vetor propagação pGEMT[®], gerando os plasmídeos *LmTrpRS2-Full/pgemT* e *LmTrpRS1-A24/pgemT*. O DNA recombinante obtido das colônias bacterianas transformantes foi seqüenciado e alinhado contra o banco de dados GeneDB. O seqüenciamento mostrou que os fragmentos clonados para o gene *LmTrpRS2-Full* e *LmTrpRS1-A24* possuíam 1473 e 1404 pb, respectivamente.

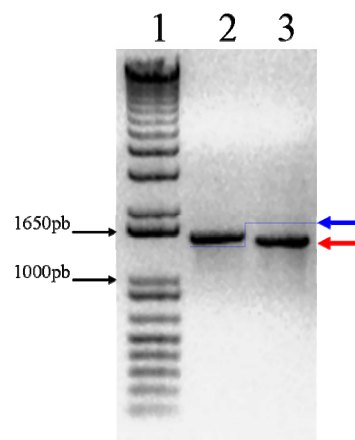


Figura I-4.11: Gel em agarose a 1% com resultado da amplificação do gene *LmTrpRS2* de *L. major*. **1** – 1 kb Plus DNA Ladder, **2** - Gene *LmTrpRS2-full* (com a região de sinalização mitocondrial), a seta azul indica a posição do gene (1473pb). **3** – Gene *LmTrpRS2-24A* (sem a região de sinalização mitocondrial), a seta vermelha indica a posição do gene (1191pb).

O seqüenciamento também mostrou que nenhuma mutação foi inserida nos genes durante o processo de amplificação. Os clones sem mutações foram propagados em bactérias *E. coli* DH5 α , e após extração plasmidial o DNA foi armazenado a -20° C para posterior utilização nos demais procedimentos de sub-clonagem e biologia molecular.

Sub-clonagem dos genes *LmTrpRS1* e *LmTrpRS2* em vetores de expressão e ensaios de expressão protéica.

A obtenção de um sistema de expressão heterólogo que produza grandes quantidades de proteína na forma solúvel ainda é uma ciência bastante empírica dentro da biologia molecular. Um somatório de fatores podem ser apontadas como determinantes para o processo de expressão protéica, como: tamanho da proteína, ponto isoelétrico, hidrofobicidade, modificações pós-traducionais, necessidade de outras proteínas para formação de complexos estáveis etc. Algumas ferramentas computacionais podem ser utilizadas para prever o perfil bioquímico de uma dada proteína e tentar adequar o sistema de expressão mais apropriado para a mesma. O sistema de expressão procarioto tem se mostrado bastante eficiente para a expressão e purificação de várias AaRSs de eucariotos (Chen *et al.*, 1999; Xu *et al.*, 2001).

Inicialmente os dois genes de interesse, codificando as proteínas LmTrpRS1 e a LmTrpRS2 com e sem a região de sinalização mitocondrial foram sub-clonados no vetor de expressão pET29a (novagem), que tem se mostrado como um sistema poderoso para a expressão de grande quantidade de proteína recombinante em *E. coli*. Os genes de interesse foram clonados de modo que as proteínas recombinantes não possuíssem cauda de histidinas C-terminal (pET29a_stop).

A expressão dos genes nestes vetores é regulada pelo promotor reconhecido pela enzima T7 RNA polimerase. Para a expressão do gene de interesse é necessário que a célula hospedeira produza a enzima T7 RNA polimerase, o que pode ser obtido a partir de diversas linhagens de *E. coli* que codificam esse gene. A expressão do gene de interesse ocorre na presença do indutor Isopropil β -D-tiogalactopiranosídeo (IPTG). A T7 RNA polimerase se liga ao promotor T7, transcrevendo o gene de interesse (Sambrook e Russel, 2000).

Os ensaios iniciais de expressão foram realizados a 37° C, variando-se a concentração do indutor IPTG (100 μ M, 400 μ M e 1 mM). Esses ensaios mostraram

que ambas a proteínas eram expressas em grandes quantidades na fração insolúvel, com nenhuma porção visível de proteína recombinante na fração solúvel.

Objetivando a expressão solúvel das proteínas recombinantes novos ensaios de expressão foram realizados. Essas novas tentativas envolveram expressão em meios de cultura nutritivos diferenciados (meio mínimo, 2XYT, LB), utilizando-se concentrações menores de indutor IPTG e também redução da temperatura de indução. A diminuição nos níveis de expressão das proteínas, obtidas em condições de expressão a temperaturas baixas (22 e 18° C) e com quantidades pequenas de indutor (menores que 10 µM) não mostraram melhora significativa na solubilidade proteica. Uma tentativa sem sucesso também foi realizada utilizando-se a expressão da proteína sem a presença de indutor, nesse caso esperava-se que o vazamento do promotor pudesse expressar a proteína em níveis basais, nos quais maquinaria celular fosse capaz de enovelar corretamente a proteína.

O emprego de cepas de *E. coli* portadoras de características diferenciadas como: códons extras para alguns aminoácidos, BL21(De3)RIL (Stratagene®) e BL21(De3)RP (Stratagene®); chaperonas (Chang e Cohen, 1978) PT-Gro e Rosetta (invitrogen®); restrição do vazamento do promotor, BL21(De3)plysS; bactérias destinadas à produção em grande escala de genes tóxicos e proteínas de membrana (Merle *et al.*, 1986), foram testadas mas também não produziram proteína solúvel.

Uma das metodologias utilizadas para aumentar a solubilidade de proteína é a fusão da proteína alvo com proteínas altamente solúveis tais como: GST (Glutathione S-transferase de *Schistosoma japonicum*), tioredoxina, MBP (Maltose Binding Protein), NusA (proteína de *E. coli*), o domínio-Gb1 (domínio GB1 da proteína G de *Streptococcus sp.*) e o duplo domínio-Z (ZZ – derivado da proteína A de *Staphylococcus aureus*), fusionadas geralmente na porção N-terminal da proteína alvo (Hammarstrom *et al.*, 2002).

Na tentativa de contornar as dificuldades técnicas para expressão e solubilidade das proteínas, algumas construções com plasmídeos bacterianos codificando proteínas de fusão com Glutathione S-transferase (pGEX2T) e Tioredoxina (pET32a) foram testadas, bem como a alteração da posição da cauda de histidina ou ainda sua remoção (figura I-4.12). No caso da proteína LmTrpRS2 também foram realizadas construções aonde a região de sinalização mitocondrial foi removida, uma vez que esta poderia influenciar o enovelamento protéico.

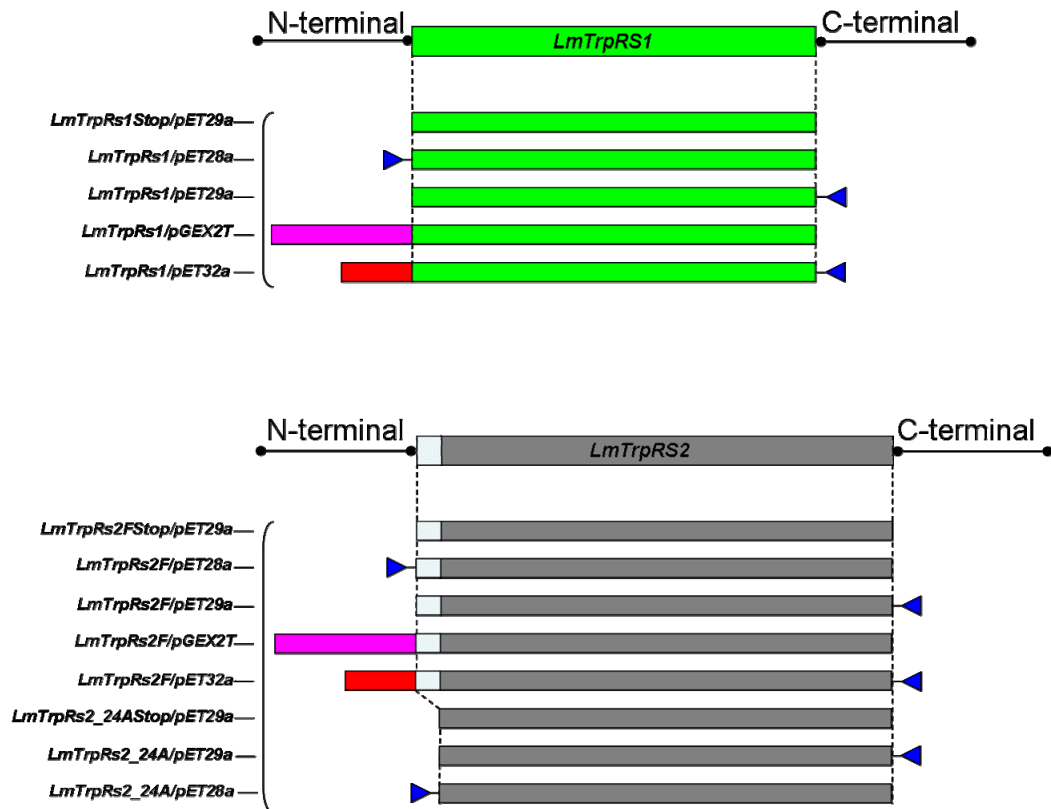


Figura I-4.12: Diagrama das construções realizadas com o gene *LmTrpRS1* e *LmTrpRS2*. Os triângulos representam cauda de histidinas (6-HIS tag), Em rosa a proteínas de fusão GST, em vermelho tioredoxina, região demarcada no início do gene *LmTrpRS2* representa sinal de localização mitocondrial.

As sub-clonagens dos genes *LmTrpRS1/pET28a*, *LmTrpRS1/pET29a*, *LmTrpRS1/pET32a* e *LmTrpRS1/pGEX2T* obtidas foram realizadas a partir de PCR utilizando-se oligonucleotídeos específicos (tabela AI-1.3 – Apêndice I) e o plasmídeo recombinante *LmTrpRS1/pgemT* como molde. Bactérias *E. coli* competentes foram transformadas com o produto da ligação. Após seleção dos recombinantes em meio sólido contendo antibiótico adequado o DNA plasmidial foi extraído e a conferência da presença do inserto foi verificada por reação de digestão seguida de gel de agarose a 1% (figura I-4.13). O seqüenciamento dos clones recombinantes mostrou que todas as construções estavam em fase de leitura e sem mutações.

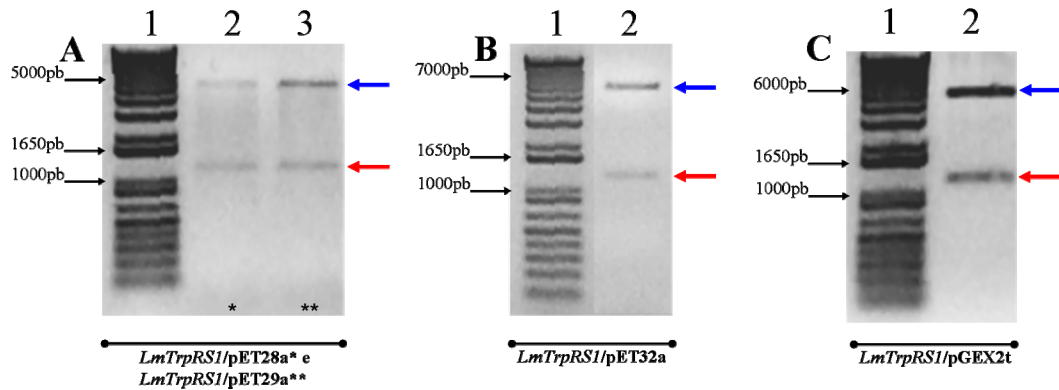


Figura I-4.13: Digestão com endonucleases de restrição das construções realizadas com o gene *Lmwars1* nos vetores bacterianos. Gel de agarose 1% com o resultado da sub-clonagem do gene pET28a, pET29a, pET32a e pGEX2t **A)** Construções *LmTrpRS1/pET28a* (poço 2) e *LmTrpRS1/pET29a* (poço 3) **B)** Construção *LmTrpRS1/pET32a* (poço 2). **C)** *LmTrpRS1/pGEX2t* (poço 2). Poço 1, 1kB plus DNA Ladder. As setas azuis indicam o gene *Lmwars1* e as vermelhas os plasmídeos.

Os ensaios de expressão realizados com as novas construções da proteína *LmTrpRS1* tanto no vetor pET28a quanto no pET29a continuam mostrando a proteína sendo expressa em grande quantidade na sua forma insolúvel, mesmo em condições de crescimento e expressão lentas, ou seja, pequenas concentrações (menores que 10 μ M) de indutor e temperatura de indução de 18° C (figura I-4.14).

Os ensaios de expressão no vetor pET32a mostraram resultado semelhante, uma grande quantidade de proteína expressa, mas nenhuma proteína solúvel foi obtida com esse vetor, mesmo após purificação com resina de níquel. O vetor pGEX2T não mostrou expressão da proteína de interesse.

As sub-clonagens da proteína *LmTrpRS2* para formar as construções *LmTrpRS2F/pET28a*, *LmTrpRS2F/pET29a*, *LmTrpRS2F/pET32a*, *LmTrpRS2F/pGEX2T*, *LmTrpRS2-24Astop/pET29a*, *LmTrpRS2-24A/pET29a* e *LmTrpRS2-24A/pET28a*, foram realizadas a partir de PCR utilizando-se oligonucleotídeos específicos (tabela AI-1.3 – Apêndice I) e o plasmídeo recombinante *LmTrpRS2Full/pgemT* como molde. Após o seqüenciamento do DNA das colônias recombinantes, os clones sem mutações foram inseridos em bactérias de expressão *E. coli*. Os resultados obtidos para as construções com proteína de fusão *LmTrpRS2* foram semelhantes aos obtidos para a citoplasmática, uma grande expressão protéica, porém totalmente insolúvel.

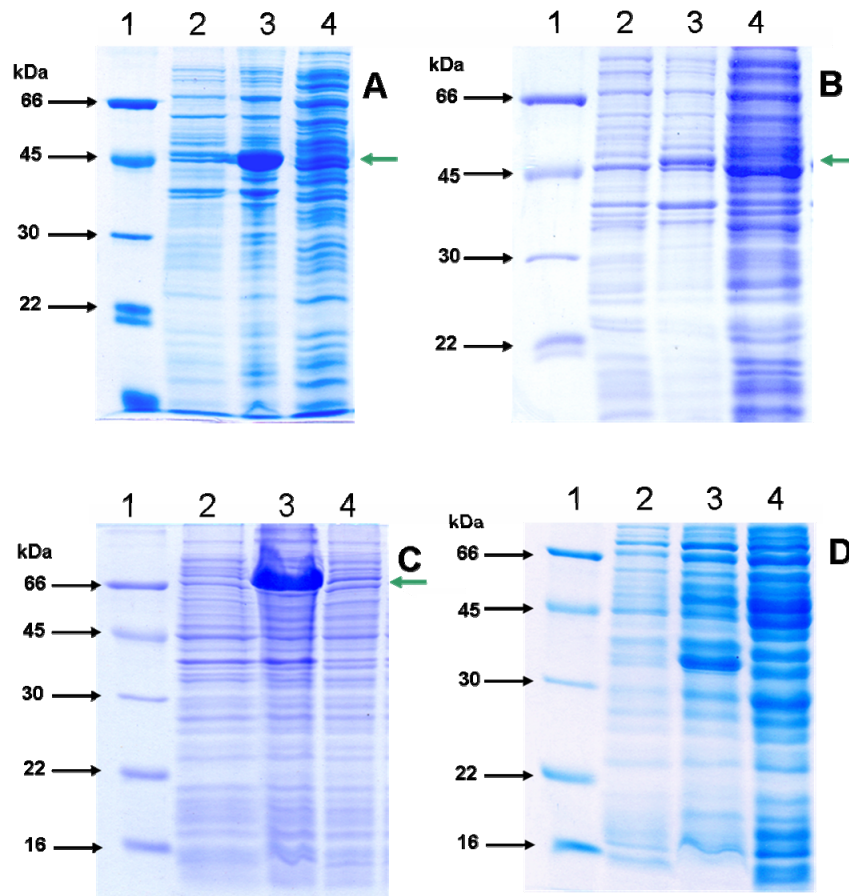


Figura I-4.14: Resultados da expressão da proteína LmTrpRS1 em vários vetores de expressão bacterianos. Eletroforese em gel de poliacrilamida a 15% mostrando a expressão da proteína LmTrpRS1 nos vetores pET28a (A), pET29a (B), pET32a (C) e pGEX2T (D). 1 – Marcador de massa molecular, 2 - Amostras de cultura celular não induzida, 3 e 4 - frações insolúveis e solúveis da expressão a 20° C. A seta verde mostra a proteína de interesse expressa nos vetores pET29a e pET28a, aproximadamente 45 kDa e no vetor pET32a e, aproximadamente 62 kDa.

Dentre as construções testadas, apenas a construção *LmTrpRS2-24A/pET28* (figura I-4.15 A) apresentou os melhores resultados de expressão e purificação. Uma característica importante conferida por essa construção (*LmTrpRS2-24A/pET28*) era a presença de uma cauda de seis histidinas na porção N-terminal da proteína, o que permitiu a purificação do extrato bruto bacteriano em uma coluna de afinidade utilizando-se resina de níquel. Uma pequena fração solúvel da proteína LmTrpRS2 pode ser observada após a eluição da resina de níquel com imidazol (figura I-4.15 B). Essa proteína eluída da coluna apresentou-se bastante instável, após sua purificação em coluna de afinidade, mesmo em concentrações diluídas. Mesmo apresentando essa instabilidade a construção *LmTrpRS2-24A/pET28* foi utilizada para a realização dos demais procedimentos de caracterização molecular, como serão discutidos a seguir.

A fim de melhorar a estabilidade protéica da construção *LmTrpRS2-24A/pET28* alguns ensaios utilizando alguns adjuvantes como glicerol, detergente, cofatores ATP, agentes antioxidantes foram realizados. Porém estes não foram capazes de estabilizar a proteína em solução, mesmo em purificações a baixas temperaturas (4° C). Experimentos de espalhamento dinâmico de luz mostraram que concentrações baixas da proteína permaneciam estáveis por mais tempo em solução, essas concentrações na ordem de 0,5 mg/mL foi suficiente para a realização dos ensaios enzimáticos.

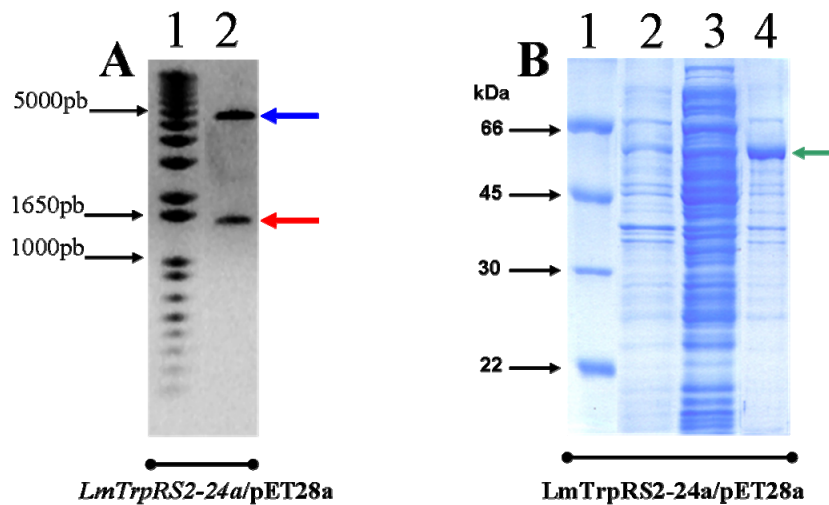


Figura I-4.15: Resultado da sub-clonagem do gene *LmTrpRS2-24A* no vetor pET28a e a expressão da proteína *LmTrpRS2* nesse vetor. A) Gel de agarose 1% com digestão da construção *LmTrpRS2-24A/pET28a* com endonucleases de restrição, 1- 1kb plus DNA ladder, 2 – Digestão do plasmídeo *LmTrpRS2-24A/pET28a* com as endonucleases de restrição *XhoI* e *NdeI*. Em azul posição do vetor pET28a em vermelho o gene *LmTrpRS2-24A*. B) Gel de poliácridamida a 15% mostrando a expressão da proteína *LmTrpRS2* no vetor pET28a, 1 – Marcador de massa molecular, 2 - Amostras de cultura celular não induzida, 3 e 4 – fração solúvel e insolúvel da expressão a 20° C da proteína *LmTrpRS2* no vetor pET28a. A seta verde mostra a *LmTrpRS2-24A*, aproximadamente 44 kDa.

As tentativas de concentrar a proteína mostraram a formação de precipitados à medida que a concentração da proteína aumentava, impossibilitando a concentração da mesma.

Paralelamente aos testes acima citados diversas construções utilizando-se outros vetores de expressão foram empregados para as tentativas de expressão das proteínas de interesse na fração solúvel. De maneira sumarizada podemos citar que foram utilizados diversos vetores de expressão bacterianos e ainda vetores de expressão para os sistemas eucarióticos de *Pichia pastoris* (vetores pPIC-Invitrogen), *Leishmania sp.* e *Crithidia fasciculata* (pNus) (Tetauld *et al.*, 2002) e *L. tarentolae*

(pF4X1.4neo - Jena[®]) (Breitling *et al.*, 2002). Dados mais detalhados sobre cada construção podem ser observados na tabela AI-1.3 – apêndice I.

Os resultados obtidos para quatro construções realizadas em *Pichia pastoris* mostraram que apenas uma construção quando expressa na cepa de *P. pastoris* Gs-115 produziu proteína solúvel. Essa expressão era visualizada somente após a purificação do extrato celular seguido de “western blotting”, aonde foi observado uma banda de reação fraca, indicando um baixo nível de expressão da mesma, o que inviabilizou sua produção para este trabalho.

Vários sistemas para a expressão heteróloga de proteínas em kinetoplastidas são atualmente disponíveis como: pTEX vetor para expressão em *Leishmania* e *T. cruzi* (Kelly *et al.*, 1992; Martinez-Calvillo *et al.*, 1997); pX, vetor para *Leishmania* (LeBowitz *et al.*, 1990; Coburn *et al.*, 1991); pNus para *Leishmania sp.* e *Crithidia fasciculata* (Tetauld *et al.*, 2002); pF4X1.4neo para *L. tarentolae* (Breitling *et al.*, 2002) e pTSO-HYG4 *T. brucei* (Sommer *et al.*, 1996). Esses vetores têm a desvantagem de produzir proteínas em pequena escala, com risco de contaminação biológica utilizando meios de cultura caros, porém a sua utilização pode ser uma alternativa para a produção de proteínas que mostram problemas durante a expressão. Nesse intuito algumas construções das proteínas LmTrpRSs foram testadas em alguns desses vetores. As construções (pNUS e pF4XI-neo) mostraram uma dificuldade de manipulação, principalmente na seleção das culturas, assim após diversas tentativas de eletroporação com plasmídeos recombinantes e seleção com antibiótico adequado nenhuma colônia recombinante foi obtida e as construções destinadas à produção heteróloga de proteína também foram descontinuadas.

A expressão insolúvel das AaRSs tem se mostrado um dos desafios aos grupos que trabalham com essas enzimas (Prof. Dr. Dieter Söll, comunicação pessoal). Um protocolo alternativo como a adição de etanol durante a indução que mostrou resultados para o aumento da solubilidade da glutaminil tRNA sintetase (Rinehart *et al.*, 2004) foi testado para as LmTrpRSs não mostrando alterações significativas na solubilidade das mesmas. Nenhum dos sistemas testados quer procaríoto ou eucarioto conseguiu produzir de maneira satisfatória as proteínas de interesse na sua forma solúvel, principalmente para as caracterizações estruturais, aonde grande quantidade de proteína e altas concentrações são necessárias. O melhor

resultado obtido foi para a proteína mitocondrial, clonada sem a região de sinalização no vetor de expressão pET28a, que será discutido adiante.

Produção de anticorpos policlonais contra as proteínas LmTrpRS1 e LmTrpRS2.

Os anticorpos policlonais para as duas proteínas de interesse foram produzidos em fêmeas camundongos Swiss (foram utilizadas as fêmeas por serem mais abundantes no biotério da UFSCar) com supervisão da prof^a. Dr^a. Heloísa Selistre Araújo. Duas imunizações foram aplicadas em cada animal, na primeira uma quantidade de aproximadamente 100 µg de proteína (macerado de um gel de poliacrilamida) juntamente com adjuvante completo de Freud (Sigma-Aldrich) foi injetada em cada animal. Tanto o adjuvante quanto os fragmentos de acrilamida injetados nos animais auxiliaram a formação da resposta imune primária, após 45 dias uma nova injeção com o mesmo macerado protéico foi injetado. Essa segunda injeção visa a produção aguda de anticorpos, no pico da produção de anticorpos, nono e décimo dia, os animais foram sacrificados em clorofórmio. O soro com a fração de anticorpos desses animais foi separado do sangue e utilizado para os experimentos de “western blotting”. Os títulos para os anticorpos foram testados com o soro de cada animal, a fim de escolher o anticorpo a ser utilizado nos demais experimentos, bem como a sua diluição. Para isso amostras das proteínas LmTrpRS1 e LmTrpRS2 recombinantes foram separadas em gel de poliacrilamida e transferidas para membrana de nitrocelulose (GE-Healthcare). Essa membrana foi corada com solução de Ponceau 0,2% e as bandas respectivas das proteínas de interesse foram marcadas. A membrana foi então bloqueada com leite desnatado a 5% por 1 hora em temperatura ambiente. Os anticorpos foram diluídos em tampão TBS (Sambrook e Russel, 2000) na proporção de 1:1000, 1:5000, 1:10000, 1:25000, 1:50000.

A detecção das reações de imunoblotting foi realizada incubando-se a membrana com um anticorpo anti-igG de camundongo conjugado com a enzima fosfatase alcalina. Após lavagem da membrana um substrato fluoróforo (AP Conjugate Substrate Kit, BCIP e NBT - Biorad) foi adicionado à membrana para produzir uma banda com coloração visível.

Uma única banda de reação com aproximadamente 44 kDa foi observado para a proteína citoplasmática (figura I-4.16 A), esse anticorpo apresentou um título de

aproximadamente 1:25000 nessa reação, sendo que as melhores reações foram observadas na proporção de 1:10000, que foi o título usado posteriormente nos experimentos de western-blotting. O anticorpo contra a proteína LmTrpRS2 reagiu com pelo menos três bandas de intensidades semelhantes com tamanhos que variavam de 50 até 75 kDa (figura I-4.16 B). A banda de aproximadamente 53 kDa corresponde à proteína LmTrpRS2, já as duas outras acima desta são reações cruzadas (figura I-4.16 B), causadas ou pela injeção de proteínas contaminantes juntamente com a proteína LmTrpRS2 no momento da inoculação dos camundongos, ou ainda por anticorpos do próprio animal contra proteínas bacterianas. Uma estratégia de neutralização das reações cruzadas foi realizada através do preparo de um lisado bacteriano da mesma linhagem da bactéria utilizada na expressão da proteína recombinantes, porém sem o plasmídeo codificante da proteína de interesse (Sambrook e Russel, 2000). Esse lisado foi incubado com os anti-LmTrpRS2 por 4 horas a 37° C, sendo então centrifugado e utilizado. Porém mesmo atenuadas as reações cruzadas ainda permaneceram visíveis. Diluições superiores a 1:50000 puderam ser observados para o anti-LmTrpRS2, uma vez que na diluição 1:50000 havia a presença de uma banda de reação muito intensa. Para verificar se as reações cruzadas poderiam interagir com as proteínas em *L. major*, um “western blotting” contra um extrato bruto com 10^6 células de *L. major* foi realizado. Nenhuma reação cruzada foi observada no extrato de *L. major* (figura I-4.16 C), nessa reação foi utilizado como controle positivo uma amostra de proteína LmTrpRS2 purificada. Esses anticorpos foram liofilizados e enviados ao nosso colaborador, o Prof. Dr. Juan D. Alfonzo – Ohio State University para a realização da localização celular das proteínas de interesse.

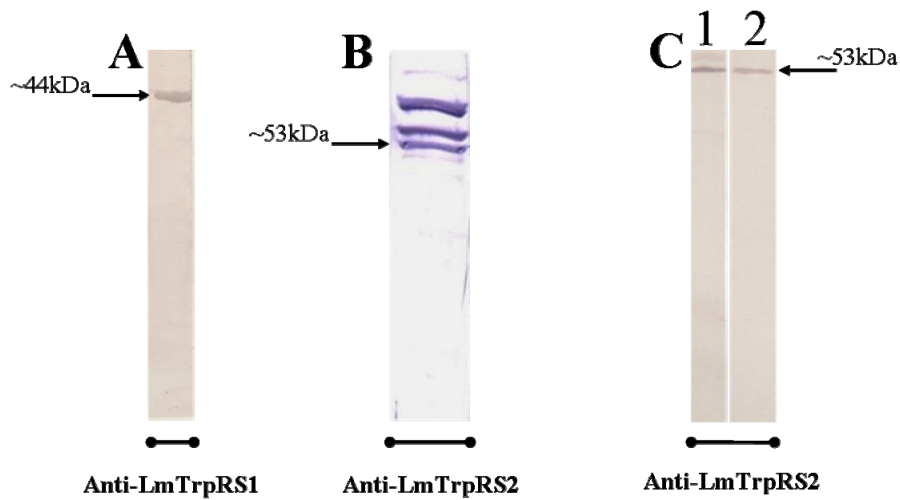


Figura I-4.16: Reação de western-blotting com os anticorpos anti-LmTrpRS1 e Anti-LmTrpRS2 contra as proteínas LmTrpRSs. A) Anti-LmTrpRS1 contra extrato total da proteína recombinante LmTrpRS1, diluição de 1:10000. B) Anti-LmTrpRS2 contra extrato total da proteína recombinante LmTrpRS2, diluição de 1:10000. C) Anti-LmTrpRS2 contra extrato total de células de *L. major* e controle positivo. 1- Controle positivo da reação (diluição de 1:10000), 2 - Extrato total de células de *L. major* (diluição 1:1000).

Localização celular das proteínas LmTrpRS1 e LmTrpRS2.

O endereçamento celular das proteínas LmTrpRS1 e LmTrpRS2 foi realizado em colaboração com o Prof. Dr. Juan D. Alfonzo e seu aluno Frank Ragone (Ohio State University), utilizando-se extratos celulares fracionados de *L. tarentolae* e anticorpos anti-LmTrpRSs produzidos.

A reação para a proteína LmTrpRS1 foi realizada com anticorpos na diluição de 1:500, diluição maior que o utilizado nos extratos bacteriano, porque a concentração da proteína no citoplasma da célula é menor que os valores expressos de forma recombinante. Uma banda fraca com massa molecular de aproximadamente 45 kDa foi observada na fração citoplasmática dos extratos celulares de *Leishmania*, nenhuma reação evidente é observada na fração mitocondrial (figura I-4.17 A). Isso mostra que a LmTrpRS1 é uma enzima com localização citoplasmática.

Os resultados com a proteína LmTrpRS2 mostram uma banda pronunciada no extrato mitocondrial (figura I-4.17 B), essa banda com massa molecular de aproximadamente 55 kDa tem massa molecular compatível com a proteína LmTrpRS2. A fração citoplasmática não apresenta essa mesma banda. Uma pequena banda com massa molecular abaixo de 15 kDa é observada na fração citoplasmática, devendo tratar-se de uma reação cruzada com alguma outra proteína de *Leishmania*.

Os controles positivos mostram que os dois anticorpos estavam reagiram fortemente contra as proteínas em estudo.

Esses fatos confirmam as predições da localização mitocondrial da proteína LmTrpRS2 e citoplasmática da proteína LmTrpRS1.

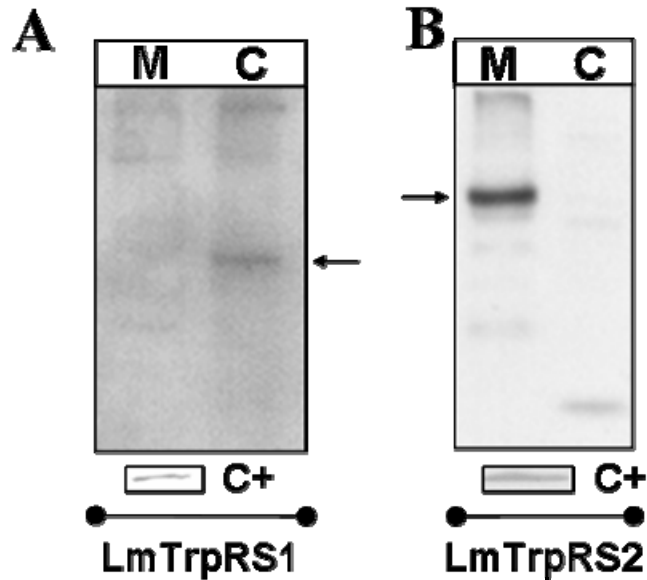


Figura I-4.17: “Western blotting” de extratos celulares de *L. tarentolae* contra os anticorpos anti-LmTrpRSs. **A)** Extrato mitocondrial (M) e citoplasmático (C) de cultura de *L. tarentolae* incubados com anti-LmTrpRS1. **B)** Extrato mitocondrial (M) e citoplasmático (C) de cultura de *L. tarentolae* incubados com anti-LmTrpRS2. As setas indicam as bandas correspondentes as proteínas LmTrpRS1 e LmTrpRS2. C+, controle positivo da reação.

Expressão e Purificação da proteína LmTrpRS2

Na realização dos ensaios enzimáticos com a proteína LmTrpRS2 foi utilizada a construção *LmTrpRS2-24A/pET28a*, que foi a única construção que apresentava uma pequena fração de proteína solúvel. Devido à instabilidade da proteína optou-se por fazer um único passo de purificação em cromatografia de afinidade com resina de níquel. A purificação da proteína LmTrpRS2 foi realizada por cromatografia de afinidade, utilizando-se a resina Ni-NTA Superflow (Qiagen). Para tanto, a fração solúvel foi aplicada manualmente na coluna de afinidade (1,0 mL de resina), previamente equilibrada com 4 volumes de coluna do tampão PBS e 10 mM de imidazol. Em seguida, a coluna foi lavada com 4 volumes do mesmo tampão para remoção das proteínas que se ligaram fracamente à resina. A proteína recombinante foi eluída da coluna com a aplicação de tampão PBS com gradiente crescente de concentração de imidazol. Três concentrações de imidazol, 50, 150 e

250 mM foram aplicadas. As frações eluídas da coluna de níquel foram analisadas em gel de poliácridamida de 15% (figura I-4.18)

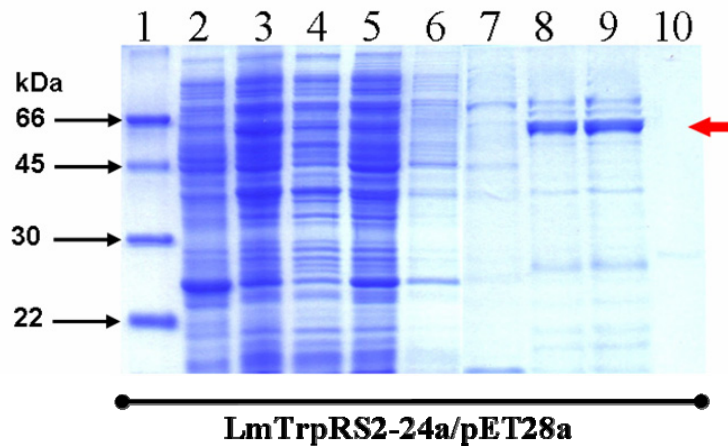


Figura I-4.18: Expressão e purificação da proteína LmTrpRS2 sem a região de sinalização mitocondrial no vetor de expressão pET28a na coluna de cromatografia de afinidade Ni-NTA. 1- marcador de massa molecular; 2- Amostras de cultura celular não induzida; 3 - extrato bruto da lise celular; 4 a 7 - lavagem da coluna com tampão PBS e 10 mM Imidazol; 8 a 10, eluição com 50, 150 e 250 mM de imidazol. A seta em vermelho mostra a proteína recombinante na fração solúvel.

Após extensivas tentativas, uma condição de expressão e purificação parcial da proteína LmTrpRS2 foi obtida. Nesta condição a proteína apresenta-se com uma pequena fração solúvel, um pouco mais estável. Todo o procedimento de purificação foi realizado a temperatura ambiente, permanecendo a proteína em gelo após eluição da resina de afinidade. Mesmo apresentando bandas de contaminação, a fração solúvel dessa purificação pôde ser empregada para os ensaios de aminoacilação para verificar a atividade da enzima contra os tRNAs^{Trp} editado e não editado. Nenhum passo cromatográfico foi adicionado ao processo de purificação, pois a proteína precipita rapidamente após sua eluição da coluna, sendo alguns precipitados podem ser observados após uma hora da eluição.

Ensaio de atividade enzimática da enzima recombinante.

As AaRSs realizam a reação de aminoacilação em duas etapas: Primeiro ATP reage com o aminoácido específico formando uma aminoacil-adenilato e difosfato (figura I-4.19 A). Em uma segunda reação o aminoácido ativado é transferido para a extremidade 3' do tRNA para formar o aminoacil-tRNA (ou tRNA carregado) e AMP (figura I-4.19 B). O monitoramento de qualquer uma dessas etapas pode ser utilizado como ensaio de atividade das AaRSs. Na primeira reação é monitorada a

degradação de ATP, através de utilização de $AT^{32}P$, já na segunda é monitorada a formação do complexo AA-tRNA (aminoácido marcado).



Figura I-4.19: Reação enzimática catalisada pelas AaRSs. **A)** Reação de ativação do aminoácido (AA) utilizando adenosina trifosfato (ATP) e a enzima aminoacil tRNA sintetase (AARS), formando aminoacil adenilato (AARS•AA-AMP) e pirofosfato inorgânico (PP_i). **B)** Aminoacilação do tRNA, aminoacil adenilato e ligado ao tRNA específico produzindo um tRNA carregado AA-tRNA e adenosina monofosfato.

Transcrição *in vitro* do tRNA^{Trp} editado e não editado, purificação e protocolo de re-enovelamento desses tRNAs^{Trp}.

O monitoramento do segundo passo da reação das AaRSs foi utilizado nos experimentos de ensaio enzimático. Um dos substratos desta reação é a molécula de tRNA^{Trp}. Os ensaios enzimático utilizaram tanto as amostras de tRNA^{Trp} editado e não editado transcritos *in vitro* quanto as frações de tRNA total e tRNA de kinetoplasto (ktRNA).

As reações de transcrição *in vitro* foram realizadas a partir da transcrição dos genes tRNA^{Trp} inseridos dentro do vetor pUC19. O cassete de transcrição consiste de um pequeno fragmento de DNA com um sítio para ligação da enzima T7 RNA polimerase viral, seguido do gene do tRNA^{Trp} e na porção final do gene um sítio para endonuclease de restrição *Bst*NI. O DNA plasmidial foi propagado e extraído de cepas de *E. coli* DH5α. Para a formação da extremidade CCA-3' comum a todos o tRNAs o DNA foi digerido com a enzima *Bst*NI por quatro horas a 65° C. Após precipitação desse DNA digerido ele foi utilizado como molde para reação de transcrição, que foi realizada a 37° C por 24 horas.

As reações de transcrição para o tRNA não editado (CCA) mostraram três bandas. As duas superiores, uma com aproximadamente 1000 pb e a outra com 250 pb, são compatíveis com fragmentos de DNA provenientes da digestão do plasmídeo utilizado como molde para a reação de transcrição, uma redução na quantidade de DNA molde manteve o rendimento da transcrição e reduziu a intensidade dessas bandas. Uma banda bem definida com massa molecular compatível com tRNA^{Trp}, 72 pb (figura I-4.20 A). O resultado para a tRNA^{Trp} editado (UCA) mostrou uma banda

única com o tamanho aproximado de 70 pb, nesta reação a quantidade de DNA molde foi reduzida e não foram observadas as bandas de contaminação com massa molecular acima de 250 pb (figura I-4.20 B).

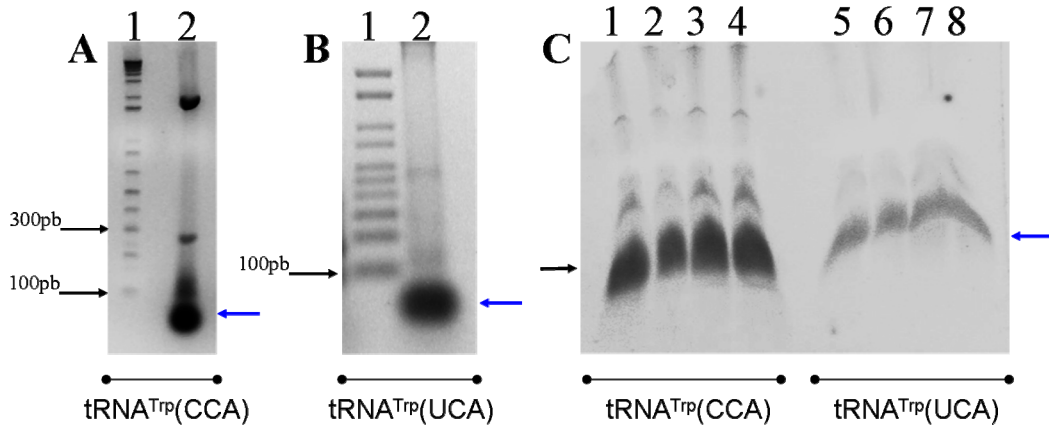


Figura I-4.20: Gel de formaldeído agarose 2% mostrando a transcrição do tRNA^{Trp}. Poço 1 mostra marcador molecular para RNA 0,1-2kb, poço 2 amostra do tRNA^{Trp}. **A)** Transcrição *in vitro* do tRNA^{Trp}(CCA) não editado. **B)** Transcrição *in vitro* do tRNA^{Trp}(UCA) editado. **C)** Purificação dos tRNAs^{Trp} editado (1 a 4) e não editado (5 a 8) em gel de poliácridamida uréia 8M.

Os tRNAs editado e não editado transcritos foram purificados em gel denaturante de poliácridamida 8% e 8 M uréia. Observa-se que a grande parte do transcrito refere-se a bandas únicas, com apenas algumas bandas de contaminação mostrando que a eficiência da reação foi alta. As bandas com pesos compatíveis com os tRNAs^{Trp} foram cuidadosamente recortadas e a eluição das mesmas do gel de poliácridamida foi realizada por incubação em solução de acetato de sódio 0,3 M a 4° C por 12 horas. Os tRNAs foram em seguida precipitados com etanol absoluto. Para o re-enovelamento, os tRNAs purificados foram dissolvidos em água livre de RNase e aquecidos em banho-maria até a temperatura de 75° C por 3 minutos, para assegurar que somente a estrutura primária do mesmo fosse mantida. Após esse tempo o banho-maria foi desligado e os tRNAs foram lentamente resfriados até a temperatura ambiente, quando tiveram sua concentração medida com espectrofotômetro a 260 nm. As amostras foram estocadas a -20° C para evitar degradação do tRNA.

Os ensaios utilizaram os tRNAs^{Trp} editado e não editado transcritos *in vitro* e re-enovelados, amostras de RNA total de kinetoplasto de *L. tarentolae* (kRNA), enviado por nosso colaborador o professor Dr. Juan D. Alfonzo, e também com

fração de RNA total de *L. major* que contém os tRNAs endógenos do parasita. Como controle da reação foi utilizada uma reação de aminoacilação sem a enzima recombinante. Os experimentos de otimização da reação de aminoacilação envolveram a variação nas concentrações de proteína e também dos substratos (ATP, triptofano e tRNA). A melhor condição reacional foi a que utilizava 50 mM Hepes, pH7,0; 10mM acetato de magnésio; 4 mM DTT; 4 mM ATP; 0,05% de albumina bovina (BSA) e uma mistura de 38 μ M de triptofano frio e 2 μ M de L- [5-³H]-triptofano (32 Ci/mmol - Amersham Life Science) em um volume final de 50 μ l.

Os dados obtidos mostram que a fração solúvel da proteína recombinante LmTrpRS2 apresenta atividade enzimática (figura I-4.21), tanto para os tRNAs^{Trp} transcritos *in vitro*, quanto para a amostra de RNA total, que contém tRNAs nativos de *L. major*. As amostras utilizando kRNA total mostraram pouca ou nenhuma atividade de aminoacilação quando comparada as demais amostras, mesmo aumentando-se a concentração total de kRNA não foi possível melhorar essas leituras.

De maneira interessante esses dados mostram que a proteína recombinante LmTrpRS2 expressa é capaz de aminoacilar efetivamente os tRNAs transcritos *in vitro*, sugerindo que alterações pós-transcricionais freqüentemente encontrados nos tRNA podem não ser importantes para o reconhecimento do tRNA pela enzima mitocondrial.

Outro ponto importante que pode ser destacado é a habilidade dessa proteína reconhecer tanto o tRNA editado quanto o não editado, uma vez que dentro da mitocôndria coexistem as duas formas do tRNA e ambas precisam ser aminoaciladas para a síntese protéica. Dados bioquímicos com a TrpRS citoplasmática de *T. brucei* (Charriere *et al*, 2005) mostraram que esta enzima homóloga é incapaz de reconhecer o códon UCA do tRNA^{Trp} editorado. No caso de *T. brucei*, a discriminação dos tRNAs^{Trp} é realizada pela enzima citoplasmática. Nossos dados apontam para a mesma hipótese, uma vez que a enzima LmTrpRS2 também reconhece os dois tRNAs, provavelmente fica a cargo da LmTrpRS1 a correta discriminação desses tRNAs^{Trp}. Todavia um estudo mais pormenorizado deve ser realizado para traçar a relação entre as duas enzimas e seus tRNAs.

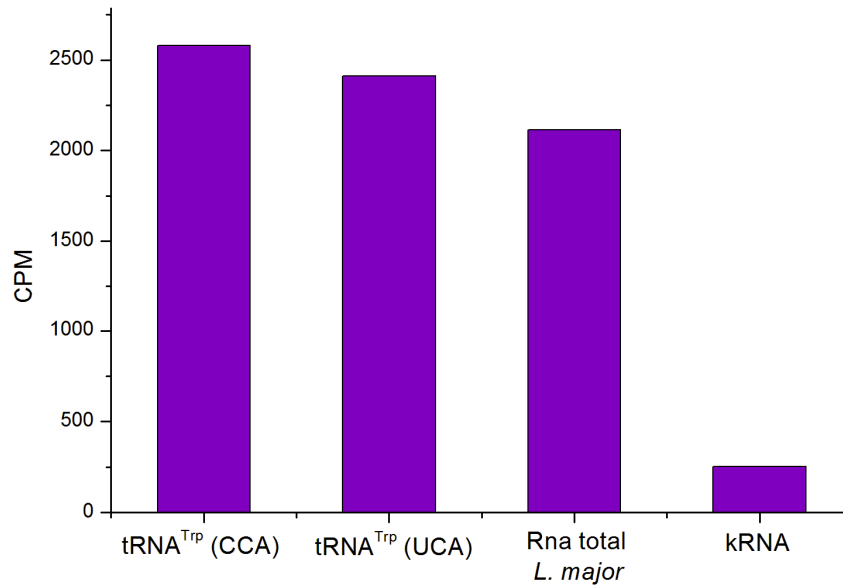


Figura I-4.21: Ensaio de aminoacilação da enzima LmTrpRS2.

Caracterização estrutural das proteínas LmTrpRS1 e LmTrpRS2.

Os avanços em técnicas experimentais e equipamentos permitiram a automatização do seqüenciamento do DNA, acelerando a obtenção de informações codificadas nos diversos projetos genomas realizados (Burley *et al.*, 1999). Transformar essa informação em produtos gênicos que afetam diretamente os sistemas biológicos é um dos grandes desafios atuais. Dentro dos diversos produtos gênicos mais estudados, as proteínas exercem um papel de destaque. Os estudos estruturais, bem como dados bioquímicos fornecem informações importantes para a compreensão da função, mecanismo de ação e ainda para o desenvolvimento de novos fármacos.

As dificuldades experimentais como: obtenção de um sistema de expressão capaz de produzir a proteína de interesse em grande quantidade, purificação dessa amostra na forma solúvel e ativa (Service, 2002), são alguns dos obstáculos para a determinação estrutural de proteínas a partir de técnicas experimentais de difração de raios-X e ressonância magnética nuclear.

A modelagem molecular de proteínas mostra-se como uma importante ferramenta para o estudo de estrutura de proteínas uma vez que não necessita diretamente de dados experimentais. Levando-se em consideração que, similaridade seqüencial pode estar implicada em similaridade estrutural, a predição de modelos estruturais de proteínas a partir de sua seqüência de aminoácidos tem se mostrado

com vários níveis de sucesso uma importante ferramenta para o estudo de vários processos biológicos. Proposições de mutações sítio dirigidas, estudo de relação proteína-proteína, proteína-ligante entre outros podem ser realizados através de modelagem molecular. A limitação deste método é a existência de uma proteína com similaridade seqüencial ou então similaridade em composição de elementos de estrutura secundária (threading protéico) à proteína em estudo, com estrutura tridimensional resolvida.

Modelagem molecular das proteínas LmTrpRS1 e LmTrpRS2.

O primeiro passo para modelagem molecular por homologia foi a busca de uma proteína de estrutura tridimensional conhecida (molécula molde) com elevado grau de similaridade seqüencial com a proteína em estudo (molécula alvo). A seleção das moléculas moldes foi realizada através de alinhamento seqüencial das proteínas LmTrpRS1 e LmTrpRS2 com as seqüências protéicas depositadas no PDB (Westbrook *et al.*, 2002), utilizando-se inicialmente a ferramenta Blastp. As seqüências de aminoácidos das proteínas com maior similaridade foram analisadas uma a uma com o programa Clustal W (Higgins, *et al.*, 1994), sendo então gerados alinhamentos seqüenciais.

Os melhores alinhamentos seqüenciais foram observados para a proteína triptofanil tRNA sintetase humana (HsTrpRSs) que apresentou 55% de identidade e 71% similaridade com a proteína LmTrpRS1 e 37% de identidade e 55% similaridade com a proteína LmTrpRS2 (figura I-4.22). O alinhamento entre as duas enzimas mostra 34% de identidade e 48% de similaridade (figura I-4.23) O alinhamento da LmTrpRS2 menos conservado deve ser considerado devido a filogenia dessa proteína, sendo de origem mitocondrial, como previamente demonstrado ela apresenta-se mais distante das TrpRSs de eucariotos, algumas inserções podem ser facilmente observadas no interior da seqüência. A seqüência de aminoácidos é conservada principalmente ao longo dos motivos “HIGH” e “KMSKS”, devido à importância desses na ativação do aminoácido. Embora o alinhamento N-terminal da proteína mitocondrial seja baixo, ele se encontra no limite para os experimentos de modelagem. A proteína mitocondrial foi modelada na sua forma madura, ou seja, sem a região de sinalização mitocondrial, uma vez que essa é

clivada após o processo de importação. Cinco estruturas de HsTrpRSs foram encontradas no PDB. Uma análise criteriosa das estruturas mostrou que a cadeia B da estrutura da HsTrpRSs, código de acesso PDB 1RT6, apresentava o maior número de aminoácidos com densidade eletrônica (Yang *et al.*, 2003), tornando-se o melhor molde para a modelagem por homologia. Quanto ao modelo da HsTrpRSs também foi realizada a substituição dos resíduos de selenometionina, empregados para o faseamento experimental da proteína, por resíduos de metionina, sem modificação das coordenadas do PDB de acesso 1R6T. As coordenadas das moléculas de água e ligantes também foram removidas.

Algumas modificações foram introduzidas manualmente no alinhamento seqüencial para adaptá-lo a molécula molde, como a remoção dos 6 resíduos iniciais da proteína bem como os quatro resíduos do C-terminal, e também a região compreendendo a Ala61 até o Glu82. Esses resíduos não foram construídos na proteína humana devido à ausência de densidade eletrônica sendo então eliminados do alinhamento seqüencial.

A modelagem molecular foi realizada utilizando-se o programa MODELLER 6a (Sali e Blundell, 1993), utilizando-se como entradas do programa os alinhamentos entre a seqüência de aminoácidos das LmTrpRS e a estrutura utilizada como molde (coordenadas PDB da estrutura HuTrpRS1).

Foram gerados inicialmente 20 modelos para cada proteína. O critério inicial para escolha dos modelos foi a seleção dos modelos com a menor pseudo energia fornecida pelo próprio programa, os quatro modelos foram selecionados para validação com o programa Procheck e Whatif e análise com o programa de visualização gráfica PyMol (Delano, 2002) e Coot (Emsley e Cowtan, 2004).

A validação foi realizada para verificar se os modelos apresentam corretos ou aceitáveis valores de estereoquímica como: comprimentos das ligações, ângulos das ligações, planaridade das ligações peptídicas e dos anéis das cadeias laterais, quiralidade, ângulos torcionais da cadeia principal e das cadeias laterais e impedimentos estéricos entre pares de átomos não ligados.

Com base na validação e na visualização da estrutura da LmTrpRS1 os seis primeiros resíduos do modelo assim como os 11 resíduos do C-terminal foram removidos, pois não apresentavam bons resultados na validação, isso é comum pois essas regiões são muito flexíveis, no caso desta modelagem ainda a homologia dessa regiões é muito baixa, o que dificulta o processo. O diagrama Ramachandran (figura I-4.24) mostra que 94,8% dos resíduos encontram-se em regiões favoráveis e 5,2% em regiões adicionais, caracterizando como um bom modelo.

O diagrama de Ramachandran do modelo obtido para a proteína LmTrpRS2, mostra que 92,0% dos resíduos encontram-se em regiões favoráveis, 6,1% em regiões adicionais, 1% em regiões generosamente permitidas e 1% em regiões desfavoráveis. A análise visual da estrutura mostrou que os resíduos A3, R61, D57 e R295 que estão em regiões desfavoráveis apresentam-se coerentes com sua vizinhança atômica, como mostrado pelo programa WhatIf e com ambiente químico favorável como mostrado pelo programa Verify 3D. Isso mostra que esses resíduos mesmo em regiões desfavoráveis no diagrama de Ramachandran, não apresentam desvios graves quanto a sua vizinhança e ao ambiente químicos. Outro fato a ser apontado é que tanto esses quatro resíduos, quanto os seis mostrados em regiões adicionais do diagrama de Ramachandran, encontram-se em regiões de baixa identidade seqüencial e em “loops”, que são regiões muito flexíveis, com grande dificuldade para a modelagem estrutural. Ainda assim podemos considerar este como um bom modelo. Tentativas de modelagem com diferentes alinhamentos não conseguiram solucionar os problemas encontrados com esses aminoácidos.

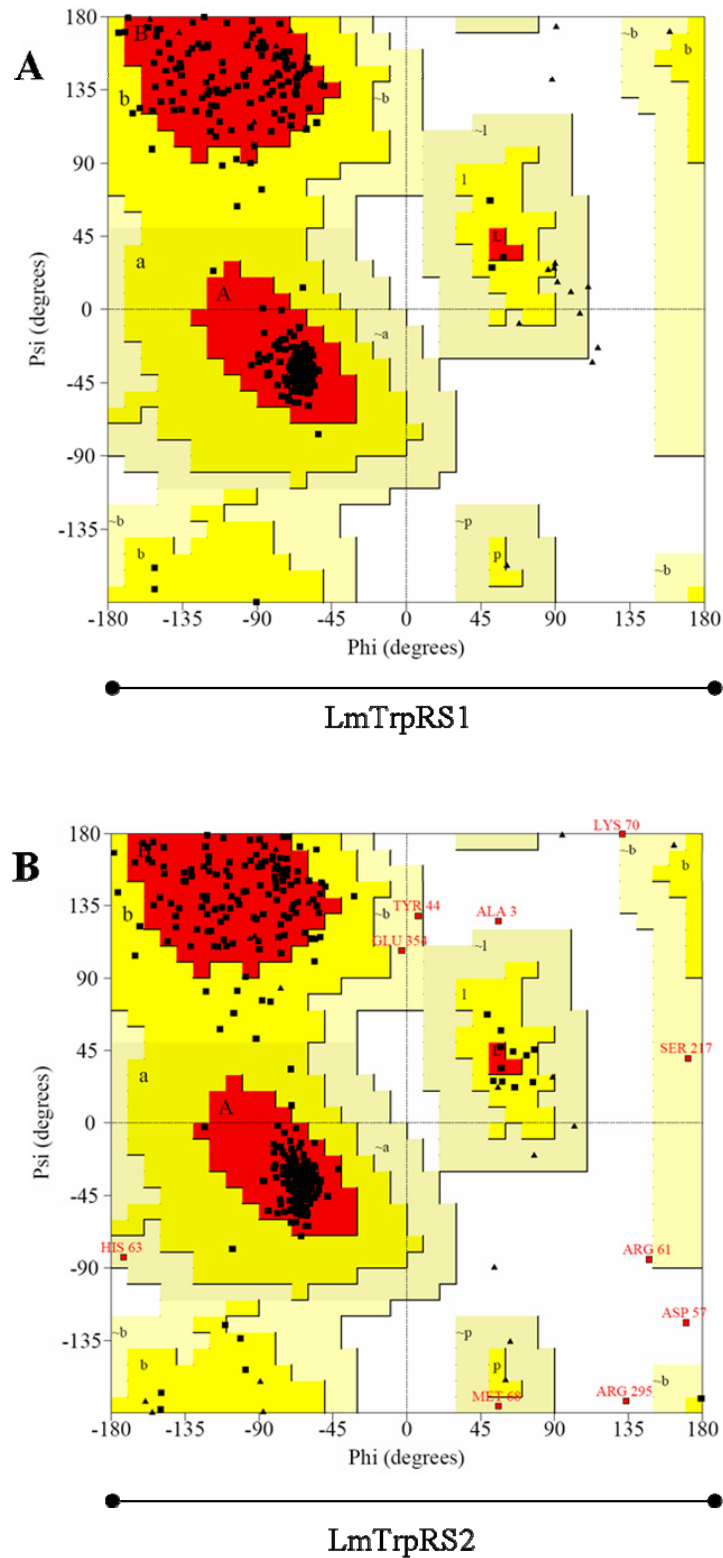


Figura I-4.24: Diagrama de Ramachandran dos modelos finais obtido da proteína LmTrpRS1 (A) e LmTrpRS2 (B). Vermelho mostra as regiões mais favoráveis dos aminoácidos, amarelo mostra as regiões adicionais e em bege as regiões favoravelmente permitidas. Os aminoácidos são representados como quadrados, os triângulos mostram o aminoácido glicina.

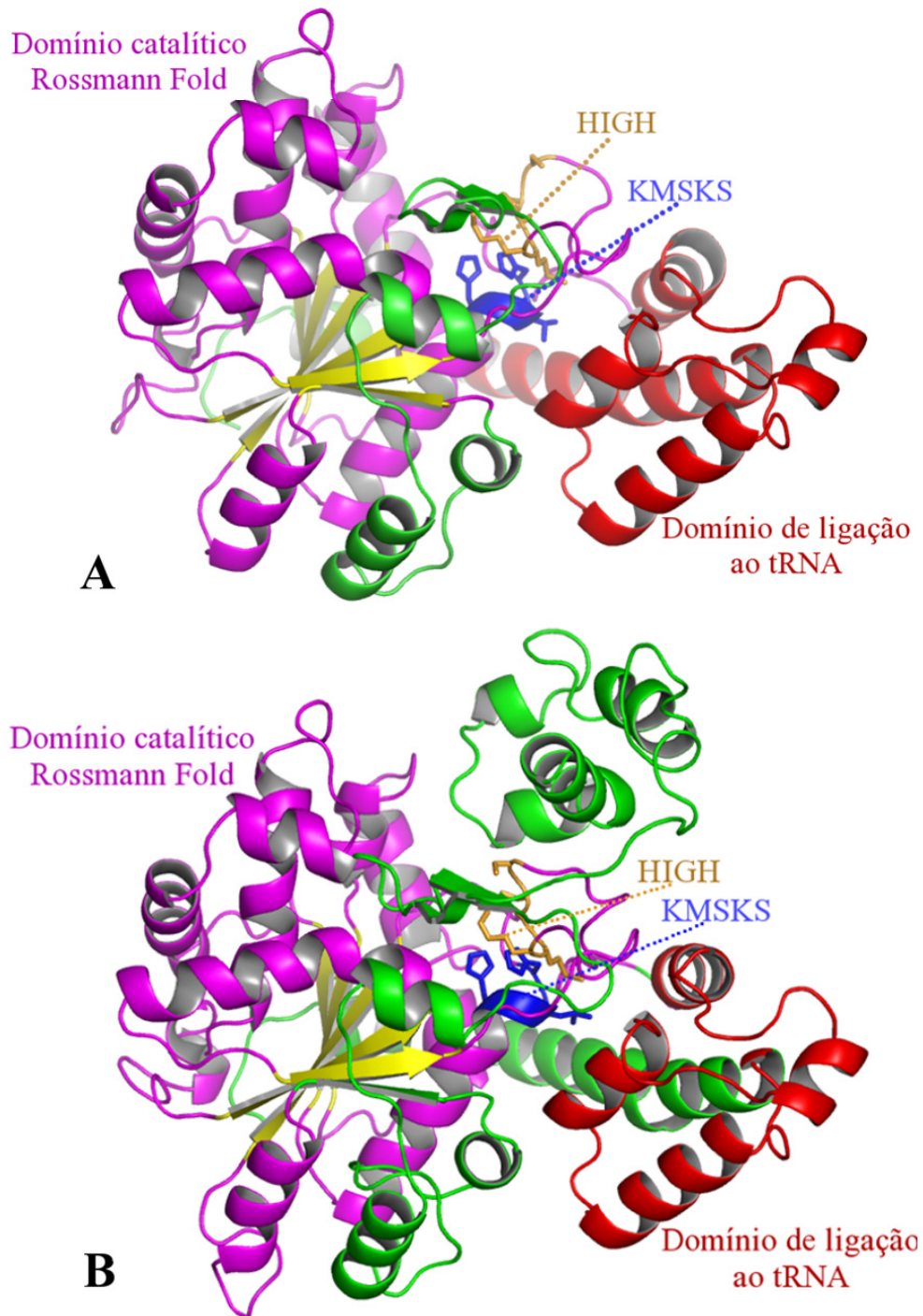


Figura I-4.25: Representação em cartoons dos modelos obtidos para as proteínas LmTrpRS1 e LmTrpRS2. A) Modelo da LmTrpRS1. B) Modelo da LmTrpRS2. Em magenta é destacado o domínio catalítico, responsável pela ativação do triptofano e a aminoacilação. O “Rossmann Fold” característico dessa classe de enzima é observado e acima desse as seqüências consenso HIGH (azul) e o “loop” KMSKS (dourado). O domínio de ligação ao anticódon do tRNA é mostrado em vermelho.

Os modelos finais obtidos das LmTrpRS1 e LmTrpRS2 apresentam dois domínios distintos, um domínio maior com aproximadamente 200 aminoácidos, no qual se encontra o motivo estrutural clássico dessa classe de AaRSs, que é o

“Rossmann fold” (Branden e Tooze, 1998) e outro domínio menor composto basicamente por hélices- α (figura I-4.25).

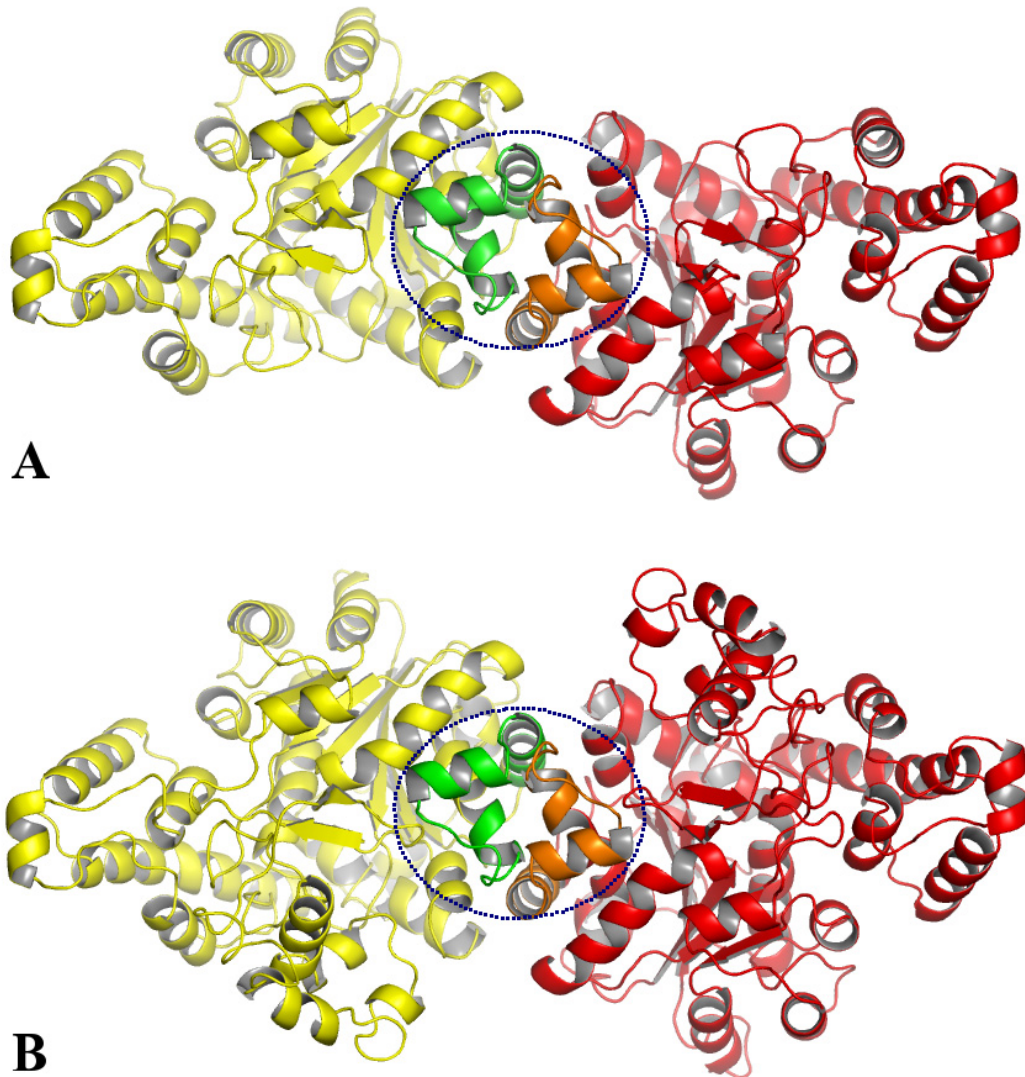


Figura I-4.26: Representação em cartoons das LmTrpRSs na forma de dímeros. A e B) Modelos diméricos da proteínas LmTrpRS1 e LmTrpRS2 respectivamente. A região destacada mostra a superfície de dimerização.

As TrpRSs são formadas por estruturas relativamente pequenas, aproximadamente 340 resíduos, quando comparados as demais AaRSs, e ao contrário da maioria das AaRSs de classe são ativas na forma de homodímeros biológicos, formando juntamente com a TyrRSs uma subclasse aparte denominada subclasse IC (Cusack, 1995). Estruturas cristalográficas mostraram que o reconhecimento e a ativação de uma molécula de tRNA é intermediada por sítios catalíticos distintos um em cada monômero. Assim enquanto a ativação do aminoácido é feita pela porção que contém o “Rossmann fold” de um monômero, o reconhecimento do anticódon é realizado pelo domínio de hélices- α do outro monômero (Yaremchuk *et al.*, 2002;

Kobayashi *et al.*, 2003, Yang, *et al.*, 2006). Desta maneira um dímero é capaz de reconhecer e ativar duas moléculas de tRNA simultaneamente.

A obtenção da estrutura dimérica das proteínas LmTrpRS foi gerada a partir da sobreposição do monômero modelado sobre a estrutura cristalográfica do homólogo humano (Yang *et al.*, 2006). A interface de dimerização entre os dois monômeros das LmTrpRS é composta por três hélices- α , provenientes de cada monômero, situadas opostas ao domínio de ligação ao anticódon (figura I-4.26).

O sítio de aminoacilação é formado por resíduos conservados na maioria da TrpRS, ele forma um grande sulco exposto ao solvente exatamente na porção superior do domínio “Rossmann fold” e pode ser dividido em dois sub-sítios: um com resíduos específicos para a ligação do ATP e outro para ligação do triptofano. Na porção lateral desse sulco estão as seqüências conservadas “HIGH” e “KMSKS”, o segundo ocupa uma posição mais exposta ao solvente (figura I-4.25). Essas duas seqüências são trazidas próximas uma a outra pelo enovelamento do “Rossmann fold”.

Uma característica interessante observada no modelo da LmTrpRS2 é a presença de duas hélices- α no N-terminal da proteína ($\alpha 1$ e $\alpha 2$). Essas apresentam em uma posição quase perpendicular em relação à proteína molde (figura I-4.27 A). Um resíduo do aminoácido prolina é observado na seqüência da primeira hélice (P37), o que pode levar a uma distorção da mesma durante a modelagem. Essa seqüência N-terminal parece ser comum nas TrpRS de eucariotos. Na estrutura da proteína humana ele assume a configuração de uma hélice-volta-hélice, sendo observada somente em um monômero de uma das oito estruturas resolvidas, mostrando grande flexibilidade da mesma. Possivelmente a ligação do substrato não hidrolisável Trp-ATP (TYM), presente na estrutura cristalográfica, está envolvida na estabilização desse domínio (Yang *et al.*, 2003). Estudos realizados com mutantes para essa região não mostraram perda de atividade da enzima quando da sua remoção, sugerindo um papel secundário para essa estrutura hélice-volta-hélice (Otani, *et al.*, 2002; Wakasugi, *et al.*, 2002). A sobreposição do modelo LmTrpRS2 com as estruturas humanas resolvidas mostram que um grande impedimento estereoquímico ocorre entre o N-terminal do modelo da LmTrpRS2 e a molécula de tRNA (figura I-4.27 B), embora na estrutura humana complexada com tRNA os 80 resíduos da região N-terminal da proteína humana encontram-se completamente

ausentes, na estrutura da humana utilizada como molde para modelagem essa região estava presente. Embora essa região do modelo apresentasse bons resultados quanto à validação não se pode inferir muito sobre sua correta orientação, uma vez que o alinhamento seqüencial apresenta-se pobre para essa região. Supondo-se que a mesma conformação do tRNA seja conservada para as LmTrpRS2 um movimento dessa região se faz necessário para evitar que a molécula do tRNA choque-se com essa região N-terminal (figura I-4.27 B).

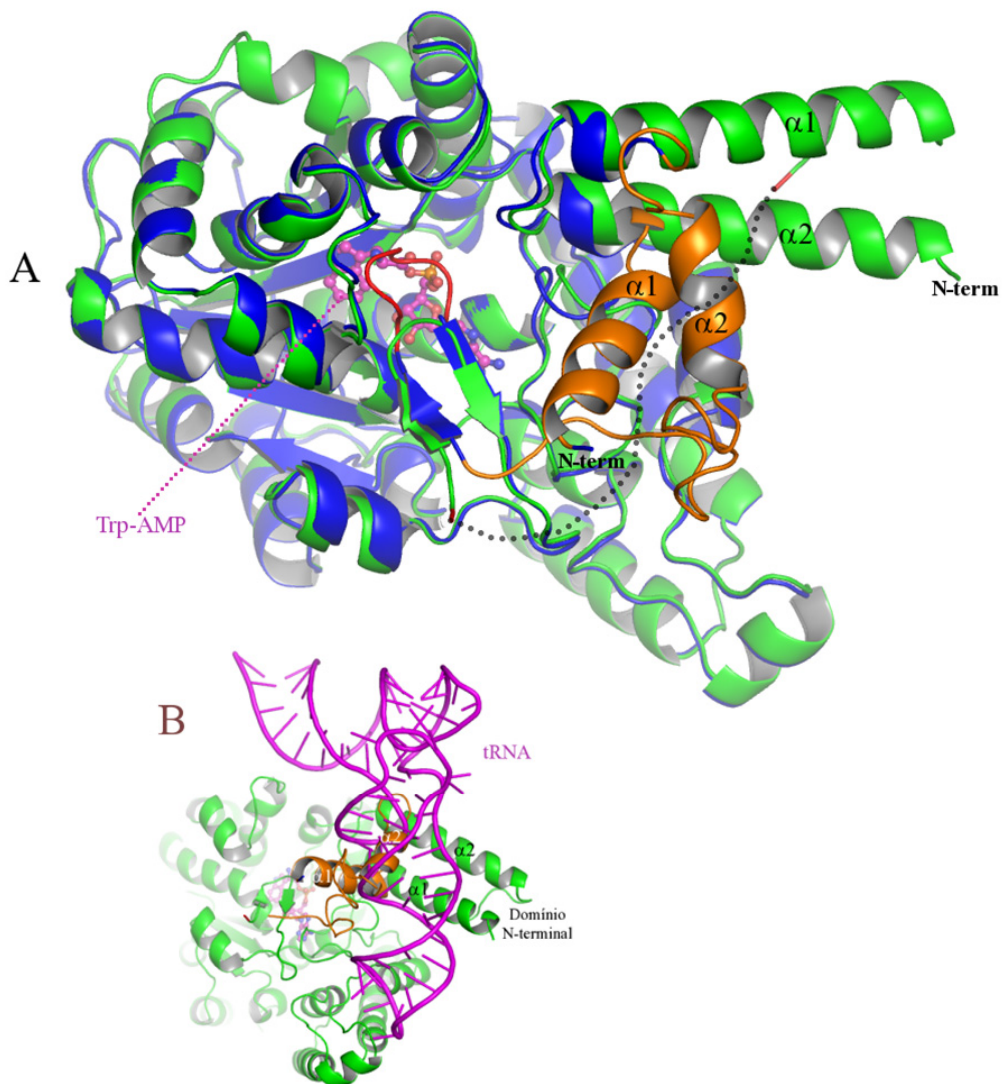


Figura I-4.27: Representação de cartoons da sobreposição da estrutura da HsTrpRS1 com a estrutura modelada da LmTrpRS2. A) HsTrpRS ligada ao substrato não hidrolisável Trp-AMP sobreposta a estrutura LmTrpRS2, mostrando diferença entre as hélices $\alpha 1$ e $\alpha 2$ na região N-terminal da duas estruturas. A linha pontilhada representa a conexão da região perdida (sem densidade eletrônica) na estrutura da proteína humana. Em vermelho mostra o loop acima do sítio catalítico. B) Sobreposição da estrutura HsTrpRS1 complexada com tRNA^{Trp} e do modelo da proteína LmTrpRS2. Região N-terminal da LmTrpRS2 (alaranjado) colidindo com molécula de tRNA.

Outra região interessante observada na LmTrpRS2 é um “loop” presente em um motivo estrutural do tipo “grampo de cabelo” que ocupa a região de entrada do sítio ativo (figura I-4.27 A). Tanto na estrutura humana quanto na LmTrpRS1 esse “loop” compreende quatro aminoácidos, já na estrutura da LmTrpRS2 ele compreende nove resíduos. Como essa é uma região bastante flexível pode-se especular que este auxilie na estabilização dos substratos no interior do sítio ativo.

As estruturas cristalográficas da proteína humana mostram que uma complexa rede de ligações de hidrogênio estabilizam o triptofano dentro do sítio ativo (Shen *et al.*, 2006; Yang *et al.*, 2006).

Na estrutura da proteína humana o reconhecimento do Trp é realizado por ligações de hidrogênio entre quatro resíduos de aminoácidos (Q194, Y159, E199 e Q284). Esses quatro aminoácidos são conservados nas seqüências das duas proteínas de *L. major* e os modelos apontam que eles ocupam a mesma configuração da proteína humana, sendo importantes para o reconhecimento do aminoácido triptofano.

Os resíduos das seqüências “HIGH” e “KMSKS” desempenham importante papel na ativação do aminoácido. A glicina e a histidina da seqüência “HIGH” fazem interações hidrofóbicas com o anel da ribose do ATP. Dois outros resíduos, D312 e K349, da estrutura humana também são conservados nas estruturas das LmTrpRS. Esses dois resíduos são responsáveis pela ligação do ATP e a remoção do pirofosfato produzido durante a ativação do aminoácido (Shen *et al.*, 2006). Um estudo da interação entre os modelos obtidos e um substrato não hidrolisável Trp-AMP foi realizado através de “docking” com o programa GOLD (Jones, 1999) mostrando que essas interações são mantidas nas estruturas modeladas.

A sobreposição dos modelos estruturais das LmTrpRS contra as estruturas cristalográficas da proteína humana ligada ao tRNA mostra que o reconhecimento do anticódon do tRNA está localizada entre as hélices $\alpha 16$, $\alpha 17$ e $\alpha 20$ (numeração idêntica à utilizada na proteína humana) essa região forma um sulco apto ao encaixe do anticódon do tRNA^{Trp} (figura I-4.28 A). Uma característica interessante da estrutura cristalográfica humana é que o nucleotídeo C34 (equivalente ao nucleotídeo editorado em *L. major*) encontra-se estendido para fora da molécula do tRNA, perdendo assim o empilhamento clássico observado nas estruturas de nucleotídeos.

Essa conformação estendida permite a interação do nucleotídeo com os resíduos da cadeia principal da proteína (figura I-4.28 A).

Uma grande identidade seqüencial é encontrada nos alinhamentos das enzimas citoplasmáticas e mitocondriais nessa região de *L. major* e *T. brucei*, principalmente os resíduos que interagem com o anticódon (figura I-4.28 B).

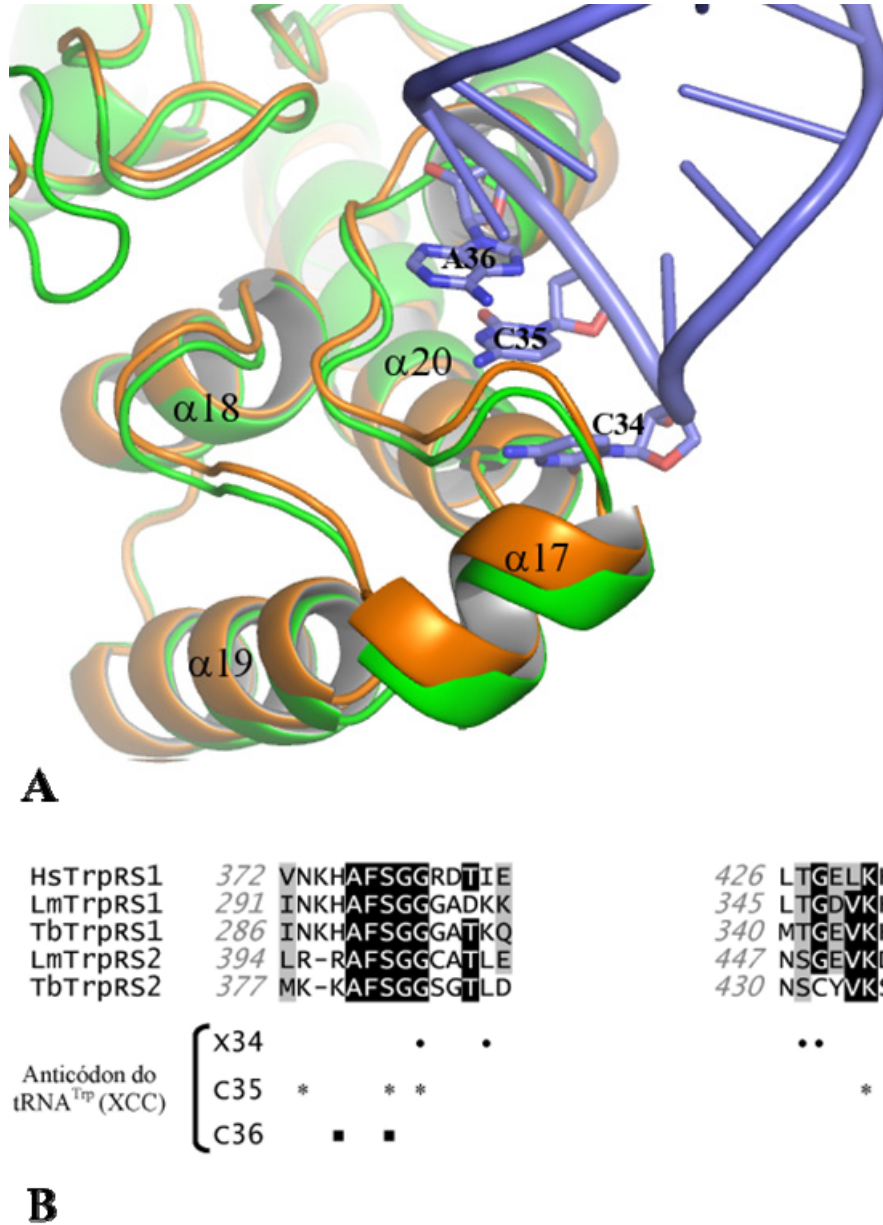


Figura I-4.28: Representação em “cartoons” da sobreposição da estrutura cristalográfica da proteína HsTrpRS1 (alaranjado) complexada com tRNA^{Trp} (azul) e da proteína LmTrpRS1 modelada (verde). **A)** Em detalhe o sulco de reconhecimento do anticódon, formado pelas hélices $\alpha 16$, $\alpha 17$ e $\alpha 20$. **B)** Alinhamento das seqüências de aminoácidos do domínio de ligação ao anticódon, baseado no complexo da proteína HuTrpRS-trp. Triptofanil tRNA sintetases citoplasmática humana (HsTrpRS1), *L. major* (LmTrpRS1), *T. brucei* (TbTrpRS1) e as homólogas mitocondriais de *L. major* (LmTrpRS2) e *T. brucei* (TbTrpRS2). Abaixo da seqüência são destacados os possíveis locais de interação com os nucleotídeos N34(●), N35(*) e N36(■).

As estruturas cristalográficas da proteína humana apontaram uma característica interessante no mecanismo de reconhecimento das bases do anticódon do tRNA^{Trp} que também pode ser aplicado para a estrutura modelada. Na estrutura humana o reconhecimento das bases é feito em grande parte por interações diretas com a cadeia principal da proteína, isso confere uma maior fidelidade à reação de aminoacilação mesmo em casos de mutações. Apenas dois resíduos fazem interações via cadeia lateral na estrutura humana, a S378 e K431, interagindo com o anel pirimidínico da base C35 do tRNA. Esses resíduos são conservados em *L. major*, *T. brucei* e *Homo sapiens*, mostrando que eles podem ter importância no reconhecimento do anticódon do tRNA.

Nenhuma característica especial pode ser observada com relação às interações entre as LmTrpRS1 e LmTrpRS2 e o anticodon do tRNA^{Trp} que sejam responsáveis pelo seu reconhecimento específico do tRNA^{Trp} (UCA), uma vez que a diferença entre os nucleotídeos resume-se ao átomo de hidrogênio no nitrogênio N3 do anel pirimidina e do grupo amino no nitrogênio N4 da uracila. Charriere *et al.*, 2006 mostraram que além da U34 uma tiouridina 33 (S²U33) também é essencial para a discriminação dos corretos tRNAs.

A interação entre o braço acceptor do tRNA^{Trp} e as LmTrpRS é realizada por elementos presente na superfície de dimerização. Como na estrutura humana, a hélice central desse domínio (Shen *et al.*, 2006) interage com os fosfatos da cadeia principal do tRNA, de maneira não específica. A presença do ribonucleotídeo A73 no braço acceptor do tRNA^{Trp} na maioria dos eucariotos e archaea é o maior elemento de identidade entre estes e as bactérias que possuem uma guanina nessa posição (Xue *et al.*, 1993; Xu *et al.*, 2001; Guo *et al.*, 2002). Trabalhos de mutagênese sítio dirigida mostram que em *Bacillus subtilis* uma lisina (K149) e um ácido glutâmico (E153) são importantes para esse reconhecimento espécie-específico, sendo o segundo extremamente conservado entre as bactérias (Jia *et al.*, 2004). No caso de eucariotos e archaea, uma arginina ocupa a posição do E153. Uma das explicações para essa especificidade advém da estrutura cristalográfica da proteína HsTrpRS complexada com tRNA, uma intrincada rede de ligações de hidrogênio estabilizam a R318, que por sua vez faz ligação de hidrogênio com o nitrogênio N1 do ribonucleotídeo A73 do tRNA^{Trp}. Essa mesma configuração pode ser observada na estrutura da LmTrpRS1 (figura I-4.29), porém de uma maneira interessante a mesma posição na

enzima citoplasmática é ocupada por uma valina (V339). Isso pode diminuir a especificidade dessa enzima para o tRNA, e talvez isso seja um indicativo da evolução dessa enzima para aceitar o tRNA^{Trp} eucariótico ou ainda sugerindo que o mecanismo de reconhecimento dessa base é diferente.

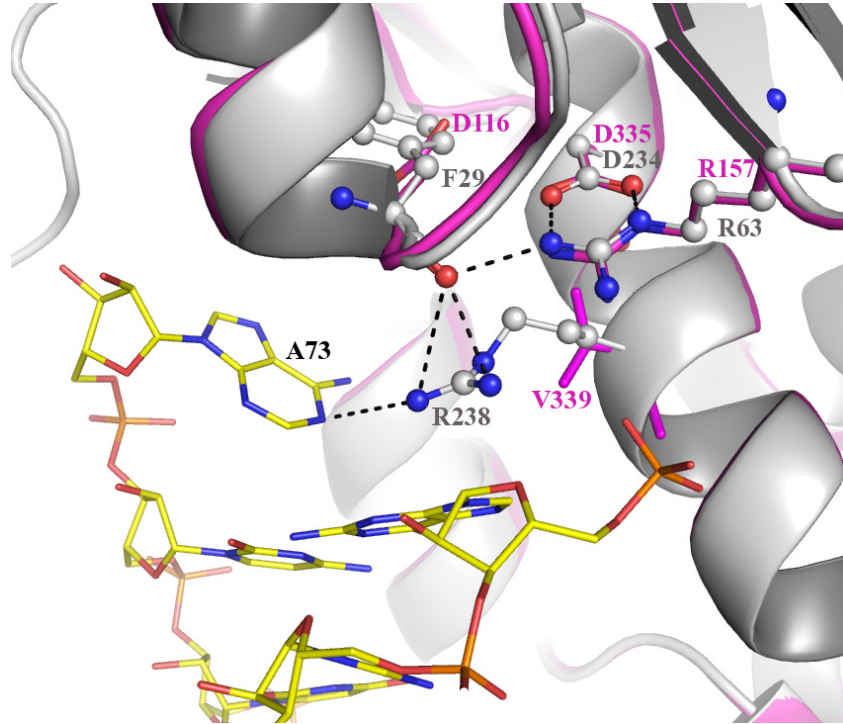


Figura I-4.29: Representação em cartoon da região de interação das proteínas LmTrpRS1 e LmTrpRS2. Em cinza proteína LmTrpRS1, elemento de identidade A73 está interagindo com a R238. Uma intensa rede de ligações de hidrogênio mantém a R238 orientada para a ligação ao N1 da base A73. A estrutura de LmTrpRS2 apresenta uma valina nessa posição.

Alinhamentos seqüenciais mostram que essa mesma valina é conservada nas enzimas mitocondriais de *T. brucei* e *T. cruzi*, corroborando a hipótese do seu envolvimento no reconhecimento do braço aceptor do tRNA^{Trp}.

Considerações finais.

O reconhecimento do aminoácido e seu respectivo tRNA é uma etapa fundamental para assegurar a fidelidade do código genético. Essa complexa tarefa envolve o reconhecimento específico do aminoácido e de seu tRNA cognato. Um grupo especial de enzimas denominadas aminoacil tRNA sintetases é responsável por esse trabalho. As AaRSs atuam inicialmente ativando o aminoácido através da ligação do mesmo a uma molécula de adenosina trifosfato (ATP), formando um

amino acil adenilato. Em um segundo passo esse aminoácido ativado é transferido para a ribose da adenina presente na extremidade CCA-3' do tRNA gerando o “tRNA carregado” que será utilizado na tradução protéica.

Em grande parte dos organismos vivos um repertório de 20 AaRSs é responsável pela formação de todos os tRNAs carregados necessários para a síntese protéica. A simples existência dos aminoácidos Pirolisina e Selenocisteína, que não possuem AaRSs descritas até o presente momento, implica necessariamente na bi funcionalidade de algumas dessas enzimas. Esse fato é complementado com a análise dos genomas das arqueobactérias *Methanococcus jannaschii* e *Methanobacterium thermoautotrophicum*, que possuem somente 16 AARSs. Duas ausências, asparagina tRNA sintetase e glutamina tRNA sintetase, podem ser prontamente substituídas com uma via alternativa de síntese de amino acil tRNA, partindo dos precursores aspártico-tRNA e glutâmico-tRNA. Porém para as demais outros mecanismos devem estar envolvidos. Dados estruturais e bioquímicos mostram que alguma flexibilidade dentro do reconhecimento enzima-aminoácido-tRNA^{aminoácido} foi a maneira encontrada pela natureza para corretamente prover todos os tRNAs necessários à síntese protéica. Nesse contexto, os estudos bioquímicos e estruturais dessas enzimas são de fundamental importância para a compreensão desses mecanismos, bem como a sua utilização para o desenvolvimento de novas terapias contra organismos patogênicos.

Uma das características interessantes observadas em kinetoplastidas é que nesses organismos o código genético mitocondrial é desviado do padrão canônico pela utilização do códon de terminação UGA para a inserção do aminoácido triptofano. Interessantemente nesses organismos apenas um tRNA^{trp} (CCA) nuclear decodificando o códon triptofano UGG foi descrito. Um mecanismo de editoração da base C₃₄ para U₃₄ convertendo o tRNA^{Trp}(CCA) em tRNA^{Trp}(UCA) foi identificado recentemente, permitindo a decodificação desse códon de terminação para o aminoácido triptofano. Uma questão que surge ao analisar esse mecanismo é a maneira pela qual a TrpRS reconhece o tRNA editado e não editado. A resposta a essa questão poderia ser respondida através da análise do genoma do parasito, aonde duas enzimas TrpRSs são codificadas. Nesse contexto se enquadram nos objetivos deste projeto que envolvem a elucidação do mecanismo catalítico e do mecanismo discriminatório entre tRNA^{trp} CCA e UCA.

A análise dos genes depositados no GeneDB caracteriza as proteínas codificadas por ambos com probabilidade de localização mitocondrial, fato não comum quando comparado aos demais kinetoplastidas. A caracterização molecular desses genes mostrou que a predição da janela aberta de leitura depositada é 252 nucleotídeos maior que o gene transcrito. Esse novo mRNA mostra que essa proteína expressa não apresenta probabilidade de exportação mitocondrial. Os ensaios de imunoblotting com anticorpos policlonais produzidos com proteína recombinante, contra extratos celulares de *L. tarentolae*, confirmaram as localizações das duas enzimas, uma com a localização citoplasmática (LmTrpRS1) e outra mitocondrial (LmTrpRS2).

Os dois genes foram subclonados em diversos vetores de expressão procarióticos e alguns eucarióticos. A grande maioria dessas construções produziu proteína recombinante, porém na maioria das tentativas na forma de corpos de inclusão. Uma construção da proteína mitocondrial foi capaz de expressar uma pequena quantidade de proteína solúvel, bastante instável, o que impossibilitou os ensaios de cristalização. Os ensaios de atividade mostram que a enzima é capaz de reconhecer os dois tRNAs. Essa característica é esperada, uma vez que no interior da mitocôndria co-existem as duas formas do tRNA^{Trp}. Um trabalho recente demonstrou em *T. brucei* que a discriminação é feita na verdade pela enzima citoplasmática, que nesse organismo foi incapaz de reconhecer efetivamente o tRNA editado. Provavelmente esse mesmo resultado pode ser aplicado à proteína citoplasmática em *L. major* que possui identidade alta com a homóloga em *T. brucei*.

Os modelos de homologia das duas enzimas não mostram alterações que poderiam justificar a escolha do tRNA apenas pela base editorada. Provavelmente outros mecanismos de reconhecimento enzima-tRNA devem estar presentes. Marcadamente modificações pós traducionais da molécula de tRNA poderiam inserir novos elementos que garantam o reconhecimento específico do tRNA. No caso de *T. brucei* uma tiouridina atua juntamente com a editoração U34 como elementos anti-determinantes para o reconhecimento da enzima citoplasmática.

Duas características dos modelos chamam a atenção na LmTrpRS2. A primeira é a posição da porção N-terminal, que ocupa uma região onde a molécula de tRNA estaria localizada, quando comparado ao modelo da proteína humana complexada com tRNA^{Trp}. Os dados de modelagem sugerem que um rearranjo dessa

região ou ainda outro mecanismo de ligação à molécula de tRNA podem ser cruciais no reconhecimento do complexo enzima-tRNA. Outra característica interessante observada no modelo da LmTrpRS2 é a substituição de uma arginina conservada por um resíduo de valina. Essa arginina tem sido descrita como fundamental para o reconhecimento do ribonucleotídeo A73 em proteínas de procariotos. A mesma substituição é observada nas seqüências das homólogas em *T. brucei* e *T. cruzi*, indicando que o mecanismo de interação da proteína com a porção final CCA-3' do tRNA pode ser diferente.

Referências do capítulo.

- Branden, C. e Tooze, J. (1998). *Introduction to protein structure*. New York: Garland Publishing, Inc.
- Braun, H. P. e Schmitz, U. K. (1995). *Trends Biochem. Sci.*, **20**: 171-175.
- Breitling, R.; Klingner, S.; Callewaert, N.; Pietrucha, R.; Geyer, A.; Ehrlich, G.; Hartung, R.; Müller, A.; Contreras, R.; Beverley, S. e Alexandrov, K. (2002). *Prot. Expr. Purific.*, **25**: 209-218
- Burley, S. K.; Almo, S. C.; Bonanno, J. B.; Capel, M.; Chance, M.R.; Gaasterland, T.; Lin D.; Sali, A.; Studier, F. W. e Swaminathan, S. (1999). *Nat. Genet.*, **23**: 151-157.
- Chang, A. C. e Cohen, S. N. (1978). *J. Bacteriol.*, **134**: 1141-1156
- Chen, J.; Li, Y.; Wang, E. e Wang, Y. (1999). *Protein Expr Purif.*, **15**: 115-120.
- Coburn, C. M.; Otteman, K. M.; Mcneely, T.; Turco, S. J. e Beverley, S. M. (1991). *Mol. Biochem. Parasitol.*, **46**: 169–180.
- Cusack, S. (1995). *Nat. Struct. Biol.*, **10**: 824-831.
- Emsley, P., e Cowtan, K. (2004). *Acta Crystallogr. D Biol. Crystallogr.* **60**: 2126-2132.
- Felter, S.; Diatewa, M.; Schneider, C. e Stahl, A. J. (1981). *Biochem. Biophys. Res. Commun.*, **98**: 727-734.
- Guo, Q.; Gong, Q.; Tong, K. L.; Vestergaard, B.; Costa, A.; Desgres, J.; Wong, M.; Grosjean, H.; Zhu, G.; Wong, J. T. e Xue, H (2002). *J. Biol. Chem.*, **277**: 14343–14349

- Hammarstrom, M.; Hellgren, N.; Van Den Berg, S.; Berglund, H. e Hard, T. (2002). *Protein Sci.*, **11**: 313–321.
- Higgins, D.; Thompson, J.; Gibson, T.; Thompson, J. D.; Higgins, D. G. e Gibson, T. J. (1994). *Nucleic Acids Res.*, **22**: 4673-4680.
- Jia, J.; Chen, X. L.; Guo, L. T.; Yu, Y. D.; Ding, J. P. and Jin, Y.X. (2004). *J. Biol. Chem.*, **279**: 41960-41965.
- Jones, D.T. (1999). *J. Mol. Biol.*, **292**: 195-202.
- Kelly, J. M.; Ward, H. M.; Miles, M. A. e Kendall, G. (1992). *Nucleic Acids Res.*, **20**: 3963–3969.
- Kobayashi, T.; Nureki, O.; Ishitani, R.; Yaremchuk, A.; Tukalo, M.; Cusack, S.; Sakamoto, K. e Yokoyama, S. (2003). *Nat. Struct. Biol.*, **10**: 425–432.
- Lebowitz, J. H.; Coburn, C. M.; McMahon-Pratt, D. and Beverley, S. M. (1990). *Proc. Natl. Acad. Sci. U. S. A.*, **87**: 9736–9740.
- Lebowitz, J. H.; Smith, H. Q.; Rusche, L. e Beverley, S. M. (1993). *Genes Dev.*, **7**: 996-1007.
- Lemire, B. D.; Fankhauser, C.; Baker, A. e Schatz, G. (1989). *J. Biol. Chem.*, **264**: 20206-20215.
- Martinez-Calvillo, S.; Lopez, I. e Hernandez, R. (1997). *Gene.*, **199(1-2)**: 71-76.
- Matthews, K. R.; Tschudi, C. e Ullu, E. (1994). *Genes Dev.* **8**: 491-501.
- Merle, M.; Trezeguet, V.; Graves, P. V.; Andrews, D.; Muench, K. H. e Labouesse, B. (1986). *Biochemistry.*, **25**: 1115-1123.
- Neupert, W. (1997), *Annu. Rev. Biochem.*, **66**: 863-917.
- Otani, A.; Slike, B. M.; Dorrell, M. I.; Hood, J.; Kinder, K.; Ewalt, K. L.; Cheresch, D.; Schimmel, P. e Friedlander, M. (2002). *Proc. Natl Acad. Sci. U. S. A.*, **99**: 178–183.
- Pfanner N. e Neupert, W. (1990). *Annu. Rev. Biochem.* **59**: 331-353.
- Rinehart, J.; Horn, K. E.; Wei D.; Soll, D. e Schneider, A. (2004). *J. Biol. Chem.*, **279**: 1161–1166.
- Sali, A. e Blundell, T. L. (1993). *J. Mol. Biol.*, **234**: 779-815.
- Sambrook, J. e Russel, D.W. (2000). *Molecular cloning: a Laboratory Manual*. 3^a ed. Cold Spring Harbor Laboratory Press.
- Service, R. F. (2002). *Science*, **298**: 948–950.

- Shen, N.; Guo, L.; Yang, B.; Jin, Y. e Ding, J. (2006). *Nucleic Acids Res.* **34**: 3246-3258.
- Sommer, J. M.; Hua, S.; Li, F.; Gottesdiener, K. M. e Wang, C. C. (1996). *Mol. Biochem. Parasitol.*, **76**: 83–89.
- Tetaud, E.; Lecuix, I.; Sheldrake, T.; Baltz, T. e Fairlamb, A. H. (2002). *Mol. Biochem. Parasitol.*, **120**: 195-204.
- Ullu, E.; Matthews, K. R. e Tschudi, E. (1993). *Mol. Cell. Biol.*, **13**: 720-725.
- Von Heijne, G. (1990). *Curr. Opin. Cell Biol.*, **2**: 604-608
- Wakasugi, K.; Slike, B. M.; Hood, J.; Otani, A.; Ewalt, K. L.; Friedlander, M.; Cheresch, D. A. e Schimmel, P. (2002). *Proc. Natl. Acad. Sci. U. S. A.*, **99**: 173–177.
- Westbrook, J.; Feng, Z.; Jain, S.; Bhat, T. N.; Thanki, N.; Ravichandran, V.; Gilliland, G. L.; Bluhm, W.; Weissig, H.; Greer, D. S.; Bourne, P. E. e Berman, H. M. (2002). *Nucleic Acids Res.*, **30**: 245-248.
- Xu, F.; Jia, J.; Jim, Y. e Wang, D.T. (2001a) *Protein Expr Purif.* **23**: 296-300.
- Xu, F.; Chen, X.; Xin, L.; Chen, L.; Jin, Y. and Wang, D. (2001b) *Nucleic Acids Res.*, **29**: 4125–4133
- Xue, H.; Shen, W.; Giege, R. e Wong, J.T. (1993). *J Biol Chem.*, **268**: 9316–9322
- Yang, X.; Otero, F. J.; Skene, R. J.; Mcree, D. E.; Schimmel, P. e Pouplana, L. R. (2003) *Proc. Natl. Acad. Sci. U. S. A.*, **100**: 15376–15380.
- Yang, X.I.; Otero, F. J.; Ewalt, K. I.; Liu, J.; Swairjo, M. A.; Kohrer, C.; Rajbhandary, U. I.; Skene, R. J.; McRee, D. E. e Schimmel, P. (2006). *EMBO J.*, **25**: 2919-2929.
- Yaremchuk, A., Kriklivyi, I., Tukalo, M. e Cusack, S. (2002). *EMBO J.*, **21**: 3829–3840.

Apêndice I

Materiais utilizados em biologia molecular.

Dentro desse apêndice encontra-se a descrição dos oligonucleotídeos utilizados para a caracterização das regiões 5' UTR e 3'UTR (Tabela AI-1.1). Ainda dentro do mesmo pode-se observar o detalhamento das construções plasmidiais realizadas em vetores de expressão de procariotos e eucariotos das proteínas LmTrpRS1 e LmTrpRS2, bem como os oligonucleotídeos utilizados para cada construção (Tabela AI-1.2).

Tabela AI-1.1: Nome e seqüência dos oligonucleotídeos utilizados na reação de RT-PCR para caracterização das regiões não traduzidas dos genes *LmTrpRS1* e *LmTrpRS2*. Os números na frente dos oligonucleotídeos representam a posição de ligação aos genes de interesse (vide figura abaixo)

Gene	Reação	Oligonucleotídeos	Seqüência dos oligonucleotídeos
<i>LmTrpRS1</i>	5'UTR	1 - SL	5'-aactaacgctatataagtatcagtttctgtactttattg-3'
		2 - LmW5UTR1	5'-ggaagtgaaggcagcgag-3'
		3 - LmW5UTR2	5'-cagagagatagacagcgagc-3'
	3'UTR	4 - LmN3UTR2	5'-gtaacggaggccgatgtggag-3
		5 - LmW3UTR1	5'-gatgggacctgcgaaagagg-3'
		3'UTR 1-a	5'-aagcgatcgagacatgtagcagttttttttttttta-3'
		6 - 3'UTR 1-c	5'-aagcgatcgagacatgtagcagtttttttttttttc-3'
3'UTR 1-g	5'-aagcgatcgagacatgtagcagtttttttttttttg-3'		
7 - 3'UTR 2	5'-ttgaattcgattgacacctgc-3'		
<i>LmTrpRS2</i>	5'UTR	1 - SL	5'-aactaacgctatataagtatcagtttctgtactttattg-3'
		2 - LmW5UTR1	5'-gcgcaaactccatggcc-3'
		3 - LmW5UTR2	5'-gatcactagtgcggccagagtac-3'
	3'UTR	4 - LmW3UTR2	5'-gtgagggtgctgcacgattgg-3'
		5 - LmW3UTR1	5'-gtggagttctgccgtatcc-3'
		3'UTR 1-a	5'-aagcgatcgagacatgtagcagttttttttttttta-3'
		6 - 3'UTR 1-c	5'-aagcgatcgagacatgtagcagtttttttttttttc-3'
3'UTR 1-g	5'-aagcgatcgagacatgtagcagtttttttttttttg-3'		
7 - 3'UTR 2	5'-ttgaattcgattgacacctgc-3'		



Estratégia adotada para a caracterização das regiões não traduzidas dos genes *LmTrpRS1* e *LmTrpRS2*. As setas indicam os oligonucleotídeos utilizados nas reações, cujas seqüências encontram-se descritos na tabela acima.

Tabela AI-1.2: Detalhamento do oligonucleotídeos utilizados na amplificação dos genes *LmTrpRS1* e *LmTrpRS2* de *L. major*, pelo método de PCR

<i>Gene</i>	<i>Detalhe da construção</i>	<i>Nome da construção</i>	<i>Oligonucleotídeos</i>	<i>Sítio de Restrição</i>	<i>Seqüência dos oligonucleotídeos</i>
<i>LmTrpRS1</i>	Gene <i>LmTrpRS1</i> *.	<i>LmTrpRS1/pgemT</i>	LM1-Full-5'* (sense)	<i>NdeI</i>	5'-atcgatcatatgatgtccccctttccccgctg-3'
			LM1-Full-3'* (antisense)	<i>KpnI</i>	5'-cggggtaccctattacttcttgactccgcctctcttc-3'
<i>LmTrpRS2</i>	Gene <i>LmTrpRS2</i> com região de sinalização mitocondrial.	<i>LmTrpRS2-full/pgemT</i>	LM2-Full-5' (sense)	<i>NdeI</i>	5' atcgatcatatgcgtcgttcacactactggccatg-3'
			LM2-3' (antisense)	<i>XhoI</i>	5' atctcgagttatcagagaagaatgtcacggatacggca 3'
	Gene <i>LmTrpRS2</i> sem região de sinalização mitocondrial, iniciando no aminoácido alanina 24.	<i>LmTrpRS2-24A/pgemT</i>	LM2-19L-5' (sense)	<i>NdeI</i>	5'-atcgatcatatggcagccgcgtactctggccgcac-3'
			LM2-3' (antisense)	<i>XhoI</i>	5' atctcgagttatcagagaagaatgtcacggatacggca 3'

* Gene *LmTrpRS1* clonado de acordo com a seqüência depositada no banco de dados do GeneDB (LmF29.060). Segundo nossos experimentos de caracterização das regiões não traduzidas a janela aberta de leitura começa no terceiro ATG desta seqüência (M32 da seqüência original).

Tabela AI-1.3: Detalhamento do oligonucleotídeos utilizados nas diversas construções dos vetores de expressão com os genes *LmTrpRS1* e *LmTrpRS2* de *L. major*.

<i>Gene</i>	<i>Detalhe da construção</i>	<i>Nome da construção</i>	<i>Oligonucleotídeos</i>	<i>Sítio de Restrição</i>	<i>Seqüência dos oligonucleotídeos</i>
<i>LmTrpRS1</i>	Gene <i>LmTrpRS1</i> *, clonado no vetor pET29a	<i>LmTrpRS1F/p29a</i>	LM1-Full-5'* (sense)	<i>NdeI</i>	5'-atcgatcatatgatgtccccctttccccgctg-3'
			LM1-Full-3'* (antisense)	<i>KpnI</i>	5'-cggggtaccctattacttcttctgactccgcctctcttc-3'
	Gene <i>LmTrpRS1</i> *, clonado no vetor pET29a. Proteína produzida com cauda de Histidina no C-terminal.	<i>LmTrpRS1F-HIS/p29a</i>	LM1-Full-5'* (sense)	<i>NdeI</i>	5'-atcgatcatatgatgtccccctttccccgctg-3'
			LM1-Full-3'* (antisense)	<i>KpnI</i>	5'-cggggtaccctattacttcttctgactccgcctctcttc-3'
	Gene <i>LmTrpRS1</i> *, clonado no vetor pET28a. Proteína produzida com cauda de Histidina no N-terminal.	<i>LmTrpRS1F-HIS/p28a</i>	LM1-Full-5'* (sense)	<i>NdeI</i>	5'-atcgatcatatgatgtccccctttccccgctg-3'
			LM1-Full-3'* (antisense)	<i>KpnI</i>	5'-cggggtaccctattacttcttctgactccgcctctcttc-3'
	Gene <i>LmTrpRS1</i> , clonado no vetor pET29a. Proteína produzida com cauda de Histidina no C-terminal.	<i>LmTrpRS1/p29a</i>	LM1T-5' (sense)	<i>NdeI</i>	5'-atcgatcatatgcgtcgcttcacactaggccatg-3'
			LM1T-3' (antisense)	<i>XhoI</i>	5-atctcgagttatcagagaagaatgtcacggataccgca 3'
	Gene <i>LmTrpRS1</i> , clonado no vetor pET28a. Proteína produzida com cauda de Histidina no N-terminal.	<i>LmTrpRS1/p28a</i>	LM1T-5' (sense)	<i>NdeI</i>	5'-atcgatcatatgcgtcgcttcacactaggccatg-3'
			LM1T-3' (antisense)	<i>XhoI</i>	5'-atctcgagttatcagagaagaatgtcacggataccgca 3'

<i>Gene</i>	<i>Detalhe da construção</i>	<i>Nome da construção</i>	<i>Oligonucleotídeos</i>	<i>Sítio de Restrição</i>	<i>Seqüência dos oligonucleotídeos</i>
<i>LmTrpRS1</i>	Gene <i>LmTrpRS1</i> , clonado no vetor pET32a. Proteína de fusão com Tioredoxina e cauda de histidina no N-terminal.	<i>LmTrpRS1/p32a</i>	LM1-p32-5' (sense)	<i>Bam</i> HI	5'-cgcgatccgaggaccccgccgagcaggaggttg-3'
			LM1-p32-3' (antisense)	<i>Hind</i> III	5'-cccaagcttctattactccttgactccgctcctctttc-3'
	Gene <i>LmTrpRS1</i> , clonado no vetor pGEX2T. Proteína de fusão com GST no N-terminal.	<i>LmTrpRS1/pgex2t</i>	LM1Tp32-5' (sense)	<i>Bam</i> HI	5'-cgcgatccgaggaccccgccgagcaggaggttg-3'
			LM1Tp32-3' (antisense)	<i>Eco</i> RI	5'-ccggaattctcattactccttgactccgctcctctttc-3'
	Gene <i>LmTrpRS1</i> , clonado no vetor pF4X1.4neo (expressão em <i>L. tarentolae</i>). Proteína com cauda de histidina no C-terminal.	<i>LmTrpRS1/pF4X-1.4neo</i>	Lm1p4-5' (sense)	<i>Nco</i> I	5'-catgcatatgaggaccccgccgagcaggaggtt-3'
			Lm1p4-3' (antisense)	<i>Pac</i> I	5'-ccttaattaatcagtggtggtggtggtggtgctcgag-3'
<i>LmTrpRS2</i>	Gene <i>LmTrpRS2</i> com região de localização mitocondrial, clonado no vetor pET29a	<i>LmTrpRS2F/29a</i>	LM2-Full-5' (sense)	<i>Nde</i> I	5' atcgatcatatgctgcgttcacactactggccatg-3'
			LM2-Full-3' (sense)	<i>Nde</i> I	5' atcgatcatatgctgcgttcacactactggccatg-3'
	Gene <i>LmTrpRS2</i> com região de localização mitocondrial, clonado no vetor pET28a. Proteína produzida com cauda de histidina no N-terminal.	<i>LmTrpRS2F/28a</i>	LM2-Full-5' (sense)	<i>Nde</i> I	5' atcgatcatatgctgcgttcacactactggccatg-3'
			LM2-Full-3' (sense)	<i>Nde</i> I	5' atcgatcatatgctgcgttcacactactggccatg-3'

<i>Gene</i>	<i>Detalhe da construção</i>	<i>Nome da construção</i>	<i>Oligonucleotídeos</i>	<i>Sítio de Restrição</i>	<i>Seqüência dos oligonucleotídeos</i>
<i>LmTrpRS2</i>	Gene <i>LmTrpRS2</i> sem região de sinalização mitocondrial, iniciando no aminoácido alanina 24. Clonado no vetor pET29a.	<i>LmTrpRS2-24A/pET29a</i>	LM2-24A-5' (sense)	<i>NdeI</i>	5'-atcgatcatatggcagccgctactctggccgcac-3'
			LM2-3' (antisense)	<i>XhoI</i>	5' atctcgagttatcagagaagaatgtcacggatacggca 3'
	Gene <i>LmTrpRS2</i> sem região de sinalização mitocondrial, iniciando no aminoácido alanina 24. Clonado no vetor pET28a. Proteína produzida com cauda de histidina no N-terminal	<i>LmTrpRS2-24A/pET28a</i>	LM2-24A -5' (sense)	<i>NdeI</i>	5'-atcgatcatatggcagccgctactctggccgcac-3'
			LM2 Δ stop -3' (antisense)	<i>XhoI</i>	5' atctcgagttatcagagaagaatgtcacggatacggca 3'
	Gene <i>LmTrpRS2</i> , com a região de sinalização mitocondrial, clonado no vetor pET32a. Proteína de fusão com Tioredoxina e cauda de histidina no N-terminal.	<i>LmTrpRS2F/p32a</i>	LM2-24A /p32-5' Sense	<i>Eco RI</i>	5'-ccggaattcgcagccgctactctggccgcac-3'
			LM2-3' (antisense)	<i>Xho I</i>	5'-atctcgagttatcagagaagaatgtcacggatacggca 3'
	Gene <i>LmTrpRS2</i> , com a região de sinalização mitocondrial, clonado no vetor pGEX2T. Proteína de fusão com GST no N-terminal.	<i>LmTrpRS2F/pgex2t</i>	LM2-24A /pgex-5' (sense)	<i>Bam HI</i>	5'-cgcgatccgcagccgctactctggccgcac-3'
			LM2-24A /pgex-3' (antisense)	<i>Eco RI</i>	5'-ccggaattctatcagagaagaatgtcacggatacggcag-3'
Gene <i>LmTrpRS2</i> , sem a região de sinalização mitocondrial, clonado no vetor pPICZ (expressão em <i>P. pastoris</i>).	<i>LmTrpRS2-24A/ppicz</i>	LM2-24A /ppicz-5' (sense)	<i>Eco RI</i>	5'-atgaattctataatgtctgcagccgctactctggccg -3'	
		LM2-3' (antisense)	<i>XhoI</i>	5' atctcgagttatcagagaagaatgtcacggatacggca 3'	

Gene	Detalhe da construção	Nome da construção	Oligonucleotídeos	Sítio de Restrição	Seqüência dos oligonucleotídeos
<i>LmTrpRS2</i>	Gene <i>LmTrpRS2</i> , sem a região de sinalização mitocondrial, clonado no vetor pPICZ α (expressão em <i>P. pastoris</i>). Proteína com cauda de histidina no C-terminal.	<i>LmTrpRS2-24A/ppiczα</i>	LM2-24A/ppicz-5' (sense)	<i>Eco</i> RI	5'-atgaattctataatgtctgcagccgctactctggccg -3'
			LM2 Δ stop/ppicz-3' (antisense)	<i>Xho</i> I	5'atctcgaggagaagaatgtcacggataggcag-3'
	Gene <i>LmTrpRS2</i> , com a região de sinalização mitocondrial, clonado no vetor pPICZ (expressão em <i>P. pastoris</i>).	<i>LmTrpRS2F-HIS/ppicz</i>	LM2F/ppicz-5' (sense)	<i>Eco</i> RI	5'-ccgaattcacgatggcccgtcgttcacactactggccatggag-3'
			LM2F Δ stop/ppicz-3' (sense)	<i>Xho</i> I	5'-atctcgaggagaagaatgtcacggataggcag-3'
<i>LmTrpRS2</i>	Gene <i>LmTrpRS2</i> , sem a região de sinalização mitocondrial, clonado no vetor pPICZ (expressão em <i>P. pastoris</i>). Proteína com cauda de histidina no C-terminal.	<i>LmTrpRS2Δ-HIS/ppicz</i>	LM2-24A/ppicz-5' (sense)	<i>Eco</i> RI	5'-atgaattctataatgtctgcagccgctactctggccg -3'
			LM2 Δ stop/ppicz -3' (antisense)	<i>Xho</i> I	5'-atctcgaggagaagaatgtcacggatacgga -3'
<i>LmTrpRS2</i>	Gene <i>LmTrpRS2</i> , com a região de sinalização mitocondrial, clonado no vetor pNUS-HcN (expressão em <i>P. pastoris</i> e <i>C. fasciculata</i>). Proteína com cauda de histidina no C-terminal.	<i>LmTrpRS2F/pNUS-HcN</i>	LM2-Full-5' (sense)	<i>Nde</i> I	5' atcgatcatatgcgtcgttcacactactggccatg-3'
			LM2 Δ stop/ppicz -3' (antisense)	<i>Xho</i> I	5'-atctcgaggagaagaatgtcacggatacgga -3'

<i>Gene</i>	<i>Detalhe da construção</i>	<i>Nome da construção</i>	<i>Oligonucleotídeos</i>	<i>Sítio de Restrição</i>	<i>Seqüência dos oligonucleotídeos</i>
<i>LmTrpRS2</i>	Gene <i>LmTrpRS2</i> , com a região de sinalização mitocondrial, clonado no vetor pF4X1.4neo (expressão em <i>L. tarentolae</i>). Proteína com cauda de histidina no C-terminal.	<i>LmTrpRS2F/pF4X-1.4neo</i>	LM2p4-5'	<i>BglIII</i>	5'- gaagatctatgcgtcgcttcacctactgg- 3'
			LM2p4-3'	<i>NotI</i>	5'-ataagaatcggcgcgctcattaatggtgatggtgatggtggagaa ,aatgcacggatagc-3'

* Gene *LmTrpRS1* clonado de acordo com a seqüência depositada no banco de dados do GeneDB (LmF29.060). Segundo nossos experimentos de caracterização das regiões não traduzidas a janela aberta de leitura começa no terceiro ATG desta seqüência (M32 da seqüência original).

Apêndice II

Artigo publicado na revista Memórias Instituto Oswaldo Cruz.

Dentro desse apêndice encontra-se o artigo publicado na revista “Memórias Instituto Oswaldo Cruz”, intitulado “*Effects of Trypanosoma brucei tryptophanyl-tRNA synthetases silencing by RNA interference*”. Esse trabalho foi realizado em colaboração com a doutoranda Liliana Tocaroma García, aluna do programa de pós graduação da Universidade Federal de São Carlos

Effects of *Trypanosoma brucei* tryptophanyl-tRNA synthetases silencing by RNA interference

Liliana Torcoroma García/*, Ney Ribeiro Leite, Juan D Alfonzo**,
Otavio Henrique Thiemann/*/+

Instituto de Física de São Carlos, Universidade de São Paulo, Caixa Postal 369, 13560-590 São Carlos, SP, Brasil *Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos, SP, Brasil **Department of Microbiology, Ohio State Biochemistry Program and the Ohio State University RNA Group, The Ohio State University, Columbus, OH, US

*The kinetoplast genetic code deviates from the universal code in that 90% of mitochondrial tryptophans are specified by UGA instead of UGG codons. A single nucleus-encoded tRNA^{Trp}(CCA) is used by both nuclear and mitochondria genes, since all kinetoplast tRNAs are imported into the mitochondria from the cytoplasm. To allow decoding of the mitochondrial UGA codons as tryptophan, the tRNA^{Trp}(CCA) anticodon is changed to UCA by an editing event. Two tryptophanyl tRNA synthetases (TrpRSs) have been identified in *Trypanosoma brucei*: TbTrpRS1 and TbTrpRS2 which localize to the cytoplasm and mitochondria respectively. We used inducible RNA interference (RNAi) to assess the role of TbTrpRSs. Our data validates previous observations of TrpRS as potential drug design targets and investigates the RNAi effect on the mitochondria of the parasite.*

Key words: tryptophanyl tRNA synthetase - RNA interference - kinetoplastid - *Trypanosoma brucei*

Parasitic protozoa represent a considerable source of human diseases predominant in the tropical and subtropical regions of the globe. Of particular relevance are parasites belonging to the family Trypanosomatidae. *Trypanosoma brucei* and its subspecies are responsible for heavy socioeconomic losses in 36 sub-Saharan Africa countries (1.6 million DALYs) and lead to an estimated 48 thousand deaths per year. Therapy against African sleeping sickness has always been difficult, especially when the disease has reached an advanced stage with central nervous system involvement, as few effective drugs are available. For these reasons, the search and validation of new pharmacological targets against trypanosomiasis is of great concern.

Trypanosomatids are early diverging flagellates that differ from other eukaryotes in a number of features. They contain a remarkable single mitochondrion (the kinetoplast) harboring a large mass of circular DNA molecules concatenated in a unique arrangement, the kinetoplast DNA or kDNA (Lukes et al. 2002). The kinetoplast genetic code, like many other mitochondrial genomes and to the exception of plants, deviates from the universal code in that 90% of tryptophans are specified by UGA instead of UGG codons, although both codons are used in translation. A single nucleus-encoded tRNA^{Trp}(CCA) is used by both the nuclear and mitochon-

drial genes. Since all kinetoplast tRNAs are encoded in the nucleus and later imported into the mitochondria (Lima & Simpson 1996), a C to U deamination event at position 34 of tRNA^{Trp}(CCA) changes the anticodon from CCA to UCA allowing the decoding of UGA stop codons to Trp (Alfonzo et al. 1999).

The identification of two tryptophanyl tRNA synthetases (TrpRSs) proteins in *T. brucei* cells was reported (Charrière et al. 2006). TbTrpRS1 (cytoplasmic, accession no. XM821841) and TbTrpRS2 (mitochondrial, accession no. XM820510), where TbTrpRS is able of aminoacylate both the unedited tRNA^{Trp}(CCA) and the edited tRNA^{Trp}(UCA). We used inducible RNA interference (RNAi) in combination with biochemical techniques to assess the role of TbTrpRSs, thereby extending previous observations (Charrière et al. 2006) and validating the TrpRSs as potential drug design targets.

MATERIALS AND METHODS

The nucleic acid and amino acid sequences of both genes, TbTrpRS1 and 2, were used to search the *T. brucei* genomic databases, and the results showed no significant identity with other *T. brucei* genes as well as with their human and *Escherichia coli* homologues. A sequence identity of 57% is shared between TbTrpRS1 and 2 distributed along the entire sequence. Due to the low sequence identity, we chose the full length ORF sequences of TbTrpRS1 and 2 for RNAi analysis, based on the idea of increasing the specificity and efficiency of RNAi knockdown (LaCount et al. 2000). *T. brucei* 29-13 procyclic cells were grown in SDM-79 supplemented with 10% fetal bovine serum in the presence of hygromycin (50 µg/ml) and G418 (15 µg/ml) at 25-26°C without agitation (Wang et al. 2000). Each specific DNA for RNAi experiments was amplified by the polymerase chain reaction (PCR) from *T. brucei* 29-13 procyclic genomic DNA using the following gene-specific primers, containing the *Xho*I and *Hind*III (underlined) restric-

Financial support: CNPq (472570/2004-8), Fapesp (98/14138-2). LTG is a Capes doctorate fellow. NRL is a Fapesp fellow (02/01883-9). JDA is supported by a grant from the American Heart Association.

*Corresponding author: thiemann@ifsc.usp.br

Received 27 April 2007

Accepted 1st August 2007

tions sites: TbTrpRs1a 5'-CGCCTCGAGATGAC-TGAAGCCGTTGTAAACACCGTGGTCT-3', TbTrpRs1b 5'-TCAAGCTTCTCACTGGGTTGCCGCCTTCTTCGCTGGAC-3', TbTrpRs2a 5'-TACTCGAGATGCGCCGGGCTTCACATGTTTGCACGTC-3' and TbTrpRs2b 5'-CTAAGCTTCTATGCCATAATGTTGCGCACCTCCGTAATA-3'. The PCRs were carried out in a PTC-100 thermocycler (MJ Research Inc.) with 0.2 units of *Taq* DNA polymerase (Biotools) according to the manufacturer conditions. The samples were subjected to 3 min denaturation at 94°C followed by 20 cycles of denaturation at 94°C for 0.5 min, annealing at 54°C for 0.5 min, and extension at 72°C for 1.5 min. A final step at 72°C for 10 min was added to maximize the full-length products. Each PCR fragment was cloned into the pCR2.1 TOPO vector and sequenced prior to cloning into the *XhoI/HindIII* sites of the pZJM vector (Wirtz et al. 1999). These constructs were linearized by *NofI* cleavage and transfected into procyclic *T. brucei* 29-13 cell line by electroporation as previously described (Wirtz et al. 1999, Wang et al. 2000). The transfected cells were cloned by the limiting dilution technique and selected under 2.5 µg/ml phleomycin until stable cell lines were obtained. To induce synthesis of the double stranded RNA (dsRNA), cells were incubated in medium containing tetracycline (1.5 µg/ml) for 24 h and transferred daily to fresh medium with tetracycline. Growth curves were obtained using cells fixed with 4% (v/v) formaldehyde in PBS and counted using a hemacytometer.

For cell morphology analysis, non-induced and tetracycline-induced cultures were examined at different time intervals after RNAi induction by light microscopy. We used different staining techniques: (a) DAPI (4,6-diamidino-2-phenylindole, Sigma) for nuclear and kinetoplast DNA. Cells were washed in PBS, and then fixed and stained with DAPI (75 ng/ml) for 1 min and analyzed by fluorescence microscopy at 350 nm of λ excitation; (b) *Mitotracker*: the mitochondrion of living trypanosomes was visualized with the potential sensitive vital stain MitoTracker Red CMXRos (λ_{Em} 579nm and λ_{Ex} 599nm from molecular probe), a derivative of X-rosamine. Cells in *log* phase of growth were incubated with 100 nM MitoTracker Red CMXRos at 28°C for 5 min in SMGHP, then, the cells were washed with fresh media and finally, incubated without dye for 20 min; (c) *DAPI-RED*: fluorescence images of cells stained with 60 µM DAPI and MitoTracker Red CMXRos 5 µM. For fixation, cells were washed with PBS and suspended in 4% freshly made paraformaldehyde in PBS. The percentage of morphologic anomalies (number nucleus, kinetoplasts, and form) was determined in 150 cells.

To determine the percentage of TbTrpRSs mRNA inhibition, total *T. brucei* RNA was isolated using Trizol (Life Technologies) after RNAi induction from the different cell lines at different time points. The TbTrpRS1 and TbTrpRS2 mRNA were estimated by semi-quantitative RT-PCR (Kit One Step RT-PCR from Qiagen) and by quantitative real-time PCR (QPCR), using the spliced leader (SLRNA) primer 5'-AACTAACGCTATATAAGTATCAGTTTCTGACTTTATTG-3' (Campbell et al. 1984) and TbTrpRs1b and TbTrpRs2b primers. For the

semi-quantitative RT-PCR experiments 100 ng of total RNA of each sample were added to a final reaction volume of 50 µl. The samples, after RT-PCR were analyzed in a 2% agarose gel. The bands were quantitated using the Kodak ds 1D digital science V.3.0.1 software. The analysis was performed comparing the intensity of the ethidium bromide stained bands of control (non-induced) and RNAi induced samples. QPCR was performed using the Absolute QPCR SYBR Green Mix (Abgene Advanced Biotechnologies Ltd.) as previously described (Yin et al. 2001) and performed in a Rotor Gene Real Time PCR (Corbett Research, Uniscience). The QPCR data obtained was analyzed using paired t-tests on the linear $2^{-\Delta Ct}$ data set (Livak & Schmittgen 2001). The RT-PCR and QPCR experiments of TbTrpRSs were performed. As a quantitation standard the *T. brucei* glycosomal glyceraldehyde 6-phosphate dehydrogenase (gGAPDH, X59955.1) gene was used with the primer pair TbGAPDH α 5'-CCTC GGG AATGAGATTGATGTCGTGCTGT-3' and TbGAPDH β 5'-CACATACTCCACACCAAGCTTTCCC-3' in a standard PCR, as described above. Samples from induced, non-induced and non-transfected (control) cells were tested.

Oxygen uptake by *T. brucei* cells was measured with a biological oxygen monitor DM4 equipped (Digimed) with oxygen probe. The rate (ppm) was calculated by assuming the oxygen content of air-saturated water is 0.276 mM at 28°C (according to manufacturer instructions). The cell concentration used was 1×10^6 cells/ml.

RESULTS

RNAi expression against TbTrpRS1 resulted in severe growth inhibition 4 h after induction (reduction of 84% relative to control cells) with cell death after 24 h of induction. TbTrpRS2 knockdown after 6 h of tetracycline induction produced a 79% growth inhibition and complete cell death after 72 h of induction. Both cell lines exhibited eventual complete arrest of growth, which was more quickly achieved in post-induced TbTrpRS1 cultures (24 h) than in TbTrpRS2 (72 h), as shown in Figs 1, 2. These cultures showed no recovery after release from tetracycline induction, confirming the lethal and irreversible effect of the gene knockdowns (data not shown).

Expression of TbTrpRS1 dsRNA led to total degradation of the specific mRNA, which was undetectable by semi-quantitative RT-PCR. QPCR analysis revealed that transcript levels were reduced more than 16 fold in cells after 24 h of TbTrpRS1 dsRNA induction (Fig. 2). Expression of TbTrpRS2 dsRNA also resulted in an efficient degradation of the cognate mRNA resulting in an approximately 3-fold decrease after 48 h of induction and a complete silencing at 72 h of induction, after which TbTrpRS2 mRNA amounts were not detectable by either semi-quantitative RT-PCR or QPCR (levels < 13 folds) (Fig. 2). These data confirmed the efficiency of silencing of expression of the TbTrpRSs by RNAi. Transcript levels were not significantly changed in the non-induced TbTrpRS2 controls or in the wild type *T. brucei* 29-13 cells ($P = 0.62$). An unrelated transcript for *T. brucei* gGAPDH remained unaffected in the samples analyzed. Furthermore, silencing of one TbTrpRS gene did not in-

fluence the expression of the other (data not shown). The non-induced *T. brucei* cells transfected with the TbTrpRS1 RNAi construct exhibited a small reduction (0.039 folds; $P = 0.06$) of the TbTrpRS1 mRNA as compared to the non-transfected *T. brucei* 29-13 cells. This phenomenon can be explained by a leakage in the T7 dual-promoter system, which has been previously observed (Durand-Dubief et al. 2003).

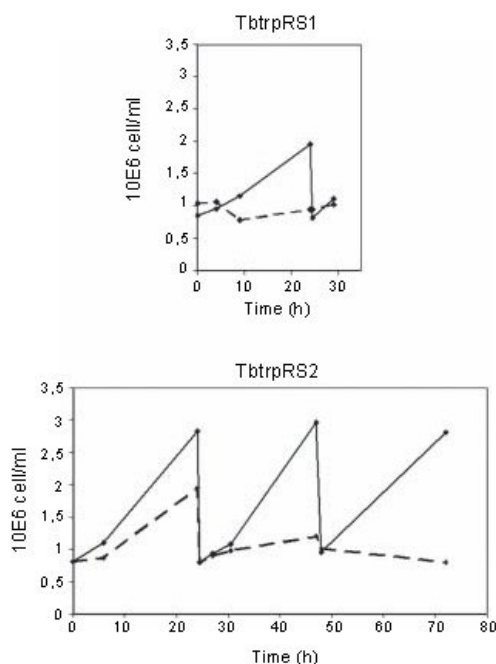


Fig. 1: growth of tetracycline induced (dotted line) and non-induced (solid line) TbTrpRS1 and 2 RNAi strains of *Trypanosoma brucei* 29-13 cells.

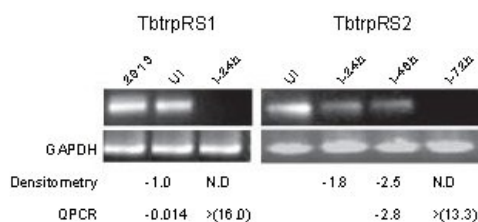


Fig. 2: reverse-transcription-polymerase chain reaction (RT-PCR) and quantitative real-time PCR (QPCR) of *Trypanosoma brucei* cytoplasmic tryptophanyl tRNA synthetase (TbTrpRS1) and 2 mRNA. 2913: non-transfected cells. UI: TbTrpRS1 non-induced cells and I-24 h, I-48 h and I-72 h cells induced after 24, 48, and 72 h, respectively. GAPDH gene was used as a control and is shown for comparison. Quantification was performed by both, band densitometry (Densitometry lane) from semi-quantitative RT-PCR gel bands shown in the figure and by QPCR (data not shown).

To analyze the cell morphology, non-induced and tetracycline induced cultures were examined. The phenotype of the control (non-induced) cells appeared normal (Figs 3A, 4A). The distribution of cell morphologies of induced and non-induced knockdowns is presented in the Table. The cells were classified in two categories according to the number of kinetoplasts (K) and nuclei (N) per cell. "Normal" constituted cells with 1K1N or with 2K1N/2K2N, representing different stages of cell division. Cells were classified as "abnormal" when containing a different number of K or N. TbTrpRS1 knockdown cells also presented a round phenotype and were classified as 'spheres'. In the control cultures, the population mainly (90-94%) constituted of 1K1N cells and a minority of dividing cells (2K1N or 2K2N) as shown in Fig. 4A and in the Table.

In the TbTrpRS1 RNAi cells, morphological abnormalities can be seen after 9 h of tetracycline induction. At 24 h after tetracycline induction the majority of cells (41%) increased in size, and contained numerous nuclei, kinetoplast, and several flagella, suggesting inhibition of cytokinesis. A significant percentage (33%) presented the 'sphere' phenotype lacking flagella (Table, Fig. 3B). The TbTrpRS2 knockdown leads to a wide range of morphologic abnormalities. These phenotypes included cells with partially detached flagella (two or more flagella) (Fig. 3B), cells with incomplete cell separations (Fig. 3C), mini-trypanosomes that lack nuclei (Fig. 4D) and cells with defects in cytokinesis (multinucleated; Fig. 4E). At 24 h after tetracycline induction, a high proportion of the cells (36%) presented abnormalities in their morphology (Figs 3B, 3C, 4B) compared to untreated cells, in which the percentage of abnormal cells was no more than 10% (Table). The depletion of TbTrpRS2 resulted in the accumulation of mini-trypanosomes that lack nuclei, suggesting an interruption of mitosis without inhibiting kinetoplast segregation, which led to cytokinesis and cell division and generated enucleated daughter cells (Ploubidou et al. 1999). In Figs 3 and 4 are shown some representative types of cells found in the knockdown of TbTrpRS1 and 2.

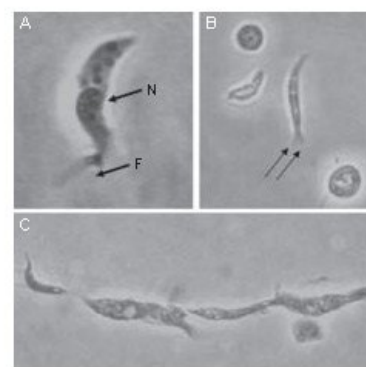


Fig. 3: phase contrast images from *Trypanosoma brucei* cells. A: non-induced TbTrpRS1 cells. Arrows indicate N: nucleus; F: flagella; B: TbTrpRS1 cells 24 h after tetracycline induction. Arrows indicate cell with several detached flagella; C: TbTrpRS2 cells 48 h after RNAi induction.

TABLE
Cell morphology of *Trypanosoma brucei* cells after RNAi induction

Cells	Normal			Abnormal						
	1K1N %	2K1N %	2K2N %	1K0N %	2K0N %	0K2N %	1K2N %	0K0N %	1K3N %	Sphere %
Tb. 29-13 control	83	11	3	3						
TbTrpRS1										
Non-induced	79	11	3	3		2	2			
Induced 24 h	9	2	3	3		3	3	3	41	33
TbTrpRS2										
Non-induced	78	12	2	4		2	2			
Induced 24 h	51	13	5	9	5	6	2	8	1	
Induced 48 h	48	8	6	16	4	2	6	8	2	
Induced 72 h	21	4	19	43	4	4	2	1	2	

K: kinetoplast, N: nucleus; T.b. 29-13 control: wild type *T. brucei* 29-13 cells; Tb TrpRS1: Tb TrpRS1 transfected *T. brucei* 29-13 cells; Tb TrpRS2: Tb TrpRS2 transfected *T. brucei* 29-13 cells.

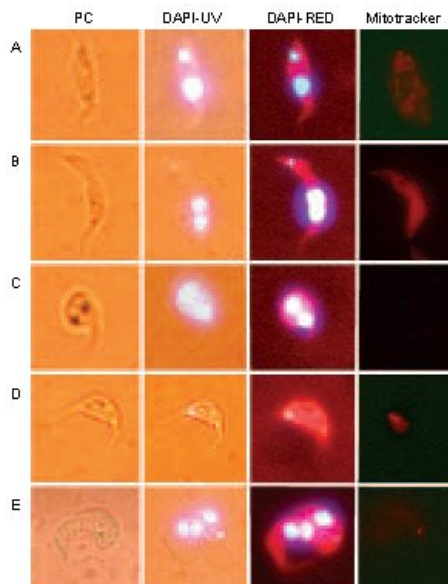


Fig. 4: cell morphology, kinetoplast, and nuclear DNA of *Trypanosoma brucei* mitochondrial tryptophanyl tRNA synthetase (TbTrpRS2) strain after different periods of incubation with tetracycline. PC: phase contrast images. DAPI-UV: DAPI stained view using both, visible and UV lights. DAPI-RED: fluorescence images of cells stained with *Mitotracker* Red CMXRos and DAPI stained observed in the DAPI channel; *Mitotracker*: fluorescence images of cells stained with the potential sensitive vital stain *Mitotracker* Red CMXRos in live cells; A: non-induced TbTrpRS2; B: 24 h after RNAi induction; C: 48 h after RNAi induction; D: example of Zooid morphology that appeared after 48 h of RNAi induction; E: 72 h after RNAi induction.

We used the mitochondria-specific stain, Mitotracker (molecular probes), in order to assay functional changes occurring during TbTrpRS2 RNAi induction. A typical reticular pattern is visible throughout the non-induced cells, representing the stained mitochondria, and this

patter is significantly altered after TbTrpRS2 RNAi induction. Induced cells show a decrease in the mitochondrial stain, in particular after 48 h of induction (Fig. 4C), indicating dramatic changes of metabolic function. To corroborate these results, the respiration rate or demand of oxygen (DO) was determined using a biological oxygen monitor DM4 (Digimed). The uninduced cells exhibit non-significant changes in their respiration rate compared with the control ($P = 0.789$). The respiration rate (ppm) was significantly decreased (64%) at 48 h post-induction (Fig. 5). This result is consistent with the three-fold decrease in TbTrpRS2 mRNA (Fig. 2). After 72 h the DO from these cells was undetectable. The quantitative mRNA analysis from TbTrpRS2 RNAi induced samples suggested that the decrease in the DO and the reduction of the mitochondrial potential occurred after the mRNA depletion (Figs 5, 6). According to these results, we can conclude that slight perturbations in the expression of TbTrpRS2 can trigger significant and immediate alterations in the mitochondrial function and redox potential. In addition, at 72 h post-induction, the complete degradation of specific mRNA in the knock-down cells was compatible with the complete lost of respiration activity and mitochondrial membrane potential (Figs 5, 6). Furthermore, the inhibition of TbTrpRS2 synthesis could decrease the levels of mitochondrial translation, with a decline of mitochondrial redox potential, respiration capability and finally triggering cell death. The TbTrpRS1 RNAi induced deleterious phenotype (Figs 1, 2) prevented an investigation of its role in mitochondria function, although such a role seems improbable.

DISCUSSION

The results we report are consistent with previous observations (Charrière et al. 2006) describing the existence of two TrpRSs genes in *T. brucei* cells and that knockdown of TbTrpRSs by RNAi effectively inhibits the replication of *T. brucei* procyclic forms in culture. In our experiments we extended those observations by the use of full length TbTrpRS1 and 2 genes in the RNAi

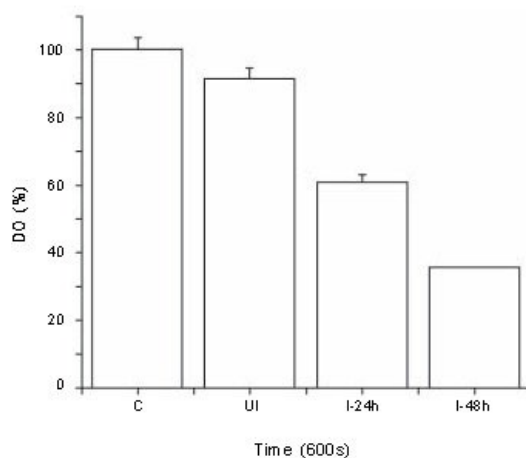


Fig. 5: assessment of respiratory function, oxygen demand (DO) of *Trypanosoma brucei* mitochondrial tryptophanyl tRNA synthetase (TbTrpRS2). C: *T. brucei* 29-13 control cells; UI: non-induced TbTrpRS2; I-24h and I-48h: TbTrpRS2 24h and 48h after induction respectively. At 72h DO was non-detectable in this assay.

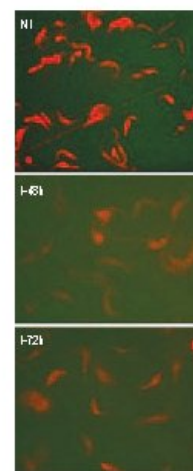


Fig. 6: fluorescence images of cells stained with Mitotracker Red CMXRos showing the accumulation of the redox potential sensitive dye by procyclic *Trypanosoma brucei* mitochondrial tryptophanyl tRNA synthetase (TbTrpRS2). NI: non-induced TbTrpRS2 cells; I-48h and I-72h: TbTrpRS2 cells 48h and 72h after tetracycline induction, respectively.

constructs. This contributed to a more pronounced sequence-specific knockdown, leading in a clearer physiological effect. Our results show that the outcome of RNA interference on TbTrpRS2 is a rapid decrease of mitochondrial activity, leading to a rapid cell death. The rapid rate which mutant phenotypes develop after expressing TbTrpRSs dsRNA indicate that these proteins have short half-lives since those enzymes are not abundant in the cells (data not shown). Our results confirm the central role that TbTrpRS1 and 2 play in cell viability, validating these proteins as targets for inhibitor development. In particular TbTrpRS2 is of interest due to its low sequence identity to other TbTrpRSs genes, especially that of humans, and for its role in blocking mitochondrial activity and resulting in cell death.

ACKNOWLEDGEMENTS

To the members of the Protein Crystallography and Structural Biology Group (IFSC-USP) for the helpful discussions in the course of this work. To Prof. Dr David Campbell Laboratory for providing pZJM vector and *T. brucei* 29-13 procyclic cells (Microbiology Immunology Molecular Genetics, University of California), Prof. Dr Heloisa Selistre de Araujo (LBBM, Federal University of São Carlos) for the QPCR support and Prof. Dr Luiz Antonio Daniel (São Carlos Engineering School, University of São Paulo) for support in the DO measurements.

REFERENCES

- Alfonzo JD, Blanc V, Estevez AM, Rubio MA, Simpson L 1999. C to U editing of the anticodon of imported mitochondrial tRNA^{Trp} allows decoding of the UGA stop codon in *Leishmania tarentolae*. *EMBO J* 18: 7056-7062.
- Campbell DA, Thornton DA, Boothroyd JC 1984. Apparent discontinuous transcription of *Trypanosoma brucei* variant surface antigen genes. *Nature* 311: 350-355.
- Charière F, Helgadóttir S, Hom, EK, Söll D, Schneider A 2006. Dual targeting of a single tRNA^{Trp} requires two different tryptophanyl-tRNA synthetases in *Trypanosoma brucei*. *PNAS* 103: 6847-6852.
- Durand-Dubief M, Kohl L, Bastin P 2003. Efficiency and specificity of RNA interference generated by intra- and intermolecular double stranded RNA in *Trypanosoma brucei*. *Mol Bioch Parasitol* 129: 11-21.
- Kim S, Lee SW, Choi EC, Choi SY 2003. Aminoacyl-tRNA synthetases and their inhibitors as a novel family of antibiotics. *Appl Microbiol Biotechnol* 61: 278-288.
- LaCount DJ, Bruse S, Hill KL, Donelson JE 2000. Double-stranded RNA interference in *Trypanosoma brucei* using head-to-head promoters. *Mol Bioch Parasitol* 111: 67-76.
- Lima BD, Simpson L 1996. Sequence-dependent *in vivo* importation of tRNA into the mitochondrion of *Leishmania tarentolae*. *RNA* 2: 429-440.
- Livak KJ, Schmittgen TD 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2-Delta Delta CT Method. *Methods* 25: 402-408.
- Lukes J, Guilbride DL, Votýpka J, Ziková A, Benne R, Englund PT 2002. Kinetoplast DNA network: evolution of an improbable structure. *Eukaryotic Cell* 1: 495-502.
- Ploubidou A, Robinson DR, Docherty RC, Ogbadoyi EO, Gull K 1999. Evidence for novel cell cycle checkpoints in trypanosomes: kinetoplast segregation and cytokinesis in the absence of mitosis. *J Cell Sci* 112: 4641-4650.
- Wang Z, Morris JC, Drew ME, Englund PT 2000. Interference of *Trypanosoma brucei* gene expression by RNA interfer-

6 Effects of *T. brucei* tryptophanyl • Liliana Torcoroma García et al.

- ence using integratable vector with opposing T7 promoters. *J Biol Chem* 275: 40174-9.
- Wirtz E, Leal S, Ochatt C, Cross GA 1999. A tightly regulated inducible expression system for conditional gene knockouts and dominant-negative genetics in *Trypanosoma brucei*. *Mol Biochem Parasitol* 99: 89-101.
- Yin JL, Shackel NA, Zekry A, McGuinness PH, Richards C, van der Putten K, McCaughan GW, Eris JM, Bidhop GA 2001. Real-time reverse transcriptase-polymerase chain Reaction RT-PCR for measurement of cytokine and growth factor mRNA expression with fluorogenic probes or SYBR Green I. *Immunol Cell Biol* 79: 213-221.

Parte II

**Estudos estruturais da Xylellaína, uma cisteíno protease da
bactéria *Xylella fastidiosa*.**

Capítulo 1

Introdução

Nesse capítulo é feita uma descrição geral sobre a bactéria *Xylella fastidiosa* e a doença Clorose Variegada do Citrus. Em seguida é realizada uma introdução sobre proteases e cisteíno proteases.

Xylella fastidiosa

A *X. fastidiosa* é uma bactéria gram negativa, não flagelada que coloniza exclusivamente os vasos do xilema das plantas infectadas (figura II-1.1) (Lambais *et al.*, 2000; Koide, *et al.*, 2004).

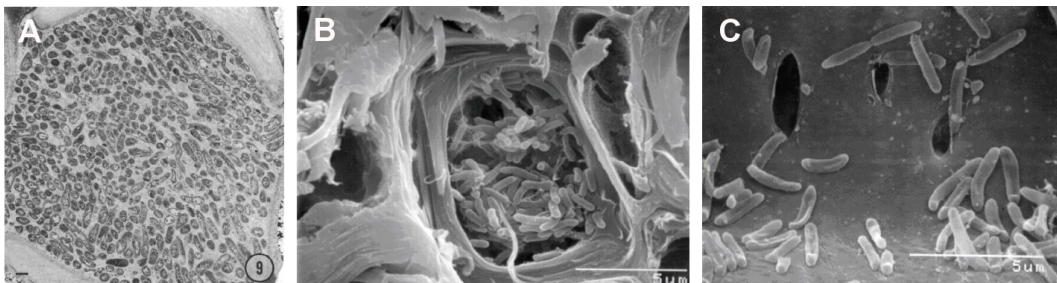


Figura II-1.1: Microscopia eletrônica vasos do xilema mostrando a distribuição das colônias de *X. fastidiosa*. A) Seção transversal de um vaso de xilema de videira (*Vitis vinifera*) completamente obstruído por colônia de *X. fastidiosa*, “Pierce’s Disease” (figura extraída de Mollenhauer, *et al.*, 1974). B e C) Detalhe de *X. fastidiosa* aderidas às paredes do vaso xilema de citrus (modificado de <http://aeg.lbi.ic.unicamp.br/xf/>).

Dentro do grupo de bactérias patogênicas em plantas a *X. fastidiosa* é caracterizada como uma bactéria “limitada ao xilema” (XLB) e de difícil cultivo por procedimentos padrões de bacteriologia (Purcell e Hopkins, 1996). O seu crescimento é lento em meio de cultura adequado, o qual lhe rendeu a denominação ‘fastidiosa’.

Diversas doenças economicamente importantes são causada por diferentes linhagens de *X. fastidiosa*, por exemplo: citrus (clorose variegada dos citrus – CVC) (Rossetti *et al.*, 1990), videiras (*doença de Pierce* - PD), mirta, amêndoa, café (Hendson *et al.*, 2000; Van Sluys *et al.*, 2003), alfafa anã, olmo, plátano, amora (Monteiro *et al.*, 2001), e crucíferas, como repolho, rabanete e nabo (Williams, 1980). Em todos os casos, a bactéria infecta o xilema da planta comprometendo a produção do fruto.

O ciclo de disseminação da bactéria nas plantas é realizado através de insetos vetores, que ao se alimentarem do xilema das plantas contaminadas adquirem a bactéria. Dentro do intestino do inseto as bactérias se reproduzem, e esse inseto infectado passa a disseminar a doença para plantas saudáveis através do aparelho picador-sugador durante sua alimentação. (Brlansky *et al.*, 1983; Roberto *et al.*, 1996).

Clorose Variegada dos Citros ou Amarelinho

A CVC é considerada uma das doenças mais devastadoras dos citrus e foi primeiramente relatada em 1987, nos estados de São Paulo e Minas Gerais (Rossetti *et al.*, 1990; Chang *et al.*, 1993; Purcel e Hopkins, 1996; Souza Pinto *et al.*, 2001). A partir dessa data a doença se espalhou rapidamente em todas as regiões citricultoras paulistas através da técnica do enxerto com ramos contaminados e por insetos vetores, tornando-se largamente distribuída.

Esta doença gera grandes prejuízos, principalmente no estado de São Paulo, que é o maior exportador de suco de laranja do mundo, sendo responsável por aproximadamente 90% da produção nacional e 30% da mundial. No Brasil, a CVC causa perdas anuais de aproximadamente 100 milhões de dólares para a indústria cítrica (Souza *et al.*, 2003).

A transmissão da doença é feita por pelo menos 11 espécies de insetos vetores homópteros (cigarrinhas) que se alimentam de seiva vegetal, algumas delas ocorrendo também na Argentina, Bolívia, Colômbia, Equador e Paraguai. (www.fundecitrus.com.br)

Os sintomas iniciais da CVC são observados na parte da árvore onde houve contato com o vetor, caracterizando-se por pequenas manchas amareladas na parte

lisa de folhas maduras. Após o período de infecção inicial, a CVC espalha-se gradativamente por toda a árvore. Nesse estágio as pequenas manchas amareladas evoluem para lesões intensas de cor palha dos dois lados da folha. As folhas tornam-se murchas, conduzindo à desfolha dos ramos (figura II-1.2). As manifestações nos frutos causam amadurecimento precoce, diminuição do fruto e em estágios mais avançados o endurecimento da casca a ponto de inviabilizar a comercialização dos frutos. (Rossetti *et al.*, 1990; www.fundecitrus.com.br). Plantas de qualquer idade são acometidas pela CVC, porém os sintomas mais severos ocorrem entre os dois e seis anos de idade.

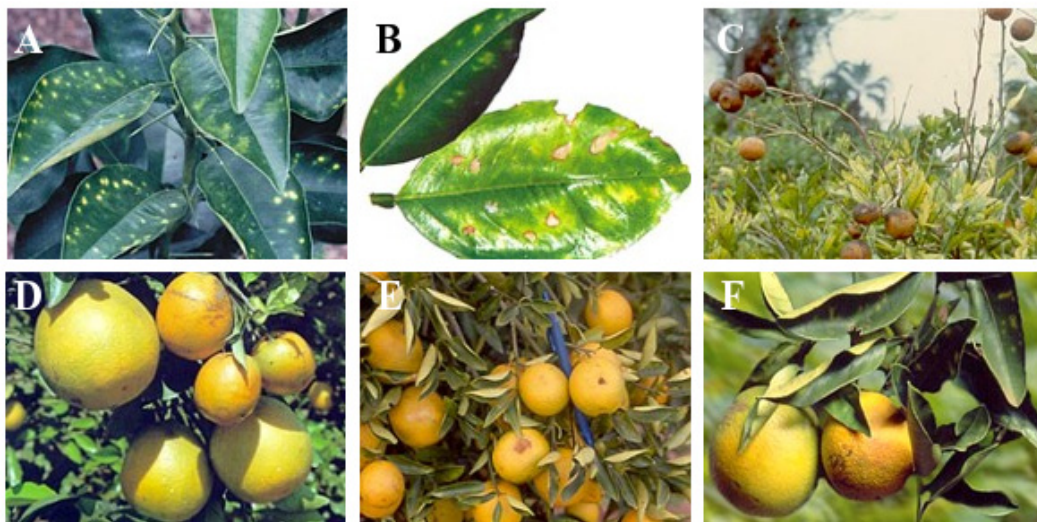


Figura II-1.2: Sintomas da CVC em folhas e frutos de citrus. A) Sintomas iniciais da CVC em folhas, pequenas manchas amareladas na face lisa da folha. B) Folhas em fase avançada de infecção, coloração de palha na parte inferior da folha. C) Perda das folhas em galho afetado, frutos endurecidos. D) Comparação de frutos normais e frutos afetados pela CVC. E) Sintomas de murcha em folhas e queimadura do sol em frutos. F) Detalhe de murcha em folhas e queimadura do sol em frutos. (Figura adaptada de www.fundecitrus.com.br)

Atualmente, o controle da CVC inclui apenas o manejo de cultura, com a poda de galhos infectados, eliminação de árvores contaminadas, utilização de mudas saudáveis e combate químicos aos insetos vetores (Souza Pinto *et al.*, 2001).

A disponibilização do genoma da bactéria *X. fastidiosa* clone 9a5c, em 2000, representou o primeiro seqüenciamento completo de um fitopatógeno (Simpson *et al.*, 2000). Muitas linhagens dessa bactéria tiveram seus genomas completamente seqüenciados, como por exemplo, duas linhagens de *X. fastidiosa* causadoras de doenças em amendoeiras e loureiros (Bhattacharyya *et al.*, 2002); *X. fastidiosa* Temecula, bactéria isolada de videiras infectadas com a Doença de Pierce (Van Sluys *et al.*, 2003); e outros estão em andamento.

A identificação e caracterização dos genomas destes fitopatógenos levam a um significativo avanço no entendimento do metabolismo, bem como mecanismos associados à virulência e patogenicidade. Estas duas linhagens de *X. fastidiosa* apresentam funções metabólicas idênticas e um conjunto de genes comuns utilizados na colonização da planta e patogênese (Van Sluys *et al.*, 2003). Contudo muito pouco se sabe sobre os mecanismos moleculares de patogenicidade da *X. fastidiosa*. Porém muitos genes possivelmente relacionados com a patogenicidade foram encontrados no genoma da *X. fastidiosa* quando comparados aos de outras bactérias (Simpson *et al.*, 2000; Lambais *et al.*, 2000; Van Sluys, 2003). Entre os fatores relacionados à patogenicidade pode-se destacar os processos de interação célula-célula, degradação da parede celular da planta, homeostase de íons, resposta antioxidante e síntese de toxinas. Algumas enzimas envolvidas nessas funções, como as celulases, xilanases, pectinases e proteases, foram identificadas neste patógeno (Simpson *et al.*, 2000). A degradação da parede celular pelas enzimas tem importante papel não só na migração das bactérias entre os vasos, como também na liberação de carboidratos, que são necessários para o crescimento bacteriano na parte interna dos vasos xilemáticos (Lambais *et al.*, 2000).

O estudo do genoma e seus produtos gênicos, em especial aqueles que podem estar envolvidos na patogenicidade da *X. fastidiosa* são de interesse para a busca de mecanismos ou inibidores seletivos que possam ser utilizados no combate a CVC.

Proteases

As proteases, também conhecidas como enzimas proteolíticas, são enzimas capazes de catalisar a degradação de proteínas através da hidrólise das ligações peptídicas. Elas são enzimas extremamente abundantes, e juntamente com seus homólogos correspondem a aproximadamente 2% dos genes codificados em alguns genomas (Rawlings e Barret, 1999; Puente *et al.*, 2003). Amplamente distribuídas em uma grande variedade de fontes, como plantas, animais e microorganismos, as proteases são vistas como importantes moléculas sinalizadoras que estão envolvidas no correto funcionamento da maioria das funções biológicas (Turk *et al.*, 1997; Turk, 2006). A atividade anormal dessas enzimas está direta ou indiretamente envolvida

com diversas patologias cardiovasculares e inflamatórias, câncer, osteoporose e desordens neurológicas (Leung-Toung, *et al.*, 2002; Rzychon *et al.*, 2004; Turk, 2006).

O envolvimento das proteases no ciclo de vida de organismos causadores de doenças as tornou um potencial alvo para o desenvolvimento de agentes terapêuticos contra doenças, como malária (Rosental *et al.*, 1993, 2002; Francis, *et al.*, 1997), Doença de chagas (Murta, *et al.*, 1990; Monteiro A. C. *et al.*, 2001), câncer e a AIDS (Turk, 2006). Além dos papéis fisiológicos, essas enzimas possuem uma grande aplicação dentro da indústria alimentícia, produtos de limpeza entre outros (Rao *et al.*, 1998).

Baseado no mecanismo de catálise do substrato pode-se subdividir as proteases em dois grupos maiores, chamados exopeptidases e endopeptidases. As exopeptidases clivam a cadeia polipeptídica próximo às terminações carboxi terminais (carboxipeptidases) ou N-terminal (aminopeptidases) do substrato. Já as endopeptidases clivam peptídeos distantes das áreas terminais do substrato, levando-se em consideração sítios de clivagem. Considerando ainda o mecanismo de ação elas podem ser subdivididas em grupos menores: serino, cisteíno, treonina, aspártico, glutâmico e metalo proteases (Rao *et al.*, 1998; Rawlings *et al.*, 2006; Turk, 2006).

Cisteíno proteases

Cisteíno proteases de vários tipos são encontradas em todos os reinos de organismos e podem ser agrupadas em pelo menos 70 famílias. Cada família contém proteínas com seqüências de aminoácidos similares conservadas evolutivamente, que refletem aos membros das famílias estruturas tridimensionais similares (Rawlings *et al.*, 2006). As propriedades físico-químicas destas proteases têm sido amplamente caracterizadas, entretanto suas funções biológicas ainda não foram completamente elucidadas.

As proteases desse grupo são comumente exemplificadas baseadas na Papaína, uma cisteíno protease isolada do látex de *Carica papaya* (Mitchell, *et al.*, 1970). A estrutura tridimensional comum a todas cisteíno proteases reportadas até o presente momento é formada por dois domínios, um composto basicamente por

hélices e o outro por fitas β . O sítio catalítico está localizado na reentrância desses dois domínios, sendo composto por aminoácidos provenientes dos dois domínios (Kamphuis, *et al.*, 1984; Rzychon, *et al.*, 2004). Os aminoácidos histidina e cisteína são diretamente responsáveis pela atividade proteolítica das cisteíno proteases. No caso de algumas famílias (família da papaína) uma asparagina também auxilia na formação do sítio ativo, orientando o anel imidazólico da histina durante a hidrólise (Kamphuis *et al.*, 1984; Menard *et al.*, 1995; Rzychon *et al.*, 2004).

Nas últimas décadas as cisteíno proteases têm sido consideradas importantes fatores de virulência de bactérias patogênicas, bem como responsáveis pelo desenvolvimento diversas de doenças (Sajid e McKerrow, 2002), tornando-se um alvo para o desenvolvimento de novas terapias (Croft *et al.*, 2005; Turk, 2006; Vicik *et al.*, 2006)

Xylellaína

O seqüenciamento total do genoma da *X. fastidiosa* 9a5c, revelou a presença de uma janela aberta de leitura com similaridade às cisteíno proteases. Esse gene foi clonado, expresso e a proteína caracterizada bioquimicamente como uma cisteíno protease, denominada Xylellaína (Nogaroto *et al.*, 2006).

Referências do capítulo.

- Bhattacharyya, A., Stilwagen, S., Ivanova, N., D'Souza, M., Bernal, A., Lykidis, A., Kapatral, V. *et al.* (2002). *Microbiology*, **99**: 12403-12408.
- Brlansky, R. H., Timmer, I. W., French, W. J. e McCoy, R. E. (1983). *Phytopathology*, **73**: 530-535.
- Chang, C. J., Garnier, M., Zreik, L. Rossetti, V. e Bove, J. M. (1993). *Cur. Microbiol.*, **27**: 137-142.
- Croft S. L.; Barrett, M. P. e Urbina, J. A. (2005). *Trends Parasitol.*, **21**: 508-512.
- Francis, S. E.; Sullivan, D. J. Jr. e Goldberg, D. E. (1997). *Annu. Rev. Microbiol.* **51**: 97-123.
- Hendson, M.; Purcell, A. H.; Chen, D.; Smart, C.; Guilhabert, M. e Kirkpatrick, B. (2000), *Appl. Environ. Microbiol.*, **67**: 895-903.

- Kamphuis, I. G.; Kalk, K. H.; Swarte, M. B. e Drenth J. (1984). *J. Mol. Biol.*; **179**: 233–256.
- Koide, T.; Zaini, P. A.; Moreira, L. M.; Vencio, R. Z.; Matsukuma, A. Y.; Durham A. M.; Teixeira, D. C.; El-Dorry, H.; Monteiro, P. B.; da Silva, A. C.; Verjovski-Almeida, S.; da Silva, A. M.; Gomes, S. L. J (2004), *Bacteriol.*, **186**:5442-9.
- Lambais, M. R.; Goldman, M. H. S.; Camargo, L. E. A. e Goldman, G. H. (2000). *Curr. Opin. Microbiol.*, **3**: 459-462.
- Leung-Toung, R.; Li, W.; Tam, T. F.; Karimian, K. (2002). *Curr. Med. Chem.*, **9**: 979-1002.
- Menard, R.; Plouffe, C.; Laflamme, P.; Vernet, T.; Tessier, D. C.; Thomas, D. Y. e Storer, A. C. (1995), *Biochemistry*; **34**: 464–471.
- Mitchell, R.; Chaiken, I. e Smith, E. (1970). *J Biol Chem.*; **245**: 3485–3492.
- Mollenhauer, H. H. e Hopkins, D. L. (1974). *J. Bacteriol.*, **119**: 612–618.
- Monteiro, A. C.; Abrahamson, M.; Lima, A. P. C. A.; Vannier-Santos, M. A. e Scharfstein, J. (2001). *J. Cell. Sci.*, **114**: 3933–3942
- Monteiro, P. B.; Teixeira, D. C.; Palma, R. R.; Garnier, M.; Bové, J. M e Renaudin, J. (2001). *Appl. Environ. Microbiol.*, **67**: 2263-2269.
- Murta, A. C. M., Persechini, P. M., Souto-Padron, T., de Souza, W., Guimarães, J. A. e Scharfstein, J. (1990). *Mol. Biochem. Parasitol.* 43, 27-38
- Nogaroto, V.; Tagliavini, S.; Gianotti, A.; Mikawa, A.; Barros, N.T.; Puzer, L.; Carmona, A.K.; Costa, P.; e Henrique-Silva, F. (2006). *FEMS Microbiol. Lett.*, **261**:187-93.
- Puente, X. S.; Sanchez, L. M.; Overall, C. M. e Lopez-Otin, C. (2003), *Nature Rev. Genet.* **4**: 544–558
- Purcel, A. H. e Hopkins, D. L. (1996). *Annu. Rev. Phytopathol.*, **34**, 131-151.
- Rao, M. B., Tanksale, A .P., Ghatge, M. S., Deshpande,V. V. (1998). *Microbiol. Mol. Biol. Rev.*,**62**:597-635.
- Rawlings, N. D. e Barrett, A. J. (1999). *Nucleic Acids Res.*, **27**: 325-331.
- Rawlings, N. D.; Morton, F. R. e Barrett, A. J. (2006). *Nucleic Acids Res.*, **34**: 270–272.
- Roberto, S. R.; Coutinho, A.; De Lima, J. E. O.; Miranda, V. S. e Carlos, E. F. (1996). *Fitopatol Brás.*, **21**: 517.

- Rosenthal, P. J.; Lee, G. K. e Smith, R. E. (1993). *J. Clin. Invest.*, **91**: 1052–1056.
- Rosenthal, P. J.; Sijwali, P. S.; Singh, A e Shenai, B. R. (2002). *Curr Pharm Des.*, **8**: 1659-1672.
- Rossetti, V. M., Garnier, J. M., Bové, M. J. G., Beretta, A. R. R., Teixeira, J. A., e Dagoberto de Negri, J. (1990). *C. R. Acad. Sci.*, **310**: 345-349.
- Rzychon, M.; Chmiel, D.; Stec-Niemczyk, J. (2004) *Acta Biochi. Pol.*, **51**:861-873.
- Sajid, M. e Mckerrow, J. H. (2002). *Mol. Biochem. Parasitol.*, **120**: 1-21.
- Simpson, A. J., Reinach, F. C., Arruda, P. *et al.*, (2000) *Nature*, **406**: 151-157.
- Souza Pinto, W. B.; Basile, G. B. e Gonzales, M. A., (2001) *CATI - Coordenadoria de Assistência Técnica Integral* - Secretaria de agricultura e Abastecimento do Estado de São Paulo.
- Souza, A. A.; Takita, M. A.; Coletta-Filho, H. D.; Caldana, C.; Golman, G. H.; Yanai, G. M.; Muto N. H.; De Oliveira, R. C.; Nunes, L. R. e Machado, M. A. (2003). *Mol. Plant Microbe Interact.*, **16**: 867-875.
- Turk, B. (2006) *Nat. Rev. Drug Discov.*, **5**:785-799.
- Turk, B.; Turk, V e Turk, D. (1997). *Biol. Chem.*, **378**: 141-50.
- Van Sluys, M. A., De Oliveira, M. C., Monteiro-Vitorello, C. B., Miyaki, C. Y., *et al.* (2003). *J. Bacteriol.*, **3**: 1018-1026.
- Van Sluys, M. A.; De Oliveira, M. C.; Monteiro-Vitorello, C. B.; Miyaki, C. Y., *et al.* (2003). *J. Bacteriol.*, **3**: 1018-1026.
- Vicik, R.; Busemann, M.; Baumann, K. e Schirmeister, T. (2006). *Curr. Top Med. Chem.* **6**: 331-353.
- Williams, P. H. (1980). *Plant Dis.*, **64**: 736-774.

Capítulo 2

Objetivos

O objetivo principal desse trabalho foi a determinação da estrutura cristalográfica da Xylellaína, uma cisteíno protease da bactéria fitopatogênica *Xylella fastidiosa*.

Objetivos específicos:

- Expressão e purificação da Xylellaína para ensaios de cristalização.
- Desenvolvimento de protocolos de cristalização da proteína Xylellaína recombinante.
- Determinação da estrutura cristalográfica da proteína
- Comparação estrutural da proteína com outros membros das cisteíno protease.

Capítulo 3

Metodologia

Neste capítulo são descritas as metodologias utilizadas para expressão e purificação da Xylellaína, ensaios de cristalização e resolução da estrutura tridimensional da enzima.

Estudos iniciais da Xylellaína.

Os trabalhos de clonagem, expressão e purificação da proteína Xylellaína foram iniciados pelo grupo do Prof. Flávio Henrique Silva (Departamento de Genética e Evolução da Universidade Federal de São Carlos) após seqüenciamento total do genoma da *X. fastidiosa* 9a5c (Simpson, *et al.*, 2000). O genoma desse fitopatógeno revelou a presença de uma janela aberta de leitura com similaridade as cisteíno proteases. O gene codificante da suposta cisteíno protease foi clonado no vetor de expressão pET28a (Novagem), produzindo uma proteína recombinante com uma cauda de 6xHIS no N-terminal, e as caracterizações bioquímicas comprovaram a Xylellaína como cisteíno protease diferencialmente expressa na cepa patogênica de *X. fastidiosa* 9a5c (Nogaroto *et al.*, 2006).

O plasmídeo pET28xylCP (Cisteíno Protease Xylellaína clonada no vetor de expressão **pET28a**) foi gentilmente cedido pelo grupo do grupo do Prof. Flávio Henrique Silva para os ensaios iniciais de cristalização.

Expressão e purificação da Xylellaína.

O plasmídeo recombinante pET28xylCP foi transformado em células *E. coli* BL-21 (DE3) competentes (Sambrook e Russel, 2000). A expressão da proteína Xylellaína nativa foi realizada como descrito por Nogaroto *et al.*, 2006, em meio líquido LB (Sambrook e Russel, 2000) e antibiótico kanamicina (25 mg/μl), partindo de uma diluição 1:100 de um inóculo inicial (colônia isolada contendo o plasmídeo

de interesse). A cultura bacteriana foi crescida a 37°C, sob agitação a 250 rpm até atingir DO_{600nm} igual a 0,5. A expressão da proteína recombinante foi induzida pela adição de IPTG 0,4 mM por 20 horas a 30° C. Já a expressão da proteína Xylellaina selenometionina substituída foi realizada pela inibição da via de biosíntese de metionina (Van Duyne *et al.*, 1993). Para tanto células de *E. coli* recentemente transformadas com o plasmídeo pET28xylCP foram crescidas em meio LB líquido até DO_{600nm} igual a 0,5. Após remoção do meio de cultura LB através de centrifugação a 4 °C, a massa foi ressuspensa em tampão PBS (Sambrook e Russel, 2002), seguida de nova centrifugação. A massa celular foi então inoculada em meio mínimo enriquecido de sais e vitaminas (BME Vitamins Solution - Sigma-Aldrich), a 20° C, 250 rpm por uma hora para adaptação celular ao novo meio de cultura. Após esse período foram adicionados: 100 mg/L dos aminoácidos lisina, fenilalanina e treonina; 50 mg/L dos aminoácidos valina e isoleucina; 60 mg/mL selenometionina e 0,1 mM de indutor IPTG. A indução foi realizada por 20 horas a 20° C.

As culturas foram centrifugadas e a massa celular ressuspensas em tampão de lise celular (Tris-HCl 10 mM, pH 8,0, 100 mM NaCl, 50 mM NaH₂PO₄). A lise celular foi realizada a 4° C com uso de lisozima (1 µL lisozima a 50 mg/mL para cada mL de tampão de lise celular) e pulsos de ultra-som (6 pulsos de 20 segundos com intervalos de um minuto entre os pulsos), seguida de centrifugação para separação das frações sobrenadante e precipitado.

A proteína recombinante presente na fração solúvel foi purificada por cromatografia de afinidade em colunas contendo resina de níquel Ni-NTA superflow (Qiagen), sendo eluída da mesma com gradiente de imidazol (Nogaroto *et al.*, 2006). Para remoção do imidazol e do excesso salino foi empregada uma coluna de desalinização (Hi-trap Desalting-GE-health care). As amostras protéicas foram analisadas em SDS-PAGE 15% (Laemmli, 1970), a dosagem protéica foi determinada através do método de Bradford (Bradford, 1976).

Cromatografia de exclusão molecular e espalhamento dinâmico de luz

A Xylellaina obtida após purificação por afinidade em resina de níquel e desalinização foi aplicada em coluna cromatográfica exclusão molecular em uma coluna xk16/20 utilizando resina superdex 200 (GE healthcare). A coluna foi pré-

equilibrada em tampão Tris-HCl 50 mM, pH 8,0 e 150 mM NaCl. A calibração da coluna de exclusão molecular foi realizada com a utilização de conjunto de proteínas padrão de baixo, médio e alto massa molecular (ferritina, aldolase, ovalbumina, catalase, albumina bovina e inibidor de tripsina de soja), seguindo concentrações proteicas sugeridas pelo fabricante.

As medidas de espalhamento dinâmico de luz (DLS), foram adquiridas utilizando-se o equipamento “DynaPro MS200” (Protein Solutions) a temperatura de 18° C com tempo de aquisição de 2,5 segundos, 30 aquisições por medida. Os dados foram analisados pelo programa “Dynamics V5.26” (Protein Solutions).

Ensaio de cristalização

Os ensaios iniciais de cristalização da proteína nativa foram realizados utilizando-se a metodologia de difusão de vapor em gotas suspensas (Hampel *et al.*, 1968; Begfors, 1999; Macpherson, 1999) a 4 e 18° C, com as condições dos conjuntos de solução de matriz esparsa “Crystal Screen I” e “II” (Hampton Research), empregando volumes iguais de proteína e solução de cristalização. A proteína recombinante foi concentrada por centrifugação em Amicon® Ultra-4 (membrana com corte de 10 kDa) até concentração de aproximadamente 7,0 mg/mL. Os ensaios de cristalização com a proteína selenometionina enriquecida foram realizados a 4 e 18° C, nas condições de cristalização da proteína nativa e do conjunto de solução de matriz esparsa “Nextal PEG suite” (Nextal-Qiagen). O refinamento das condições iniciais de cristalização foi obtida através de preparação de soluções variando-se concentração dos agentes precipitante e pH.

Coleta e processamento de dados de difração de raios X

As coletas de dados foram realizadas em temperatura criogênica (100 K). Os cristais de Xylellaina foram transferidos para soluções crio protetoras (solução de cristalização com 20 a 25% de etileno-glicol) e montados em laços de “Nylon” (Cryo loops – Hampton Research). Os dados de difração de raios-X dos cristais da proteína nativa e derivados de átomos pesados foram coletados em detector do tipo placa de imagem, modelo Mar345dtb (MAR Research), utilizando radiação K_{α} do cobre produzida em uma fonte convencional de raios X (ultraX 18 - RIGAKU/MS),

operando a 50 kV de tensão e 90 mA de corrente. A preparação de cristais derivados de átomos pesados foi realizada utilizando-se o método de “*quick cryo-soaking*” (Dauter *et al.*, 2000; Nagem *et al.*, 2001; 2003). O cristal de Xylelláina selenometionina derivado foi coletado na borda de absorção do selênio (peak- 0.923 Å), na linha de difração de raios-X X12C do Laboratório Nacional de Luz Síncrotron de Brookhaven (BNL, Upton, NY - EUA), em um detector do tipo CCD, modelo QUANTUM 4 (ADSC). Os conjuntos de dados coletados para proteína nativa foram processados com os programas “Mosflm” (Leslie, 1992) e “Scala” (Evans, 1997) do pacote “CCP4” (Collaborative Computational Project, Number 4, 1994), já o cristal selenometionina derivado foi processado com o pacote “HKL2000” (Otwinowski e Minor, 1997).

Resolução, refinamento e análise da estrutura tridimensional da Xylelláina

Tentativas iniciais de resolução da estrutura foram realizadas pelo processo de substituição molecular (Haas e Rossmann, 1970), utilizando-se os programas “AMoRe” (Navaza, 1994), “MOLREP” (Vagin e Teplyakov, 1997) e “PHASER” (Read, 2001). Os modelos para a substituição molecular foram preparados com o programa “CHAINSAW” (Collaborative Computational Project, Number 4), utilizando-se as coordenadas das estruturas cristalográficas de cisteíno proteases humanas (PDB 1NQC) e de *T. cruzi* (PDB 1AIM).

As posições iniciais dos átomos pesados no cristal foram encontradas pelos pacotes de programa “SHELX” (Schneider e Sheldrick, 2002) e as fases iniciais foram calculadas com o programa “SOLVE” (Terwilliger e Berendzen, 1999), seguido de modificação de densidade eletrônica e reconstrução automática do modelo com o programa “RESOLVE” (Terwilliger, 2000) O modelo inicial construído foi refinado ciclicamente com o programa “REFMAC5” (Collaborative Computational Project, Number 4) seguido de visualização dos mapas de densidade eletrônica e reconstrução manual no programa gráfico “COOT” (Emsley e Cowtan, 2004). Moléculas de água foram adicionadas e visualmente inspecionadas com programa “COOT”. As figuras e animações foram geradas com o programa “Pymol” (Delano, 2002).

Seqüenciamento do N-terminal da proteína Xylellaína

Cristais da proteína Xylellaína foram dissolvidos em água destilada e enviados para análise da seqüência de aminoácidos NH₂-Terminal, utilizando o método de degradação de Edman (Edman, 1950). O seqüenciamento do N-terminal foi realizado pelo grupo da Prof. Dr^a. Heloísa Selistre de Araújo (Centro de Biotecnologia Molecular e Estrutural, Departamento de Fisiologia, Universidade Federal da São Carlos).

Referências do capítulo

- Bergfors, T. M. (1999). *Protein Crystallization*, Publisher: International University Line.
- Bradford, M. M. (1976). *Anal. Biochem.*, **72**: 248-254.
- Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. D Biol. Crystallogr.*, **50**: 760–763
- Dauter, Z.; Dauter, M. e Rajashankar, K. R. (2000). *Acta Crystallogr. D Biol. Crystallogr.*, **56**: 232-237.
- Delano, W.L. The PyMOL Molecular Graphics System (2002). DeLano Scientific, Palo Alto, CA, USA. <http://www.pymol.org>
- Edman, P., *Acta Chem. Scand.* (1950) ,**4**:283.
- Emsley P. e Cowtan, K. (2004). *Acta Crystallogr. D Biol. Crystallogr.*, **60**: 2126-2132
- Evans, P. R. (1997). *Joint CCP4 and ESF-EACBM Newsletter*, **33**: 22-24.
- Haas, D. J. e Rossmann, M.G. (1970). *Acta Crystallogr., B*, **26**: 998 – 1004
- Hampel, A.; Labananskas, M.; Conners, P. G.; Kirkegard, L.; Raj Bhandary, U. L. e Sigler, P. B. (1968). *Science*, **162**: 1384-1387.
- Laemmli, U. K. (1970). *Nature*, **227**: 680-685.

- Leslie, A. G. W. (1992) *Jt. CCP4 ESF-EADBM Newsl. Protein Crystallogr.*, **26**: 27-33.
- McPherson, A. (1999). *Crystallization of Biological Macromolecules*, Publisher: Cold Spring Harbor Laboratory Press.
- Nagem, R. A. P.; Dauter, Z. e Polikarpov, I. (2001) *Acta Crystallogr. D Biol. Crystallogr.*, **57**: 996–1002.
- Nagem, R. A. P.; Polikarpov, I. e Dauter, Z. (2003) *Meth. Enzymol. Macrom. Cryst.*, **374**:120-137.
- Navaza, J. (1994). *Acta Crystallogr., A, Found. Crystallogr.*, **50**: 157-163.
- Otwinowski, Z. e Minor, W. (1997). *Meth. Enzymol.*, **276**: 307 – 326.
- Read, R. J. (2001) *Acta Crystallogr. D Biol. Crystallogr.*, **57**: 1373-1382.
- Sambrook, J. e Russel, D.W. (2000), *Molecular cloning: a Laboratory Manual*. 3^a ed. Cold Spring Harbor Laboratory Press.
- Schneider, T. R., Sheldrick, G.M. (2002). *Acta Crystallogr. D Biol. Crystallogr.*, **58**: 1772-1779.
- Simpson, A. J., Reinach, F. C., Arruda, P. *et al.*, (2000) *Nature*, **406**: 151-157.
- Terwilliger, T. C.(2000). *Acta Crystallogr. D Biol. Crystallogr.*, **56**: 965-972.
- Terwilliger, T.C., Berendzen, J. (1999). *Acta Crystallogr. D Biol. Crystallogr.*, **D55**: 849-861.
- Vagin, A. e Teplyakov, A. (1997). *J. App. Cryst.*, **30**: 1022-1025.
- Van Duyne, G. D.; Standaert, R. F.; Karplus, P. A.; Schreiber, S. L. e Clardy, J. (1993). *J Mol Biol.*, **229**:105-124.

Capítulo 4

Análise Estrutural e Discussões

Neste capítulo são abordados os resultados obtidos desde a cristalização da Xylellaína nativa até a resolução da estrutura tridimensional dessa enzima através da técnica de Dispersão Anômala a um único comprimento de onda, utilizando cristais de proteína selenometionina substituída. A análise detalhada da estrutura resolvida a 1.65Å e comparações com outras cisteíno proteases mostram características que a Xylellaína é expressa na forma de uma pró-proteína com a pró-região mais curta que a grande maioria das enzimas da família papaína. Um ribonucleotídeo Uridina di-fosfatado parece orientar o enovelamento da região N-terminal, bem como sua manutenção em conformação de inativação enzimática. Dois loops próximos ao sulco de ligação ao substrato também são observados na estrutura, um deles semelhante ao loop de oclusão da catepsina B.

Expressão, purificação e cristalização da Xylellaína.

Os ensaios iniciais de cristalização da Xylellaína nativa foram realizados pela aluna de iniciação científica Maria Amélia Vilela Oliva (Departamento de biologia, Universidade Federal de São Carlos) utilizando proteína purificada enviada pelo grupo do prof. Flávio Henrique Silva. Devido ao comportamento instável dessa proteína, optou-se por sua expressão e purificação em nosso laboratório. Para tanto, o plasmídeo pET28xylCP (Cisteíno **Protease Xylellaína** clonada no vetor de expressão **pET28a**) foi gentilmente cedido pelo Prof. Dr. Flávio Henrique Silva. Uma pequena alteração foi introduzida no protocolo original de expressão e purificação da proteína Xylellaína (Nogaroto *et al.*, 2006). Essa alteração consistia na substituição do passo de diálise após a purificação da proteína em resina de níquel por uma cromatografia em coluna de exclusão molecular G25 (Hitrap Desalting – Amershan Bioscience) para a remoção do excesso salino.

A expressão e purificação mostraram que a proteína era expressa em grandes quantidades na fração solúvel (figura II-4.1), e uma banda única com massa

molecular de aproximadamente 32 kDa era observada na eluição com 100 mM de imidazol (figura II-4.1). Uma banda de massa molecular levemente menor que a Xylellaína também foi observada durante todo processo de purificação, a qual diminui a partir da eluição com 50 mM de imidazol (figura II-4.1).

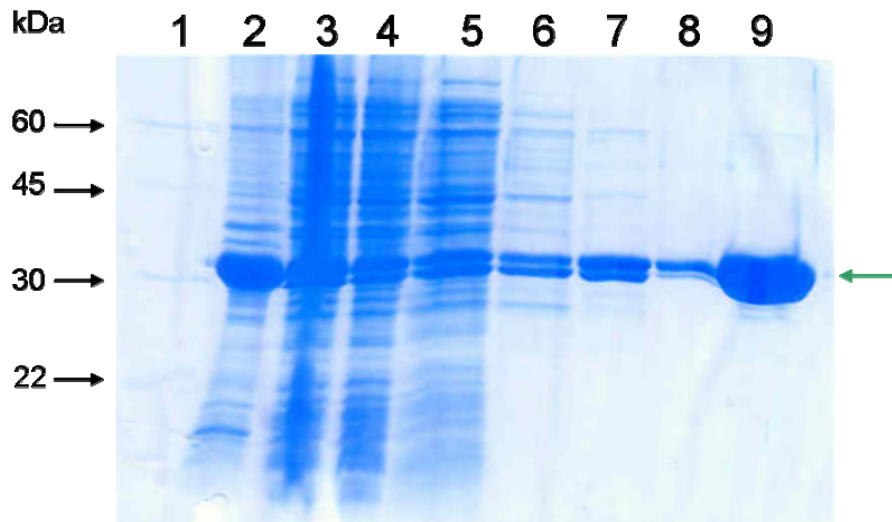


Figura II-4.1: Gel em poliacrilamida a 15% mostrando a expressão e purificação da Xylellaína em cromatografia de afinidade. 1 – Marcador de massa molecular, 2 e 3 – Fração insolúvel e solúvel da expressão a 30° C da proteína Xylellaína no vetor pET28a, 4 – Fração de proteínas aplicada na coluna de Níquel-NTA, 5 e 6 - Lavagem da coluna de afinidade com tampão de lise celular e 10 mM imidazol, 7 a 9 – Eluição da proteína Xylellaína com gradiente de imidazol (25, 50 e 100 mM). A seta verde mostra a Xylellaína, aproximadamente 32 kDa purificada.

A concentração salina (acima de 50 mM NaCl) e principalmente o imidazol atuam em conjunto precipitando a proteína poucas horas após sua eluição da coluna de níquel. então essa alteração objetivava acelerar o processo de preparo da amostra para os experimentos de cristalização, evitando a precipitação da proteína que era bastante comum após 24 horas do início da purificação. A proteína em tampão Tris-HCl 10 mM, pH 8,0, NaCl 50 mM, NaH₂PO₄ 50 mM foi concentrada inicialmente até aproximadamente 7,0 mg/mL em concentrador Amicon[®] Ultra-4.

Análises de espalhamento dinâmico de luz foram utilizadas para aferir a homogeneidade da amostra antes dos experimentos de cristalização. A montagem das condições de cristalização foi realizada pelo método de difusão de vapor com gota pendurada, utilizando volumes iguais de solução de mãe e proteína (gotas de 4 µL), à temperatura de 18° C.

Pequenos agregados cristalinos foram observados após duas semanas em duas condições de cristalização distintas. A primeira utilizando 30% de polietilenoglicol

4000 (PEG 4000), 100 mM citrato de sódio pH 5,5, 200 mM de acetato de amônio (condição 9) e a segunda com 20% de polietilenoglicol 8000 (PEG 8000), 100 mM cacodilato de sódio pH 6,5 e 200 mM de acetato de magnésio (condição 18) do conjunto de matriz esparsa “Crystal Screen I”. Esses aglomerados cristalinos iniciais foram cuidadosamente separados com auxílio de ferramentas de manipulação de cristais e submetidos a difração de raios-X, utilizando uma fonte convencional, disponível em nosso laboratório, modelo ultraX 18 (RIGAKU/MSU) com detector do tipo placa de imagem, modelo Mar345dtb (MAR Research), caracterizando-os como cristais de proteína porém com baixo poder de difração acima de 6Å.

O refinamento dessas condições de cristalização foi realizado alterando-se o pH, concentração dos agentes precipitantes e da proteína.

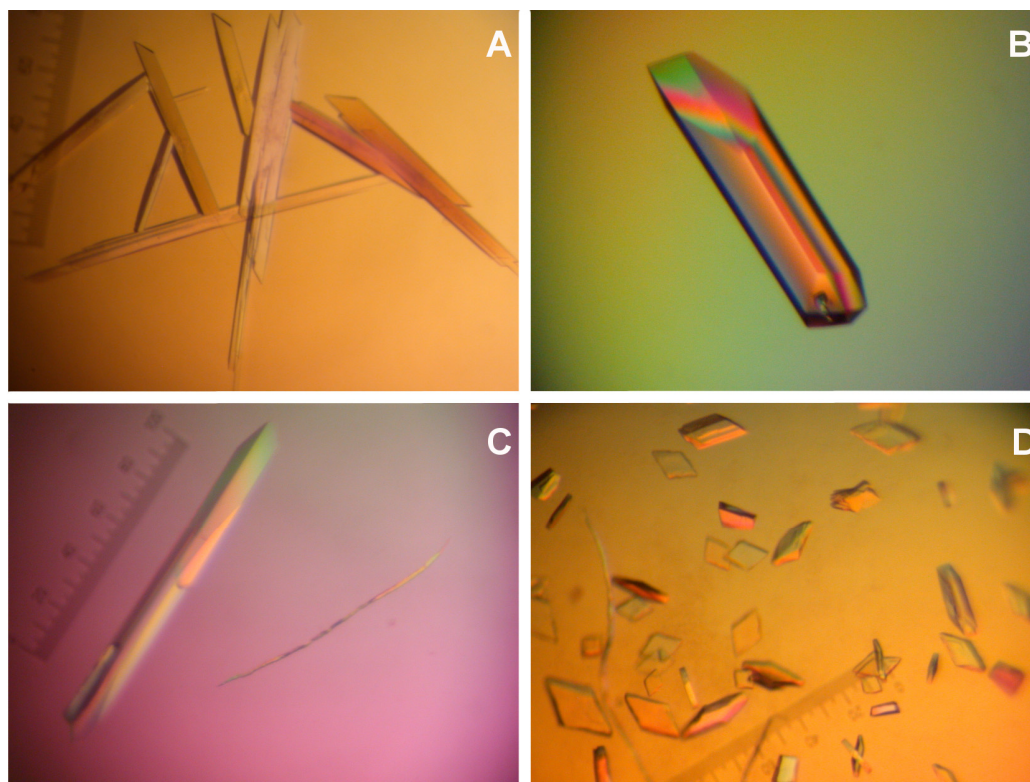


Figura II-4.2: Ensaio de cristalização da proteína Xylellaia. **A)** Condição com 15 % de PEG 4000, 100 mM citrato de sódio pH 5,6, 200 mM de acetato de amônio. **B e C)** Condição com 20 e 22 % de PEG 4000, 60 mM citrato de sódio pH 5,6, 134 mM de acetato de amônio. **D)** Condição com 25 % de PEG 4000, 60 mM citrato de sódio pH 5,6, 134 mM de acetato de amônio.

O refinamento da condição de cristalização 18 mostrou a formação de cristais em forma de pequenas agulhas. Os melhores resultados foram obtidos para o refinamento da condição de cristalização 9 (figura II-4.2), utilizando proteína na concentração de 7 mg/mL. Cristais com tamanho de aproximadamente 0,05-0,1 mm

foram observados nas condições de 20 a 22% de PEG 4000, 60 mM citrato de sódio pH 5,6, 134 mM de acetato de amônio (figura II-4.2B e 3.2C). Esses cristais apresentam grande reprodutibilidade nessa condição com crescimento em aproximadamente 48 horas.

A expressão da Xylellaína selenometionina substituída foi realizada pelo protocolo de inibição da via de biosíntese de metioninas (Van Duyne *et al.*, 1993), através do bloqueio de asparto kinases bacterianas, mediante altas concentrações dos aminoácidos isoleucina, lisina e treonina presentes no meio de cultura celular. A utilização de cepas bacterianas auxotróficas pode ser então descartada, de maneira que pequenas alterações no protocolo inicial de expressão são suficientes para produção de quantidades apropriadas de proteína recombinante (Doublíé, 2003).

As condições de indução previamente estabelecidas produziram proteína selenometionina substituída na fração insolúvel. A redução da temperatura de expressão para 20° C e da concentração do indutor para 0,1 mM propiciaram a produção da proteína na fração solúvel.

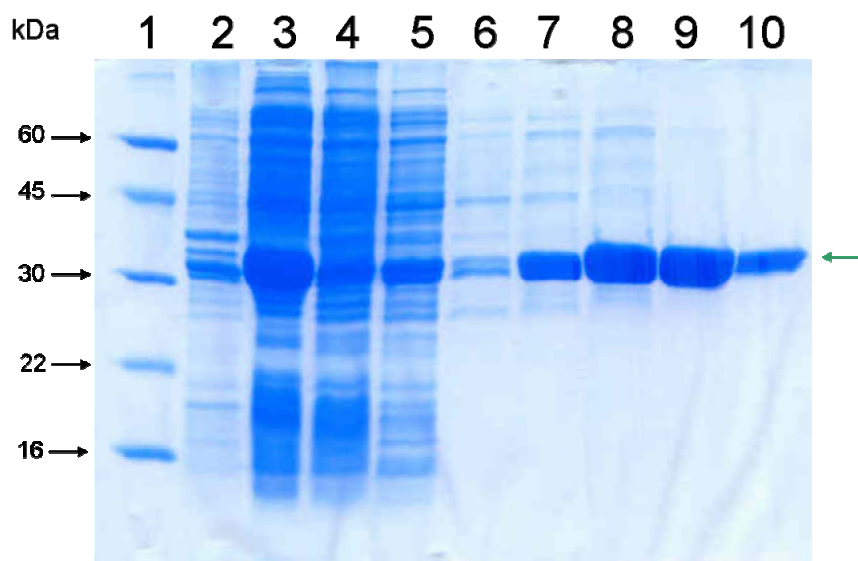


Figura II-4.3: Gel em poliácridamida a 15% mostrando a expressão e purificação da Xylellaína selenometionina substituída em cromatografia de afinidade, 1 – Marcador de massa molecular, 2 e 3 – fração insolúvel e solúvel da expressão a 20° C, 4 – Fração de proteínas aplicada na coluna de Níquel-NTA, 5 e 6 - Lavagens da coluna com tampão de lise e 10 mM Imidazol, 7 a 10 – Eluição da proteína Xylellaína com gradiente de imidazol (25, 50, 100 e 250 mM). A seta verde mostra a Xylellaína purificada.

A etapa de purificação foi realizada de maneira idêntica à proteína nativa, com produção aproximadamente de 8 mg de proteína por litro de cultura (figura II-4.3).

O tampão de eluição foi substituído para Tris-HCl 10 mM, pH 8,0, NaCl 50 mM, NaH₂PO₄ 50 mM, através de coluna de exclusão molecular G25, sendo então a proteína concentrada até aproximadamente 8,0 mg/mL em concentrador Amicon[®] Ultra-4. Os experimentos iniciais de cristalização foram realizados nas temperaturas de 4 e 18° C, utilizando método de difusão de vapor com gota pendurada com volumes iguais de solução de cristalização e proteína (gotas de 4µL) nas condições de cristalização próximas às obtidas para a proteína nativa. Nenhum cristal ou precipitado cristalino foi obtido nessas condições, mesmo com alterações na concentração de agentes precipitante e pH. Uma nova bateria de testes de cristalização foi realizada com o conjunto de soluções Nextal PEG Suite[®]. Um aglomerado cristalino não uniforme foi observado após 24 horas na condição de cristalização 32 deste conjunto (100 mM MES, pH 6,5 e 25% de polietilenoglicol 8000) (figura II-4.3 A).

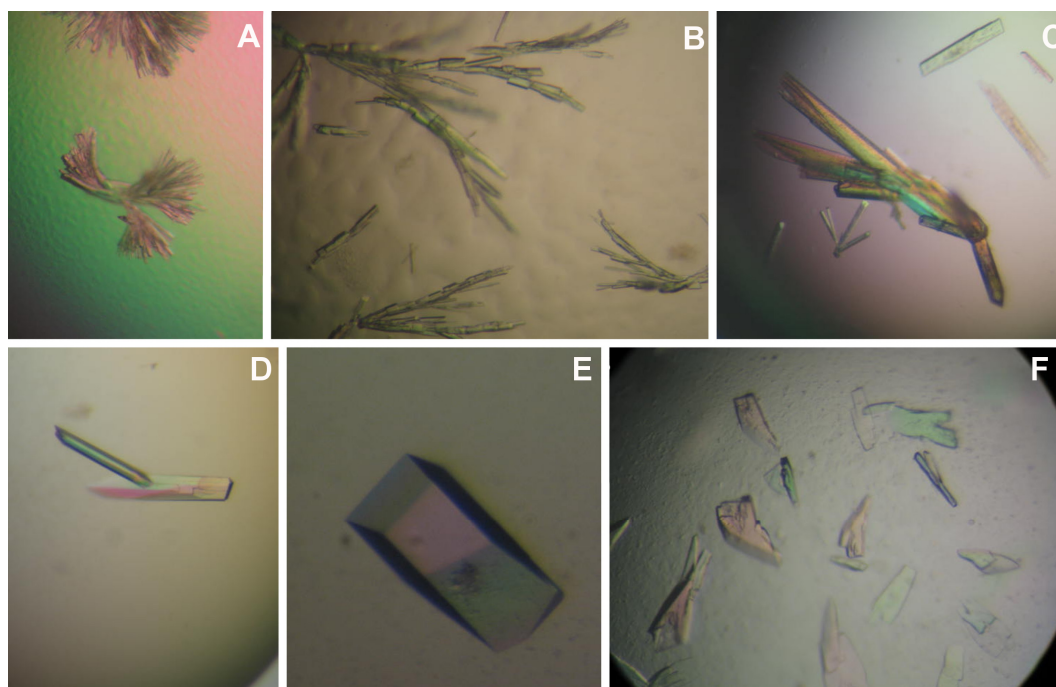


Figura II-4.4: Ensaio de cristalização da proteína Xylellaína selenometionina substituída. A) 100 mM MES, pH 6,5 e 25% de PEG 8000. B) 100 mM MES, pH 6,0 e 22% de PEG 8000, C) 100 mM MES, pH 5,7 e 20 % de PEG 8000 D e E) 100 mM MES, pH 5,6 e 17% de polietilenoglicol. F) 100 mM MES, pH 5,3 e 15 % de PEG 8000.

O refinamento dessa condição foi realizado com varredura da concentração de PEG 8000 de 15 a 30 % e com variação de pH de 5,2 até 6,5 (figura II-4.3 B a F). Cristais alongados e finos semelhantes aos obtidos para Xylellaína nativa (figura II-4.3D) e também cristais em formato de trapézio (figura I-4.3 E) foram encontrados

concomitantemente nas condições 17-18% PEG 8000 e pH 5,6-5,7. Diversos cristais provenientes dessa última condição foram selecionados e utilizados para coleta de dados.

Coleta e processamento de dados.

Os cristais da proteína obtidos foram crioprotetidos com solução de cristalização acrescida de etilenoglicol 15 a 20% (v/v) e congelados a 100 K em fluxo de nitrogênio. Esses foram testados utilizando o anodo rotatório disponível em nosso laboratório. Diversos cristais da proteína nativa foram testados, todos pertencentes ao grupo espacial P1 e com boa resolução.

A melhor coleta dados de um cristal nativo foi realizada com um cristal obtido em 18% de PEG 4000, 60 mM citrato de sódio pH 5,6, 134 mM de acetato de amônio a 18° C (figura II-4.2B). As imagens mostram padrão de difração bem definido e com a resolução de aproximadamente 1.6 Å (figura II 3.5 A e B). Dois conjuntos de dados foram coletados com esse cristal, o primeiro composto por 360 imagens foi coletado até 1.8 Å de resolução. Cada imagem foi adquirida com oscilação de um grau e tempo de exposição de 4 minutos. Já o segundo conjunto de dados foi coletado até 1.5 Å de resolução, com um total de 400, cada uma adquirida com oscilação de um grau e tempo de exposição de 6 minutos.

A integração das imagens de cada conjunto foi realizada com o programa “MOSFLM” (Leslie, 1992), no escalonamento e na redução dos dados os dois conjunto de dados foram unidos pelo programa “SCALA” (Evans, 1997) ambos parte do pacote de programas cristalográficos “CCP4” (Collaborative Computational Project Number 4 (1994)), gerando um conjunto de dados proveniente de duas coletas de dados. Os parâmetros e estatísticas de coleta de dados e processamento desse conjunto são mostrados na tabela a seguir (tabela 3.1).

Cristais da proteína nativa também foram utilizados para a obtenção de cristais isomorfos de átomos pesados, através da metodologia de “quick cryo soaking” e “soaking” com soluções de diversos sais de átomos como iodo, céσιο, platina, mercúrio, ouro, gadolínio. O resultado comum à maioria dos sais testados foi a quebra do cristal, mesmo quando pequenas concentrações salinas eram utilizadas. Os melhores resultados obtidos foram com as tentativas de “quick cryo soaking”.com

pequenas concentrações de iodeto de sódio (150 mM) por 30 segundos de incubação. Dois conjuntos de dados com alta redundância, provenientes de diferentes cristais foram coletados porém após processamento de dados, o sinal anômalo mostrou-se baixo ou inexistente impossibilitando o faseamento experimental.

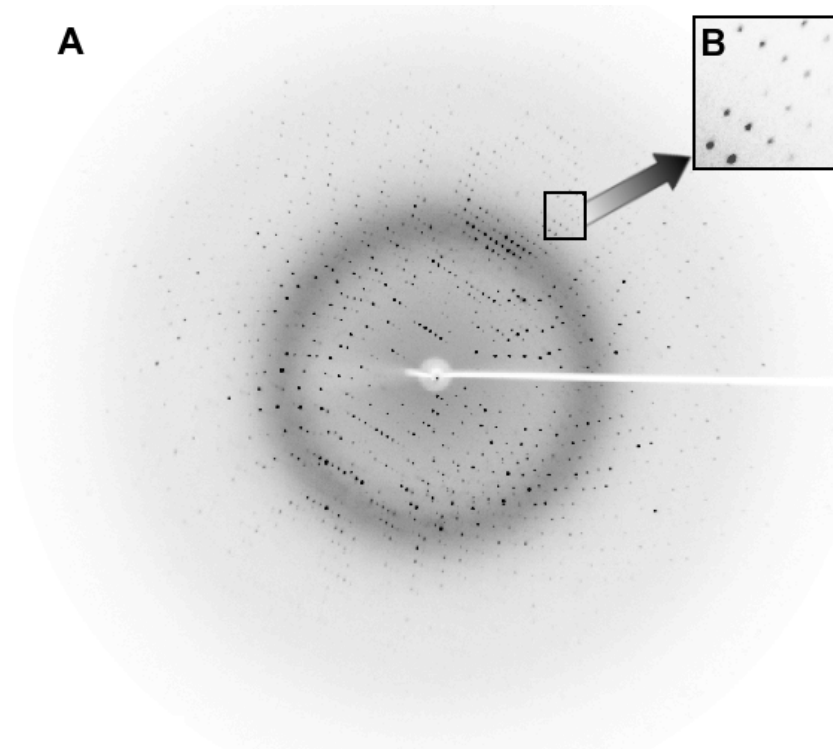


Figura II-4.5: Padrão de difração dos cristais nativos da Xylellaína. A) Imagem de padrão de difração do cristal da Xylellaína. B) Ampliação do padrão de difração.

Tabela II-4.1: Estatísticas de coleta e processamento de dados do cristal de Xylellaína nativo.

Grupo espacial	P1
Parâmetros da rede	a = 55,1 Å, b = 69,3 Å, c = 82,3 Å
	$\alpha = 75,8^\circ$, $\beta = 75,5^\circ$, $\gamma = 66,5^\circ$
Limites de resolução (Å)	23,10 – 1,65 (1,74 - 1,65)
Número de imagens	797
$\Delta\phi$ (°)	1
Mosaicidade (°)	1,02
Reflexões únicas	117070 (16308)
Multiplicidade	8,1(4,4)
R_{sym} (%)	5,4 (40,0)
$\langle I \rangle / \sigma(I)$	27,0 (3,2)
Completeza (%)	91,0 (86,7)

Valores em parênteses são referentes à camada mais externa de resolução

Os cristais de Xylellaína selenometionina derivados foram crioprotetidos com solução mãe acrescida de 15% de etilenoglicol e submetidos a experimentos de difração de raios X na linha X12B, Laboratório Nacional de Luz Síncrotron (Laboratório Nacional de Brookhaven, Upton, NY - EUA), durante o curso “RAPID

DATA 2006”. Experimento de fluorescência de raios-X mostrou a incorporação dos selênios na proteína e o máximo de absorção para os cristais foi obtido a um comprimento de onda de 0,9792 Å. Dois conjuntos de dados foram coletados, o primeiro utilizando-se os maiores cristais com a morfologia de trapézio (figura II-4.4 E) e o segundo com os cristais com morfologia semelhante aos da Xylellaína nativa (figura II-4.4 D).

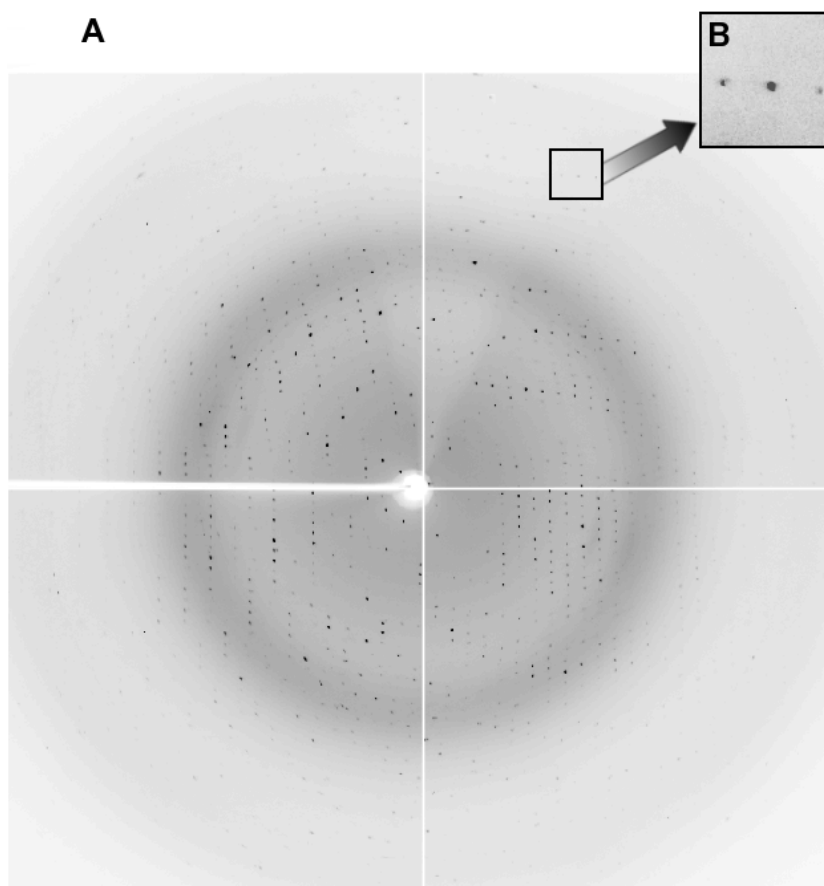


Figura II-4.6: Padrão de difração dos cristais Xylellaína selenometionina derivada. A) Imagem de padrão de difração do cristal da Xylellaína. **B)** Ampliação do padrão de difração mostrando pontos bem definidos.

O primeiro, de maneira diferente que os cristais nativos, pertencia ao grupo espacial $P2_1$ com parâmetros de célula $A=68.06\text{Å}$, $B=98.13\text{Å}$, $C=81.58\text{Å}$ e ângulos $\alpha = \gamma = 90^\circ$ e $\beta = 101.88$, com máximo de resolução em aproximadamente 2,5 Å e mosaicidade acima de 1.9° , um conjunto de dados com 360 imagens foi coletado, e os dados foram processados com o programa “HKL2000” (Otwinowski e Minor, 1997). Já o segundo possuía parâmetro de célula semelhante ao observado para estrutura nativa, além de apresentar menor mosaicidade e uma melhor resolução (figura II-4.6 A e B).

Inicialmente foram coletadas 396 imagens com 20 segundos de exposição e oscilação de 1°, após essa coleta o cristal foi re-orientado na cabeça goniométrica e mais 119 imagens foram coletadas. As estatísticas de coleta e processamento de dados estão sumarizadas na tabela a seguir (tabela II-4.2).

Tabela II-4.2: Estatísticas de coleta e processamento de dados do cristal de Xylellaína selenometionina derivado.

Grupo espacial	P1
Parâmetros da rede	a = 55,3 Å, b = 69,3 Å, c = 82,3 Å $\alpha = 75,8^\circ$, $\beta = 75,4^\circ$, $\gamma = 66,5^\circ$
Limites de resolução (Å)	50,00 – 1,83 (1,91 - 1,83)
Comprimento de onda (Å)	0,9792 (“Peak”)
Número de imagens	515
$\Delta\phi$ (°)	1
Mosaicidade (°)	0.952
Reflexões únicas	86291 (7651)
Multiplicidade	4,8(3,1)
R_{sym} (%)	5,4 (40,0)
$\langle I \rangle / \sigma(I)$	15,1 (2.5)
Completeza (%)	99,5 (97,8)

Valores em parênteses são referentes à camada mais externa de resolução

Resolução e refinamento da estrutura da Xylellaína.

As tentivas de faseamento inicial foram realizadas pela metodologia de substituição molecular (Haas e Rossmann, 1970). Os modelos de busca foram encontrados através de alinhamento seqüencial entre a Xylellaína e as proteínas depositadas no “PDB”. Os melhores alinhamentos foram obtidos para as cisteíno proteases catépsina S Humana (PDB-1NQC) e cruzaina (PDB-2AIM), porém essas apresentaram uma similaridade seqüencial baixa, aproximadamente 19% e 28 % respectivamente, quando comparados com a Xylellaína. Os modelos de busca foram preparados com programa “CHAINSAW” (Collaborative Computational Project, Number 4) e a substituição molecular foi realizada com o conjunto de dados da proteína nativa coletada até 1.65 Å de resolução. O número de moléculas presentes na unidade assimétrica foi estimado através do coeficiente de Matthews (Matthews, 1968) utilizando-se o programa “Matthews_coeff” (Collaborative Computational Project, Number 4).

Considerando-se a massa molecular da Xylellaína como sendo aproximadamente 32 KDa, estimou-se quatro moléculas como o valor provável da constituição da unidade assimétrica do cristal, resultando volume de Matthews equivalente a 2,3 Å³/Da e conteúdo de solvente no cristal de 45,1 %. Nenhuma

solução clara foi encontrada para qualquer programa utilizado. A inspeção visual do empacotamento cristalino das primeiras soluções obtidas nessas tentativas sistematicamente apontavam choques entre cadeias próximas. Ciclos iniciais de refinamento com o programa “REFMAC5” (Collaborative Computational Project, Number 4) não mostrou queda significativa nos valores de R_{factor} e R_{free} (valores da ordem de 0,45 – 0,50)

O problema das fases foi resolvido pelo método de dispersão anômala a um único comprimento de onda (SAD), utilizando os dados do cristal da Xylellaina selenometionina substituída. A posição dos sítios de selênio foi determinada com auxílio do pacote de programas “SHELX” (Schneider e Sheldrick, 2002) e “HKL2MAP” (Pape e Scheneider, 2004). Oito sítios claros de Selênio com ocupância acima de 80% foram encontrados no conjunto de dados. As fases iniciais foram calculadas utilizando as coordenadas dos átomos de selênio com o programa “SOLVE” (Terwilliger e Berendzen, 1999).

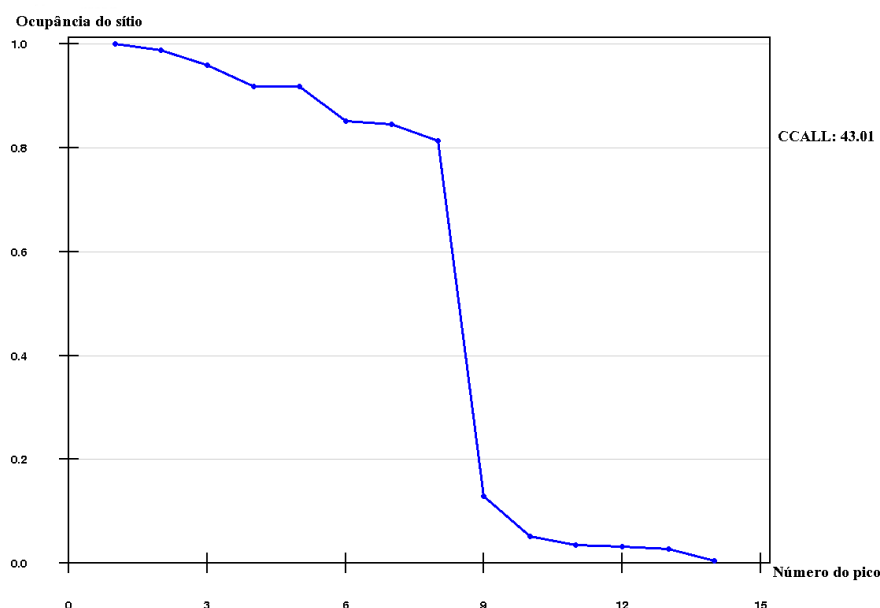


Figura II-4.7: Gráfico mostrando ocupância dos sítios de selênio.

A análise inicial mostrou que os mapas iniciais apresentam algumas características de elementos de estrutura secundária, além de algumas cadeias laterais para aminoácidos aromáticos. Os mapas de densidade foram otimizados através de ciclos de modificação de densidade e achatamento de solvente usando o programa “RESOLVE” (Terwilliger, 2000), seguido de construção automática do modelo inicial, também realizada pelo programa “RESOLVE”. Após o ciclo de reconstrução

automática, foi realizada a inspeção visual dos mapas e do modelo contruído. Os resíduos incoerentes com os mapas de densidade foram removidos, gerando um modelo inicial composto por fragmentos das quatro moléculas presentes na unidade assimétrica. As fases obtidas desse conjunto de dados foram combinadas com o conjunto de dados nativo utilizando o programa “CAD” (Collaborative Computational Project, Number 4), e juntamente com o modelo inicial foi realizado um novo ciclo de modificação de densidade e achatamento de solvente para a extensão das fases. O modelo construído após esse ciclo apresentou 665 resíduos dos 1164 esperados (420 com cadeia lateral e 245 resíduos sem cadeia lateral) com valor de R_{fator} igual a 43% e figura de mérito 0,55. Os mapas de densidade eletrônica mostraram melhora significativa e a partir desse momento a construção do modelo foi realizada interativamente com ciclos de manipulação no programa “COOT” e refinamento com o programa “REFMAC 5”. Moléculas de água foram introduzidas manualmente, com o auxílio do programa “COOT” durante o processo de refinamento, utilizando-se como critério distâncias mínima e máxima de 2,2 a 3,5 Å e fator de temperatura menor que 50 Å².

A análise do diagrama de Ramachandram (figura II 3.9), gerado pelo PROCHECK, mostra que 794 resíduos (88,3 %) encontram-se nas regiões mais favoráveis, 103 (11,5 %) estão em regiões adicionais permitidas, 2 resíduos, Tyr 33 cadeias B e C, (0,2 %) em regiões generosamente permitidas e nenhum na região desfavorável (exceto as glicinas). Os dados de refinamento são sumarizados na tabela II-4.3.

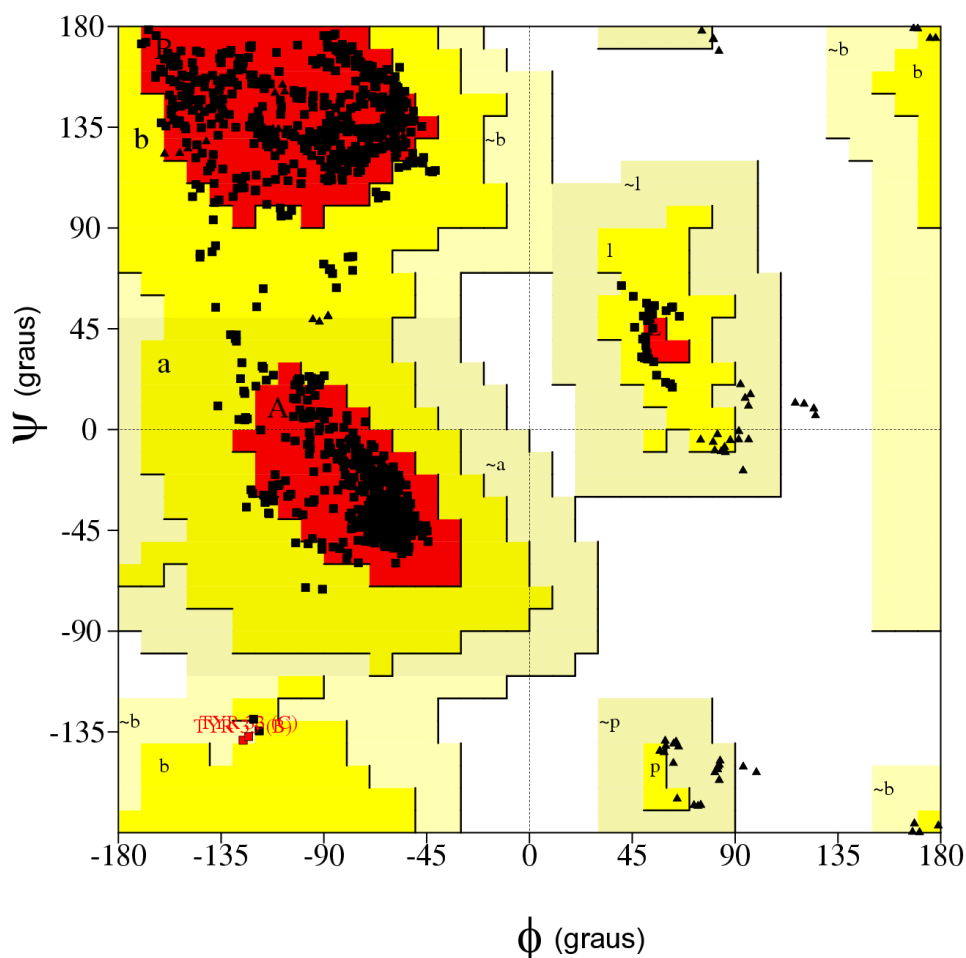


Figura II-4.8: Diagrama de Ramachandram da estrutura refinada da Xylellaína.

Tabela II-4.3: Estatísticas do refinamento da estrutura da Xylellaína.

Limites de resolução (Å)	23,10 – 1,65
Número de átomos	10504
R_{fator}	17,3
R_{free}	21,0
B_{factor} (Å ²)	22,1
Moléculas de água	1416
RMSD ângulos (°)	1.967
RMSD distância C-C (Å)	0.022

Estrutura cristalográfica da Xylellaína

A estrutura refinada é composta por 4 monômeros na unidade assimétrica (totalizando 899 aminoácidos), 4 Ribonucleotídeos di-fosfatados (UDP) e 1416 moléculas de água (figura II 3.8).

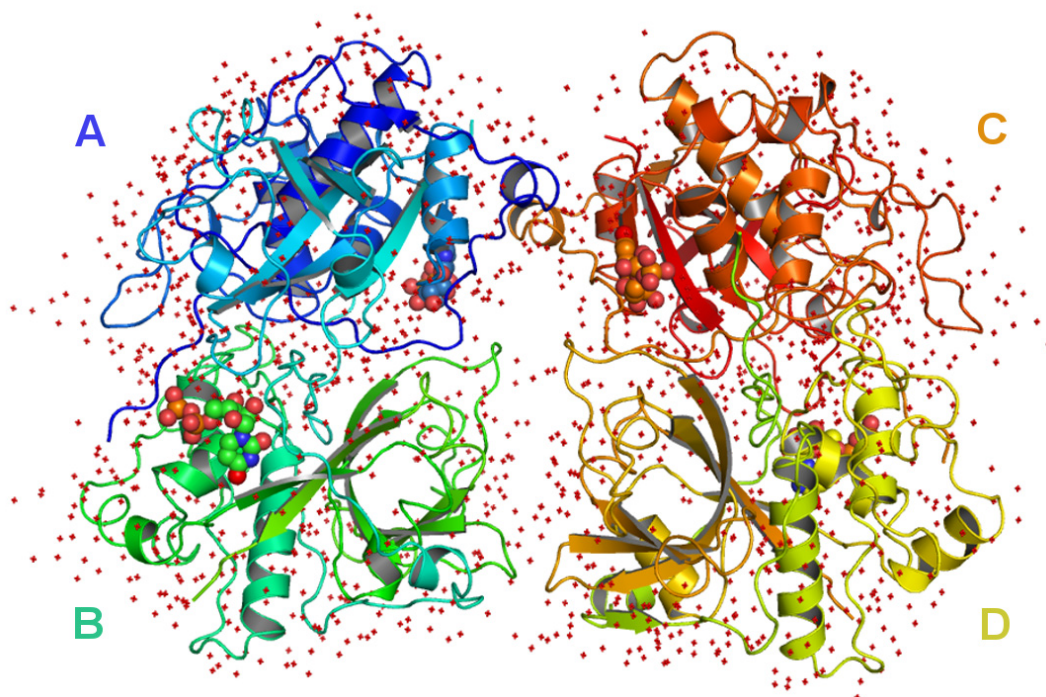


Figura II-4.9: Representação na forma de cartoons da estrutura refinada da Xylellaína. Quatro monômeros da proteína xylellaína, 1416 moléculas de água, um ribonuclotídeo difosfatado próximo à região N-terminal de cada monômero (representação de esferas).

Nenhuma densidade eletrônica foi observada para os 22 resíduos do N-terminal (MGSSHHHHHHSSGLVPRGSHMQT₂₃VLKR) da proteína recombinante nas cadeias A e D, e para os 23 resíduos iniciais das cadeias B e C, bem como para o último resíduo do C-terminal (lisina 271). O seqüenciamento N-terminal de cristais da proteína Xylellaína mostraram que essa região foi perdida durante o processo de cristalização ou então após sua purificação em coluna de afinidade. Como os 20 resíduos iniciais correspondem a artefatos de clonagem, somente os dois resíduos iniciais da proteína nativa foram perdidos. Metodologicamente os resíduos foram numerados considerando a metionina iniciadora da janela aberta de leitura como o resíduo número um.

Interface de dimerização da Xylellaína

A análise do empacotamento cristalino mostra que uma grande superfície de contato conecta os monômeros A com B e C com D (figura II-4.9). Uma análise pormenorizada da área de contato entre os monômeros, dos resíduos que realizam contatos polares ou hidrofóbicos, bem como sua acessibilidade ao solvente foi realizada com o auxílio do servidor PISA (Krissinel e Henrick, 2005). Uma das

ferramentas implementadas no PISA é o cálculo da probabilidade da interface analisada estar relacionada com a relevância biológica da proteína, nesse caso é atribuindo um valor entre zero e um para essa probabilidade.

Os contatos intermoleculares entre os pares de monômeros AB e CD no cristal são amplos formando uma superfície de aproximadamente 2018 \AA^2 (15,8% do total área acessível ao solvente na proteína). Valores acima de 1400 \AA^2 tem sido preditos como representativos de papel biológico da interface (Lo Conte *et al*, 1999). Os contatos polares envolvem pelo menos 19 resíduos de cada monômero formando interações de hidrogênio (13 resíduos) e ligações salinas (6 resíduos) (figura II-4.10). Os dados analisados pelo servidor sugerem que essa interface poderia ter alguma implicação biológica (valor da probabilidade igual a um). Com esse empacotamento uma característica importante a ser observada é que os monômeros estão com os sítios catalíticos praticamente no interior dos monômeros, um de frente para o outro e com uma distância de aproximadamente 23 \AA em relação às cisteínas catalíticas (figura II-4.11).

A grande maioria das cisteíno proteases atua como monômeros biológicos, uma exceção conhecida, catepsina C (Horn *et al*, 2002) que atua como tetrâmero. Durante os experimentos de espalhamento dinâmico de luz, realizados para aferir a homogeneidade da amostra da Xylellaína, o valor do raio hidrodinâmico calculado foi de $3,63 \text{ nm}$, valor condizente com uma proteína de aproximadamente $68,7 \text{ kDa}$, indicando estado dimérico da Xylellaína que possui massa molecular de aproximadamente 33 kDa . Já os dados de cromatografia de exclusão molecular mostram a proteína com massa molecular aparente de 51 kDa , sugerindo estado monomérico da proteína, que é um valor acima da massa molecular esperada para as cisteíno proteases.

Embora os dados da cromatografia de exclusão molecular não confirmem o estado dimérico da proteína um estudo aprofundado desse estado de oligomerização deve ser realizado uma vez que um dímero biológico semelhante ao observado no empacotamento cristalino seria justificável como uma maneira da proteína manter-se inativa, evitando a degradação celular.

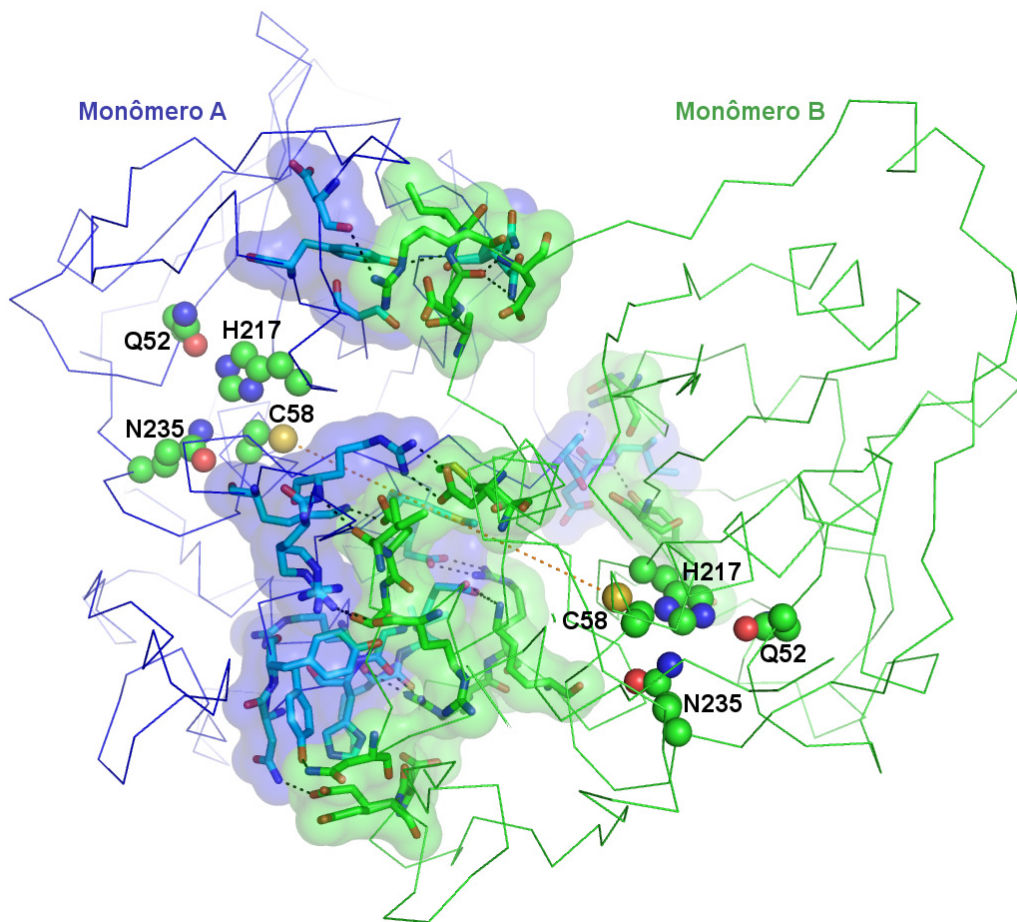


Figura II-4.10: Detalhe da superfície de contato entre monômeros A e B da Xylellaína. Representação de linhas dos carbonos α dos monômeros A (azul) e B (verde). Os resíduos que interagem na superfície da proteína são representados como “sticks”, com as respectivas interações de hidrogênio como linhas pontilhadas. As superfícies verde e azul mostram o contato entre os dois monômeros.

Enovelamento da Xylellaína.

A estrutura da Xylellaína mostra-se semelhante às cisteíno proteases da família papaína (C1), em especial as do grupo das catepsinas (cisteíno proteases lisosomais mamíferos), sendo composta por dois domínios característicos denominados classicamente como “L” (Left) e “R” (Right) (Drenth *et al.*, 1968). O domínio-L é formado por hélices e “loops”, enquanto o domínio-R é formado predominantemente por fitas β (figura II-4.11), na interface desses domínios é formando um sulco que se estende por toda a região superior da proteína no qual esta localizado o sítio ativo da enzima (figuras II-4.11 e 4.12).

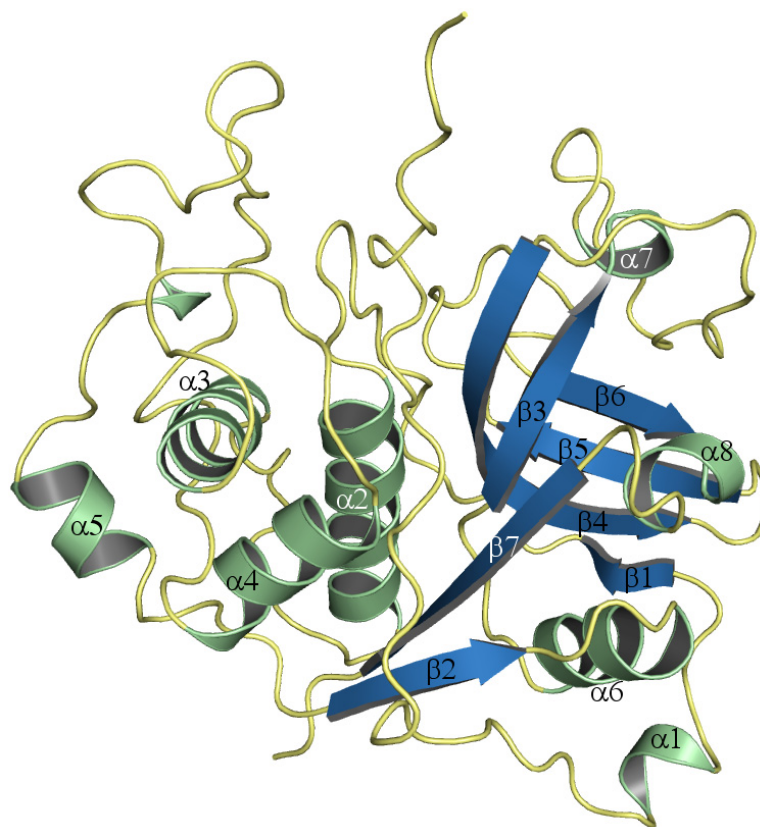


Figura II-4.11: Representação em cartoons da estrutura cristalográfica da Xylellaína. A estrutura da Xylellaína assume o enovelamento característico da família da Papaína, sendo composta pelos domínios L (Left domain) e R (Right domain). O domínio L é formado predominantemente por hélices e loops, já o domínio R é composto por fitas β (direita).

Sítio ativo da Xylellaína.

O sítio ativo característico das enzimas da família papaína é observado na estrutura tridimensional da Xylellaína. Os resíduos catalíticos cisteína (C58) e histidina (H217) são observados respectivamente na porção N-terminal da α -hélice 2 e na porção intermediária da fita β 4, formando a díade catalítica. Dois resíduos complementam o sítio ativo, uma glutamina (Q52) que precede a cisteína catalítica e é responsável pela formação do bolsão oxi-ânion durante a catálise da ligação peptídica e uma asparagina (N235) que direciona o correto posicionamento do anel imidazólico através de uma ligação de hidrogênio entre o $N\epsilon 2$ da histidina e o $O\delta 1$ da glutamina (figura II-4.12).

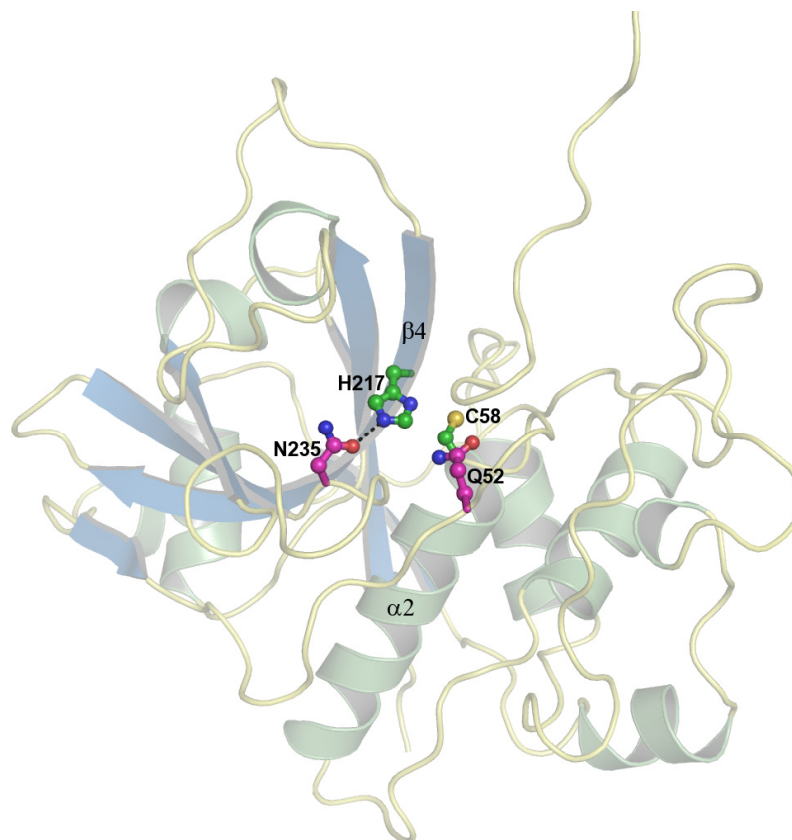


Figura II-4.12: Representação em cartoons da estrutura da Xylellaína mostrando em esferas e sticks os aminoácidos do sítio catalítico. O sítio ativo da enzima é formado na interface entre os domínios L e R. Díade catalítica formada pelos aminoácidos C58 e H217 (verde), e os aminoácidos complementares N235 e Q52 (magenta). Ligação de hidrogênio entre N235 e H217 mantém o N1 do anel imidazólico da histidina em uma orientação favorável a catálise.

Interação da região N-terminal da Xylellaína com o sulco de ligação ao substrato, mecanismo de inibição enzimática, pró-Xylellaína.

A regulação da atividade proteolítica está diretamente envolvida nos diversos mecanismos celulares. Esse controle é finamente ajustado por processos celulares que abrangem desde a regulação da expressão gênica, secreção, maturação e degradação da enzima, até o bloqueio de sua atividade por proteínas específicas (Wiederanders, 2003; Rzychon et al, 2004), assegurando que a enzima seja ativada quando sua função é requerida ou então inativada na situação oposta, geralmente em resposta a estímulos específicos. Exemplos de proteólise inadequada têm sido demonstrados em doenças neurodegenerativas, bem como no câncer (Turk, 2006).

Um mecanismo de inibição através de pró-região pode ser observado para estrutura cristalográfica da Xylellaína. Um segmento da região N-terminal composto por aproximadamente 37 aminoácidos inicia-se na forma de uma estrutura alongada acima da hélice α_2 , percorre o sulco que separa os dois domínios, envolvendo parte da superfície externa do domínio L (figura II-4.13).

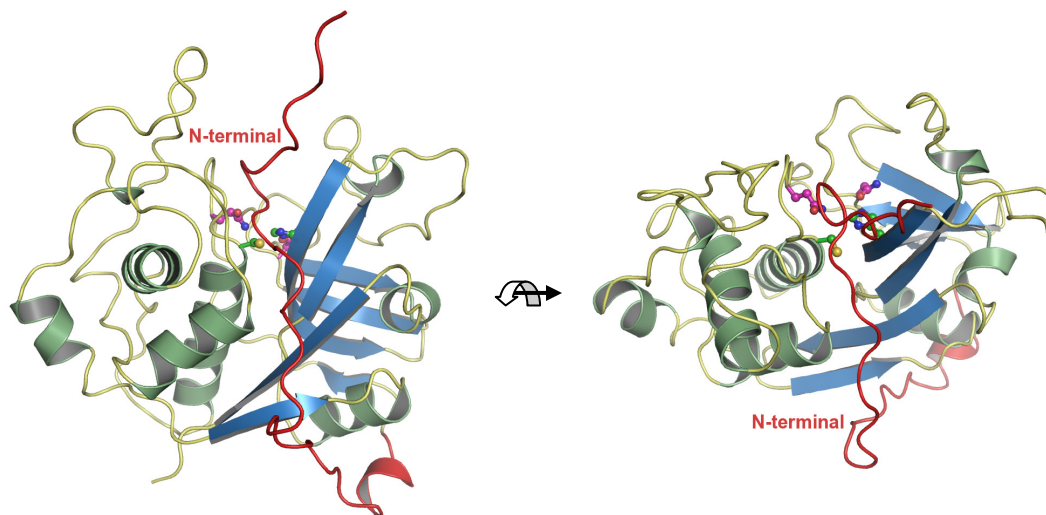


Figura II-4.13: Representação em “cartoons” da estrutura da Xylellaína mostrando porção N-terminal encobrindo o sítio catalítico. Sítio catalítico é representado como esferas e “sticks”.

O mecanismo de inibição presente na Xylellaína tem sido extensamente observado na grande maioria das cisteíno proteases da família da papaína (Cygler e Mort, 1997), nessas além de atuar inibindo a atividade enzimática a pró-região também está envolvida no correto enovelamento das enzimas (Tao *et al.*, 1994; Wang *et al.*, 2006). No caso da Xylellaína a remoção de 30 resíduos do N-terminal tornou a expressão da proteína completamente insolúvel (Prof. Dr. Flávio Henrique Silva, comunicação pessoal), mostrando que essa região também representa um importante mecanismo para o direcionamento do enovelamento protéico.

O alinhamento das seqüências das catepsinas mostra que a grande maioria dessas enzimas possui uma pró-região formada por segmentos de 60 a 100 aminoácidos com baixa identidade seqüencial (Cygler e Mort, 1997; Rzychon *et al.*, 2004), que compartilham alguns elementos de estrutura secundária.

A sobreposição da Xylellaína com as estruturas cristalográficas de algumas pró-proteínas da família papaína como as procatepsinas B (código PDB – 3PBH) (Turk *et al.*, 1996), K (código PDB – 7PCK) (Sivaraman *et al.*, 1999), L (código

PDB – 1BY8) (Lalonde *et al.*, 1999) e procatepsina-X (código PDB -1DEU) (Sivaraman *et al.*, 2000), mostram em todas as estruturas que a porção C-terminal da pró-região é composta por um segmento com estrutura secundária não definida de aproximadamente 20 aminoácidos. Essa região atravessa o sulco de ligação ao substrato formado na interface entre os domínios L e R, percorrendo a lateralmente a parte externa do domínio R até contactar-se com o mesmo em sua porção inferior. Na Xylellaína esse segmento N-terminal está compreendido entre a S11 e a L37 (figura II-4.13 e 4.14),

A porção N-terminal da pró-região da Xylellaína é completamente diferente da observada nas pró-catepsinas B, K e L. Na Xylellaína, ela é menor, formada por uma estrutura não ordenada de 10 resíduos que faz uma curva de aproximadamente 90° logo acima do sítio ativo, ficando praticamente paralela a hélice α_2 . Essa orientação provavelmente é resultante do empacotamento cristalino com as hélices α_3 e α_4 do monômero vizinho na unidade assimétrica. Nas estruturas de pró-catepsinas B, K e L essa região é maior, variando de 74 aminoácidos (procatepsina K) a 40 (procatepsina B). Uma pequena hélice- α ancorada nos sítios S1' e S2' (acima do sítio ativo) conecta o segmento C-terminal da pró-região a um motivo estrutural do tipo “grampo de cabelo” formado entre a hélice- α N-terminal com uma fita β antiparalela (figura II-4.14). Nas procatepsinas K e L uma pequena hélice- α adicional esta empacotada contra a hélice do “grampo de cabelo” (figura II-4.14). Interações de Vander walls entre as cadeias laterais do “grampo de cabelo” e a superfície da proteína mantém essa estrutura ligada à proteína, formando uma estrutura denominada PBL (Proregion Binding Loop) (Cygler *et al.*, 1996).

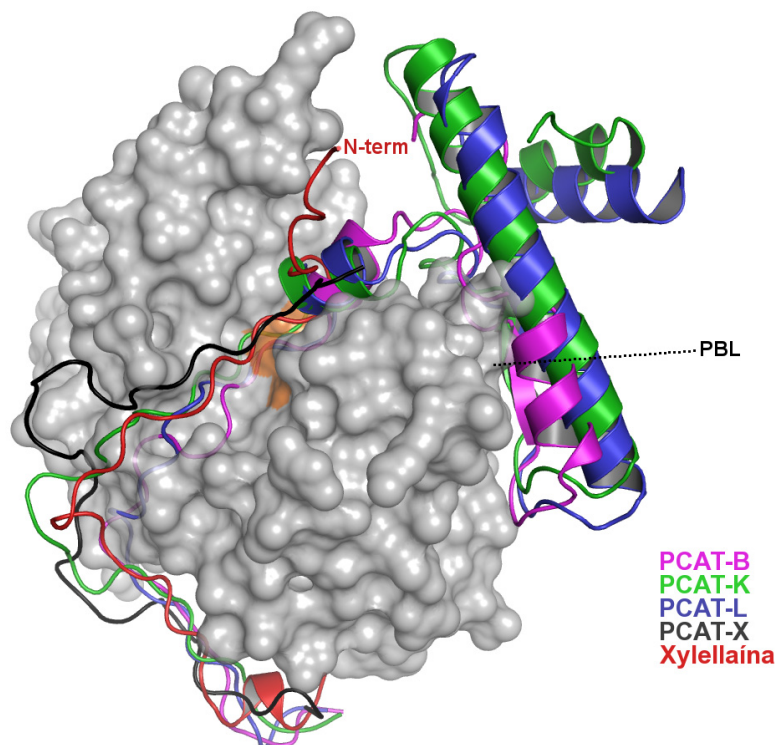


Figura II-4.14: Sobreposição das estruturas de pró-catepsinas com a Xylellaína. A cadeia principal das pró-catepsinas foi removida. Em cinza superfície da Xylellaína, estruturas em cartoons mostram a pró-região das proteínas, em magenta a pró-catepsina B humana, em verde a pró-catepsina K humana, em preto, pró-catepsina X humana e em vermelho a Xylellaína. Superfície em alaranjado mostra posição do sítio ativo. PBL é o “loop” de ligação do pró-domínio.

A interface de ligação entre a pró-região da Xylellaína e o restante da a enzima envolve principalmente os resíduos 11 ao 18 (figura II-4.15), as cadeias laterais dos aminoácidos S11, Y13, Y15, I16, D18 interagem diretamente com os sítios S2', S1', S1, S2 e S3 respectivamente (Schester e Biger, 1967) (figura II-4.15), fazendo a cadeia principal cobrir competamente o sulco. O modo de ligação entre as cadeias laterais da pró-região e o sulco de ligação ao substrato é semelhante ao esperado complexo enzima-substrato, mas a direção da cadeia principal da pró-região é feita de maneira oposta. Dessa maneira a porção N-terminal fica localizada no bolso catalítico S1' enquanto a porção C-terminal no sítio S1, impedindo a hidrólise da ligação peptídica (Wiederanders, 2003).

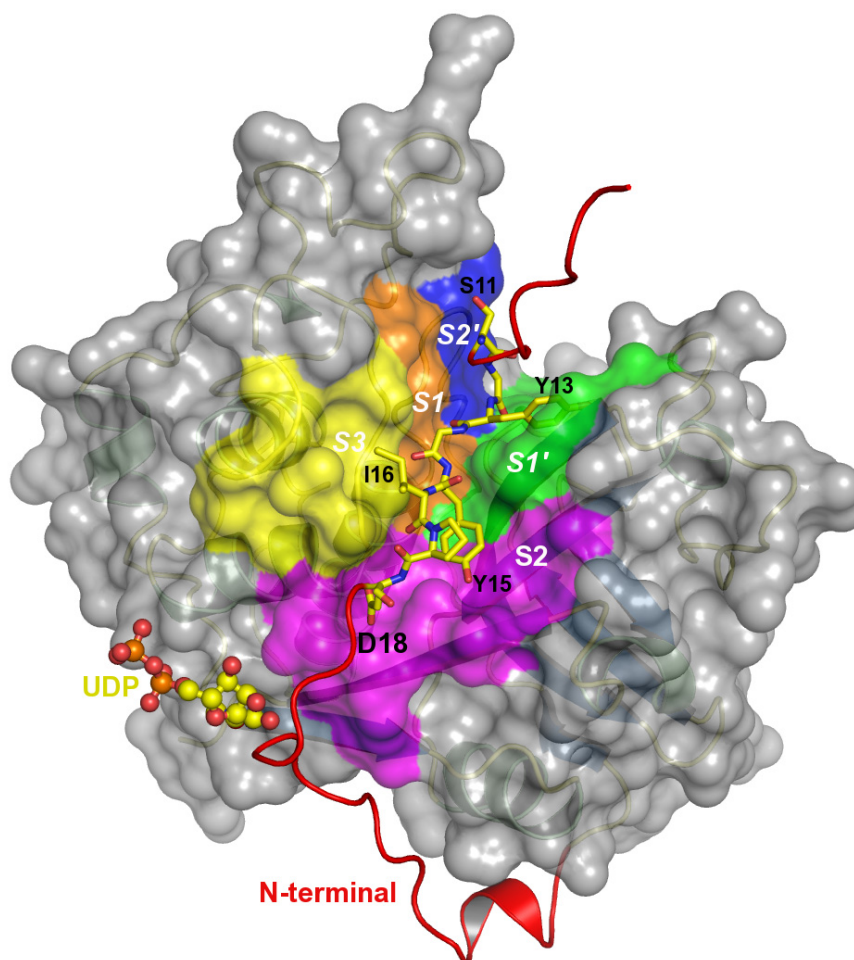


Figura II-4.15. Representação da superfície da Xylellaína, mostrando interação da pró-região com a proteína. Superfície da provável proteína madura é representada com os sítios de ligação ao substrato S1', S2', S1, S2 e S3 coloridos. A pró-região é representada em “cartoon” vermelho, em “sticks” aminoácidos que interagem com o sítio de ligação ao substrato. Molécula de UDP em esferas e “sticks”.

Sítio de ligação do ribonucleotídeo uridina di-fosfato

Durante o processo de refinamento uma região de densidade eletrônica não esperada foi observada próxima da hélice $\alpha 4$ e fita $\beta 3$ (figura II-4.16 B). Dois ribonucleotídeos difosfatados contendo uma base pirimidínica, citosina ou uracila, foram modelados satisfazendo completamente a densidade eletrônica. A escolha do ribonucleotídeo foi realizada através da análise das interações entre a proteína e o ligante, que foram satisfeitas pela Uridina (figura II-4.16 A-C).

O bolsão de ligação ao Uridina di-fosfato (UDP) é formado por oito aminoácidos, quatro provenientes da pró-região da Xylellaína, D21 e D24, R23 e

F25, e os demais aminoácidos, R107, K111 e K115 e Y164, provenientes da hélice $\alpha 4$ (três primeiros) e fita $\beta 3$ (último). Este bolsão está localizado pouco abaixo da porção C-terminal do sulco de ligação ao substrato (figura II-4.15). Os resíduos da pró-região formam parede externa do sítio de ligação de modo que a R23 forma a lateral do sítio, que acomoda a ribose do ribonucleotídeo. A cadeia lateral da R23 é estabilizada por ligações de hidrogênio entre o N ϵ 1 e NH1, com o OH1 e OH2 do D21 respectivamente. A região N-terminal é aproximada da hélice $\alpha 4$ por duas ligações de hidrogênio entre o D23 com a R107, formando assim a porção distal do sítio. A F25 compõe a parte superior do sítio. As cadeias laterais da F25 e da Y164 estão paralelamente dispostas respectivamente acima e abaixo da base nitrogenada uridina, em uma espécie de empacotamento do tipo empilhamento semelhante ao observado com as bases nitrogenadas do DNA. Interações de hidrogênio são observadas entre os O2 e O4 da base nitrogenada com os nitrogênios da cadeia principal da F25 e Y164 respectivamente, e também entre o N3 da base com o oxigênio da cadeia principal da Y164, mantendo o UDP firmemente ligado à proteína (figura II-4.16 A-C).

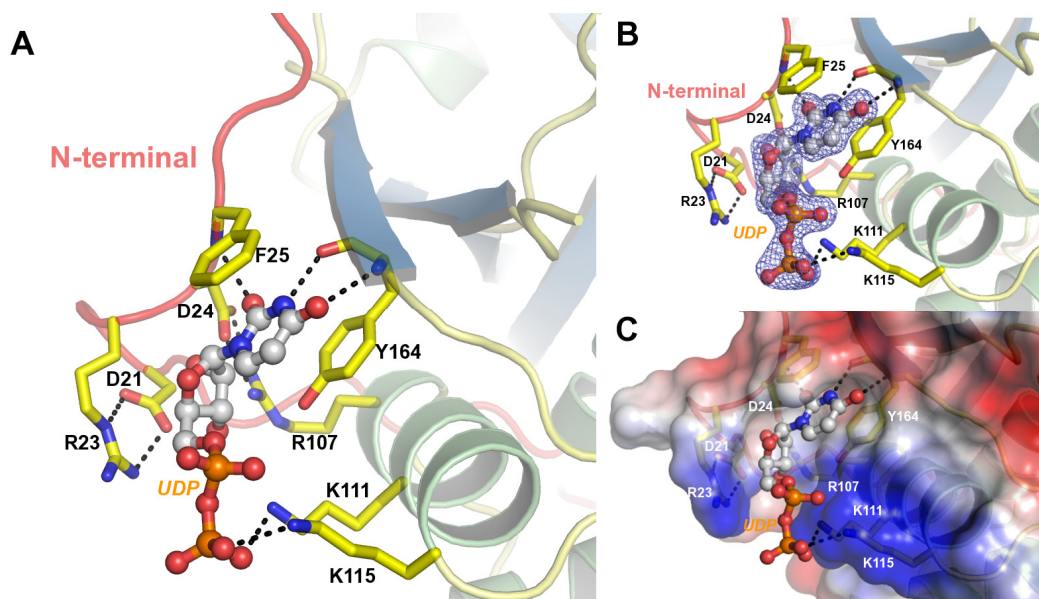


Figura II-4.16: Região de ligação ao ribonucleotídeo UDP. A) Detalhe da região de ligação ao UDP. B) Região de ligação ao UDP, mapa de densidade eletrônica (2Fo-Fc) contornando a 1 σ . C) Superfície potencial do bolsão de ligação ao UDP mostrando disposição de cargas.

Os resíduos K111 e K115 fazem interações de hidrogênio com os oxigênios do grupamento fosfato (figura II-4.16 A-C), a representação da superfície potencial mostra uma grande região positiva formada por esses dois resíduos juntamente com

os nitrogênios da cadeia lateral da R23, permitindo interação com os grupos fosfatos negativamente carregados do UDP (figura II-4.16 C).

Um dos possíveis papéis para esse nucleotídeo na estrutura da Xylellaína seria o seu envolvimento no direcionamento e estabilização da pró-região. Nas pró-catepsinas com estrutura resolvida a pró-região é consideravelmente mais longa que a Xylellaína, formando uma espécie de “grampo”, com a sua parte central localizada sobre o sulco formado pelos domínios L e R, assim o sítio ativo da enzima fica completamente obstruído (figura II-4.14). Nessas estruturas, o PBL atua fechando a porção N-terminal da pró-região próxima ao sítio ativo, através de interações com a superfície da proteína, enquanto do outro lado a pró-região está covalentemente ligada. Com esse mecanismo, esse grampo está firmemente ligado a proteína mantendo a mesma na sua configuração inativa. Na estrutura cristalográfica da pró-catepsina X, que da mesma maneira que a Xylellaína não possui essa região N-terminal longa como as demais procatepsinas, um interessante mecanismo conecta a porção N-terminal do pró-domínio com a proteína (figura II-4.17A). Nessa proteína, a cisteína catalítica (C31), presente na hélice α_2 , forma uma ligação de dissulfeto com a pró-região (C28p), a qual mantém a mesma ligada ao sulco formado entre os domínios L e R, compensando o efeito da perda da porção N-terminal comum a grande maioria das pró-catepsinas. A pró-região da Xylellaína de maneira semelhante à procatepsina-X, também não faz interação com o PBL (figura II 3.17A), a sobreposição das duas estruturas mostra que em Xylellaína a posição semelhante a C28p da catepsina é ocupada pela G14, o que exclui alguma interação direta com a cisteína catalítica. Entretanto a localização do UDP na região aonde o N-terminal da Xylellaína faz uma curva para entrar no sulco do sítio ativo (figura II 3.17A), somada à intensa relação do ribonucleotídeo com a proteína torna essa região menos flexível. Esse sítio de ligação ao ribonucleotídeo poderia então impedir um grande deslocamento de todo o segmento anterior, funcionando como mecanismo compensatório para a perda da região N-terminal comum as catepsinas.

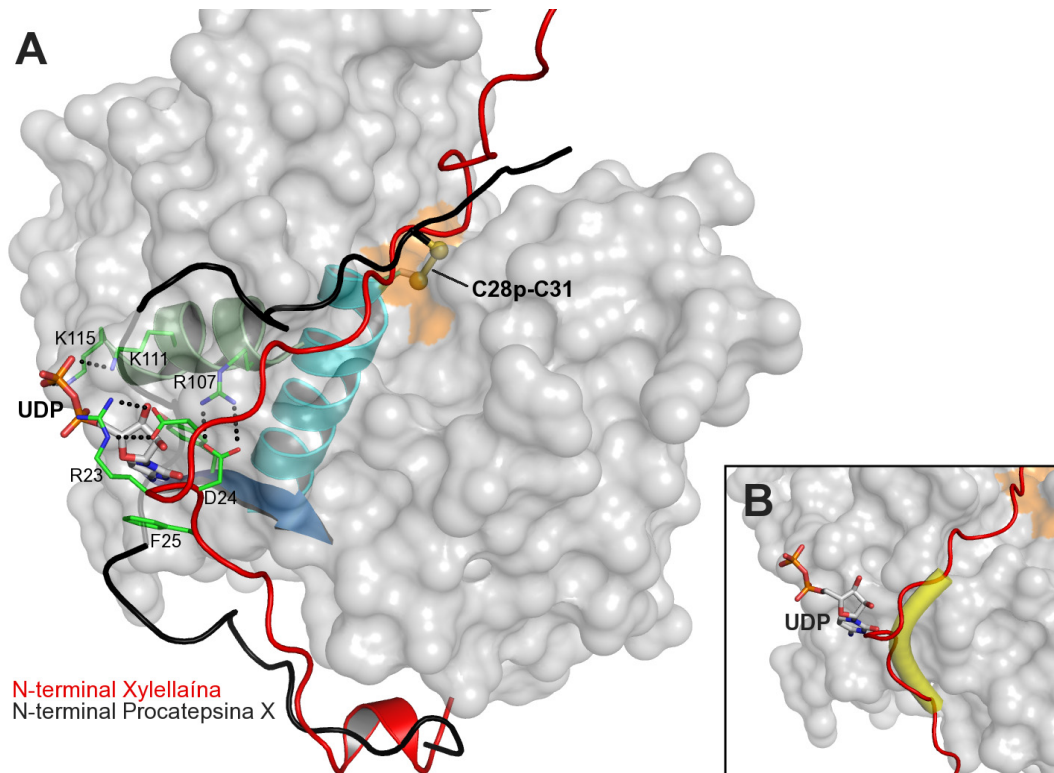


Figura II-4.17: Superfície da Xylellaína com sobreposição da pró-região da pró-catepsina X e da Xylellaína. A) Representação em cartoons mostra pró-região das duas enzimas, preto a pró-catepsina e vermelho a Xylellaína. Interações entre o UDP e as cadeias laterais dos aminoácidos formam o bolsão de ligação ao UDP. Ligação de dissulfeto entre a C28P com a C31 é mostrada em “sticks”. **B)** Detalhe do posicionamento do nucleotídeo na região de onde a porção-terminal faz uma curva (amarelo).

Marcadamente as cisteíno proteases da família papaína apresentam atividade de endopeptidases, algumas exceções a essa regra geral podem ser explicadas por alguns elementos estruturais característicos como o “loop” de oclusão na catepsina B e mini “loop” na catepsina X, que conferem a estas, atividade de endo e exopeptidase (Musil et al, 1991, Illy *et al.*, 1997, Nagler et al, 1999, Klemencic *et al.*, 2000).

Os ensaios de atividade enzimática da Xylellaína (Prof. Dr. Flávio Henrique Silva, comunicação pessoal) mostram que essa enzima possui uma pequena atividade de exopeptidase. Na estrutura da Xylellaína dois “loops” são observados na extremidade S’ do sulco de ligação ao substrato, o “loop” S’ (D128-A144) e o mini “loop” S’ (V195-L200). A sobreposição das estruturas da catepsina B com a Xylellaína mostra que os pontos de inserção do “loop” S’ são compatíveis com as do “loop” de oclusão da catepsina B, porém na xylellaína esse é mais curto (17 aminoácidos contra 21 da catepsina B) e ocupa uma posição aberta e próxima ao domínio-R (figura II 3.17). Com isso essa estrutura forma uma espécie de extensão

acima do sulco de ligação ao substrato, enquanto no caso de catepsina B o “loop” de oclusão está voltado para dentro do sulco. Já o mini “loop” S’ está em posição oposta, próximo ao sítio S’ (figura II 3.17).

As estruturas cristalográficas da catepsina B de rato (Jia *et al.*, 1995) e da Humana (Turk *et al.*, 1995), complexadas com inibidores irreversíveis, mostram que duas histidinas sequenciais (H111 e H112) presentes no “loop” atuam como aceptores negativos no sítio S2’ (figura II-4.18), sendo aptas a receber o carboxilato do resíduo C-terminal P2’, auxiliando na atividade de exopeptidase. Um mecanismo semelhante pode ser observado na estrutura da catepsina X. Um pequeno “loop” de 4 aminoácidos inserido após a Q22, posiciona a H23 em uma região próxima ao sítio S2’ (figura II-4.18), que por um mecanismo de rotação ao longo do carbono beta da H23 posiciona a mesma em uma orientação idêntica a H111 da catepsina B (Klemencic *et al.*, 2000), permitindo assim a atividade de exopeptidase. As sobreposições também mostram que a estrutura do mini “loop” S’ da estrutura da Xylellaína encontra-se próxima da porção final do “loop” de oclusão, com a R128 com sua cadeia lateral voltada para o interior sítio S1’, posição próxima à localização da H111 da catepsina B.

Como a grande maioria das cisteíno proteases, a Xylellaína mostra atividade endopeptidase (Nogaroto *et al.*, 2006), sendo caracterizado uma pequena atividade de exopeptidase até o momento. A presença desses dois loops não conservados e próximos à região do sítio ativo sugere que os mesmos podem estar relacionados com a interação com o substrato. Um rearranjo do “loop” S’ poderia trazer o mesmo a uma conformação semelhante à do “loop” de oclusão presente em catepsina B. Embora nenhum resíduo de histidina seja encontrado na sequência do “loop” S’ alguns resíduos carregados como ácido glutâmico e arginina estão presentes nesse loop, em especial a R134 poderia ter um papel semelhante as H111 da catepsina B e a H23 na catepsina X. Novos ensaios enzimáticos serão realizados com a Xylellaína para confirmar essa atividade de exopeptidase, e experimentos de mutagênese poderão confirmar o papel desses loops na referida atividade da enzima.

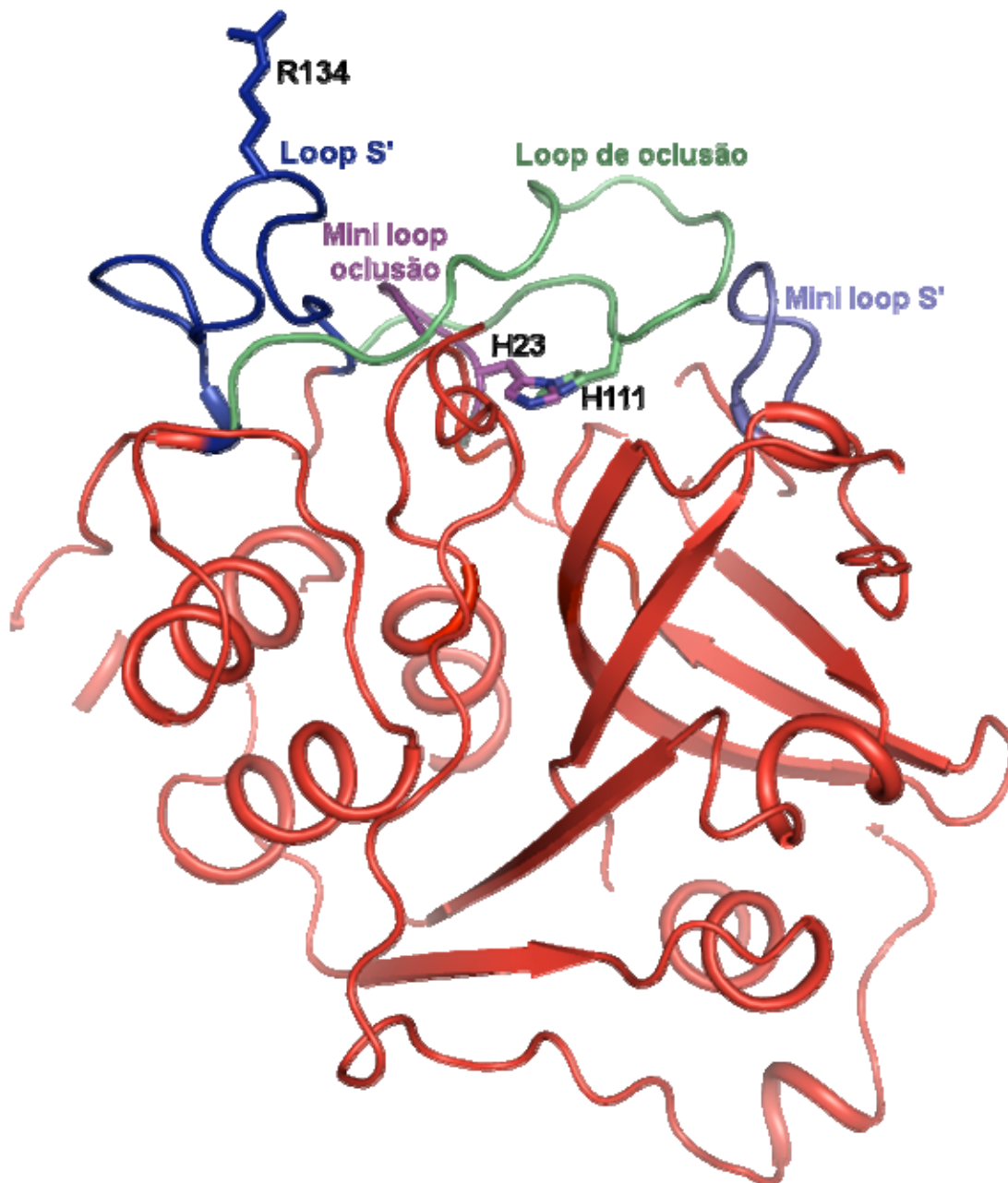


Figura II-4.18: Representação de linhas dos carbonos α da Xylellaína (vermelho) e a sobreposição dos “loops” de oclusão das estruturas cristalográficas da catepsina B (verde) e K (magenta). Em destaque regiões de “loops” pouco conservadas dentro das cisteíno protease. Em azul, “loop” S’ e Mini “loop” S’ presentes na estrutura da Xylellaína; em verde “loop” de oclusão característico da catepsina B; em magenta mini “loop” da catepsina K. Histidinas envolvidas na atividade de exopeptidase das enzimas catepsina B e K são representadas em “sticks”.

Considerações finais.

As cisteíno proteases têm se mostrado como enzimas envolvidas direta ou indiretamente em vários processos de patogenicidade, isso chama à atenção dessas

enzimas como alvos promissores para o planejamento racional de compostos seletivos contra essas enzimas.

A bactéria *X. fastidiosa* é causadora de doenças economicamente importantes, acarretando grandes prejuízos aos agro negócios. Mesmo após o seqüenciamento completo do genoma da bactéria muito pouco ainda se sabe sobre os seus mecanismos moleculares de patogenicidade, além disso, um combate eficiente ao patógeno ainda não é conhecido, sendo importante a realização de estudos que gerem alguma perspectiva de controle desta praga.

Recentemente o grupo do Prof. Dr. Flávio Henrique Silva, identificou e caracterizou uma cisteíno protease (denominada Xylellaína) desse fitopatógeno. As análises iniciais mostram que essa proteína é diferencialmente expressa na cepa patogênica de *X. fastidiosa*, sendo ausente ou pouco expressa na cepa não patogênica. Considerando que esta proteína pode ter participação no desenvolvimento e patogenicidade desta doença, a busca de inibidores seletivos através do planejamento racional de fármacos e/ou uso de moléculas com potencial inibidor pode contribuir para o desenvolvimento de mecanismos de combate a este patógeno.

A proteína Xylellaína é a quarta proteína da bactéria *X. fastidiosa* com estrutura resolvida, e apresenta algumas características estruturais interessantes e pouco usuais dentro da família de cisteíno proteases da papaína.

A estrutura resolvida da Xylellaína mostrou que essa janela aberta de leitura depositada na verdade trata-se de uma forma inativa ou menos ativa, uma “pró-Xylellaína”. Uma região não estruturada de aproximadamente 37 aminoácidos funciona como um inibidor da atividade da enzima ligando-se ao longo de quase metade da proteína, obstruindo o sítio ativo da enzima. A pró-região está posicionada sobre o sítio ativo de uma maneira inversa ao esperado para a posição do substrato, produzindo assim uma orientação não produtiva impedindo a auto catálise da pró-região da enzima. Um intrigante nucleotídeo di fosfatado (UDP) parece direcionar e estabilizar essa pró-região próxima ao sítio de ligação ao substrato. Alguns mutantes pontuais envolvendo resíduos chave dessa interação UDP-Xylellaína foram construídos após a resolução da estrutura e ensaios estão sendo realizados para confirmar importância do mesmo na estrutura da proteína.

Outra característica interessante observada é a presença de dois pequenos loops próximos a porção S' da proteína. Pelo menos um desses loops está em uma posição semelhante ao “loop” de oclusão presente em catepsina B, essa estrutura está diretamente ligada à presença de atividade exopeptidase dessa enzima. No caso da Xylellaina uma pequena atividade exopeptidase foi observada durante os experimentos de atividade enzimática. Novos ensaios enfatizando essa atividade exopeptidase serão realizados para confirmar esses dados preliminares. Um resíduo de histidina é diretamente relacionado com a atividade exopeptidase tanto na estrutura da catepsina B quanto em catepsina X, embora o “loop” S' da Xylellaina não apresente histidinas em sua seqüência de aminoácidos, um resíduo de arginina poderia substituir essa histidina, conferindo a essa enzima uma característica interessante.

Novos experimentos estão em andamento para confirmar algumas das observações apontadas pela estrutura cristalográfica. Esses experimentos juntamente com a estrutura cristalográfica da enzima trarão novas informações sobre o papel biológico dessa enzima e a sua possível relação com a patogenicidade da *X. fastidiosa*.

Referências do capítulo

- Collaborative Computational Project Number 4 (1994). *Acta Crystallogr. D Biol. Crystallogr.*, **50**: 760–763
- Cygler, M e Mort, J. S. (1997), *Biochemie*, **79**: 645-52.
- Cygler, M.; Sivaraman, J.; Grochulski, P.; Coulombe, R.; Storer, A. C. e Mort, J. S. (1996) *Structure*, **4**: 405-416
- Doublié, S. (2003) *Meth. Enzymol.*, **276**: 523-530.
- Drenth, J.; Jansonius, J. N.; Koekoek, R.; Swen, H. M. e Wolthers, B. G. (1968). *Nature*, **218**: 929-932.
- Evans, P.R. (1997). *Joint CCP4 and ESF-EACBM Newsletter*, **33**: 22-24.
- Haas, D .J. e Rossmann, M. G. (1970). *Acta Crystallogr.*, B, **26**: 998 – 1004.

- Horn, M.; Baudys, M.; Voburka, Z.; Klüh, I.; Vondrasek, J. e Mares, M. (2002). *Protein Sci.* **11**: 933-943.
- Illy, C.; Quraishi, O.; Wang, J.; Purisima, E.; Vernet, T. e Mort, J. S., (1997) *J Biol Chem.* **272**:1197-1202.
- Jia, Z.; Hasnain, S.; Hiramata, T.; Lee, X.; Mort, J. S.; To, R. e Huber, C. P.(1995) *J. Biol. Chem.* **270**, 5527–5533.
- Klemencic, I.; Carmona, A. K.; Cezari, M. H.; Juliano, M. A.; Juliano, L.; Guncar, G.; Turk, D.; Krizaj, I.; Turk, V. e Turk, B. (2000). *Eur. J. Biochem.*, **267**: 5404-5412.
- Krissinel, E. e Henrick, K., (2005) In: M.R. Berthold *et.al.* (Eds.): CompLife 2005, LNBI 3695, pp. 163--174. Springer-Verlag Berlin Heidelberg.
- LaLonde, J. M.; Zhao, B.; Janson, C. A.; D'Alessio, K. J.; McQueney, M. S.; Orsini, M. J.; Debouck, C. M. e Smith, W.W. (1999), *Biochemistry*, **38**: 862-869.
- Leslie, A. G. W. (1992). *Jt. CCCP4 ESF-EADBM Newsl. Protein Crystallogr* **26**: 27-33.
- Lo Conte, L.; Chothia, C. e Janin, J. (1999). *J. Mol. Biol.*, **285**, 2177–2198.
- Matthews, B.M. (1968). *J. Mol. Biol.* **33**: 491 – 497.
- Musil, D.; Zucic, D.; Turk, D.; Engh, R. A.; Mayr, I.; Huber, R.; Popovic, T.; Turk, V.; Towatari, T.; Katunuma, N. e Bode, W. (1991). *EMBO J.*, **10**, 2321-2330.
- Nagler, D. K.; Tam, W.; Storer, A. C.; Krupa, J. C.; Mort, J. S.; e Menard, R. (1999). *Biochemistry*, **38**, 4868-4874.
- Nogaroto, V.; Tagliavini, S.; Gianotti, A.; Mikawa, A.; Barros, N. T.; Puzer, L.; Carmona, A. K.; Costa, P.; e Henrique-Silva, F. (2006). *FEMS Microbiol. Lett.*, **261**:187-93.
- Otwinowski, Z. e Minor, W. (1997). *Meth. Enzymol.*, **276**: 307 – 326.
- Pape, T e Schneider, R. (2004). *J. Appl. Crystal.* **37**: 843-844.
- Rzychon, M.; Chmiel, D. e Stec-Niemczyk, J. (2004). *Acta Biochim Pol.*, **51**: 861-873.
- Schneider, T. R. e Sheldrick, G. M. (2002), *Acta Crystallogr. D Biol. Crystallogr.*, **58**: 1772-1779.
- Sivaraman, J.; Lalumiere, M.; Menard, R. e Cygler, M., (1999) *Protein Sci.*, **8**: 283-290.

- Sivaraman, J.; Nagler, D.K.; Zhang, R.; Menard, R. e Cygler, M. (2000). *J.Mol.Biol.*, **295**: 939-951.
- Tao, K.; Stearns, N. A.; Dong, J.; Wu, Q. L. e Sahagian, G. G. (1994). *Arch. Biochem. Biophys.* **311**: 19-27.
- Tao, K.; Stearns, N. A.; Dong, J.; Wu, Q. L. e Sahagian, G. G. (1994), *Arch. Biochem. Biophys.*, **311**:19-27.
- Terwilliger, T. C.(2000). *Acta Crystallogr. D Biol. Crystallogr.*, **56**: 965-972.
- Terwilliger, T. C., Berendzen, J. (1999). *Acta Crystallogr. D Biol. Crystallogr.*, **55**: 849-861.
- Turk, D.; Podobnik, M.; Kuhelj, R.; Dolinar, M. e Turk, V. (1996), *FEBS Lett*, **384**: 211-214.
- Turk, D.; Podobnik, M.; Popovic, T.; Katunuma, N.; Bode, W.; Huber, R. e Turk, V. (1995) *Biochemistry*, **34**: 4791–4797.
- Turk, B. (2006) *Nat. Rev. Drug Discov.*, **5**: 785-799.
- Van Duyne, G.D.; Standaert, R. F.; Karplus, P. A.; Schreiber, S. L. e Clardy, J. (1993). *J. Mol. Biol.*, **229**:105-24.
- Wang, S. X.; Pandey, K. C.; Somoza, J. R.; Sijwali, P. S.; Kortemme, T.; Brinen, L. S.; Fletterick, R. J.; Rosenthal, P. J. e McKerrow, J. H. (2006). *Proc. Natl. Acad. Sci. U. S. A.*, **103**:11503-11508.
- Wiederanders, B. (2003). *Acta Biochim. Pol.*, **50**: 691-713.

Apêndice III

Colaborações realizadas durante o período do doutorado.

Dentro desse apêndice encontra-se de maneira resumida as colaborações em outros trabalhos realizados durante o período do doutorado. Dois trabalhos foram realizados durante esse período. O primeiro foi a determinação da estrutura cristalográfica de uma proteína de ligação a odores “Odorant Binding Protein” de *Aedes aegypti*, e o segundo foi a determinação de parâmetros cinéticos e da estrutura cristalográfica de uma Glioxalase II da bactéria *Salmonella typhimurium*. O primeiro trabalho está em fase de finalização do artigo para submissão a revista “Acta crystallographica F”. O trabalho com a Glioxalase II está com a estrutura tridimensional da proteína depositada no banco de dados PDB (código acesso 2QED) e artigo aceito para publicação na revista “Biochemistry”.

1- Estudos Estruturais de uma Proteína de ligação a moléculas odoríficas (Odorant Binding Protein – OBP) de *Aedes aegypti*.

A transmissão do vírus da dengue e também da febre amarela ao hospedeiro vertebrado dentro do ciclo de vida da doença é realizado por mosquitos da espécie *Aedes aegypti*. Os principais países endêmicos são localizados na África tropical e na América do Sul.

Feromônios e moléculas odoríferas são importantes sinais químicos para a vida dos insetos. Isso pode ser comprovado pela existência de um sistema extremamente especializado de recepção, transporte e degradação, que auxiliam os insetos a encontrar alimento, localização do parceiro sexual, sítios de ovoposição entre outros. Alguns grupos de proteínas em especial como: os receptores de odores (ORs); as proteínas de ligação a odores (OBPs); e as enzimas de degradação de odores (ODEs) desempenham um importante papel na recepção, transporte, localização e degradação dessas moléculas sinalizadoras essenciais para a vida dos insetos.

Dados experimentais sugerem que essas proteínas são capazes de ligação seletiva a compostos fisiologicamente importantes, permitindo o transporte de substâncias hidrofóbicas através do meio aquoso presente na linfa dos órgãos olfativos, bem como a sua liberação próxima aos receptores odoríferos (ORs) ligados às membranas celulares. Dados bioquímicos e estruturais mostram que um mecanismo pH dependente é responsável pela liberação/captação dessas moléculas odoríferas.

O conhecimento da estrutura tridimensional dessa enzima pode conduzir a abordagens para a utilização dessa proteína como um alvo no planejamento de novas entidades químicas com potencial de prevenção ou combate a esses insetos.

A resolução da estrutura da proteína de ligação a odores de *A. aegypti*, AegOBP (*A. aegypti* Odorant Binding Protein) faz parte de uma colaboração de nosso laboratório com o Prof. Dr. Walter Leal, Department of Entomology University of California-Davis.

Os trabalhos envolveram a cristalização da proteína através da metodologia de difusão de vapor; coleta e processamento de dados de difração de raios-X no

Brookhaven National Laboratory -Upton, NY, resolução e refinamento da estrutura até a resolução de 1,95Å. Os parâmetros e as estatísticas de coleta e processamento de dados são resumidos na tabela AIII-1.1.

Tabela AIII-1.1: Estatísticas de coleta dados, processamento e refinamento.

Conjunto de dados		X12C - BNL
Grupo espacial		P2 ₁
Cela unitária	<i>a, b, e c</i> (Å)	34.37, 47.93, 69.12
	β (°)	96.58
	Comprimento de onda (Å)	0.9
	Resolução (Å)	50.00 (1.95)
	R_{merge} (%)	9.4 (38.6)
	$I/\sigma(I)$	17.8 (4.5)
	Completeza (%)	100 (100)
	Redundância	7.2 (7.5)
	Nº.de reflexões observadas	126447
	Nº. de reflexões únicas	17789
Refinamento		
	Número de átomos	2337
	$R_{\text{work}}/R_{\text{free}}$ (%)	16/22
r.m.s.d.	Ângulos de ligações (°)	1.633
	Comprimento de ligações (Å)	0.018
Ramachandran plot (%/No.)	Mais favoráveis	91.8
	Região adicional	8.2

Números entre parênteses representam a camada de resolução mais externa (1.87 - 1.95Å)

Nesse trabalho a estrutura da AegOBP1 foi determinada até a resolução de 1.95 Å pela metodologia de substituição molecular. A estrutura compartilha enovelamento semelhante às OBP de inseto, composta por seis α -hélices conectadas por três ligações de dissulfeto (figura AIII-1.1). Uma longa molécula, compatível com polietilenoglicol (PEG) foi encontrada ligada ao bolso de ligação do substrato de duas moléculas presentes na unidade assimétrica (figuras AIII-1.2 A-C). Ensaios de espectrometria de massa com cristais da proteína dissolvida não apontaram qualquer outra substância presente em quantidades aceitáveis.

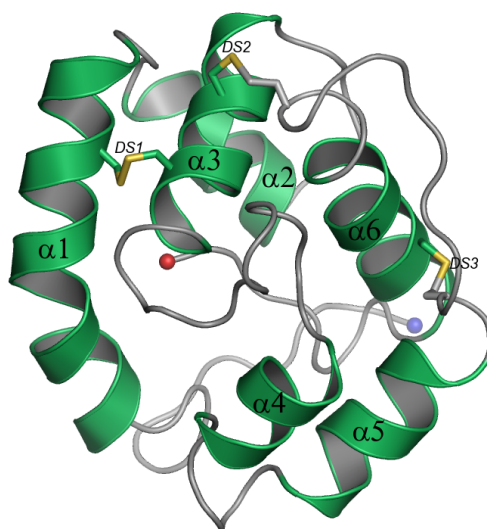


Figura AIII-1.1: Representação em “cartoons” do monômero AegOBP1. Estrutura comum as OBP’s pode ser observada pela presença de 6 hélices- α (verde) conectadas por 3 ligações de dissulfeto. O interior das hélices é composto por resíduos hidrofóbicos, formando um bolsão hidrofóbico. A porção C-terminal (esfera vermelha) está inserida dentro do bolsão. N-terminal é assinalado com uma esfera azul.

Uma característica observada para essa proteína, bem como para a homóloga em *Anopheles gambiae*, é que a porção C-terminal da proteína está localizado dentro do bolso de ligação ao substrato.

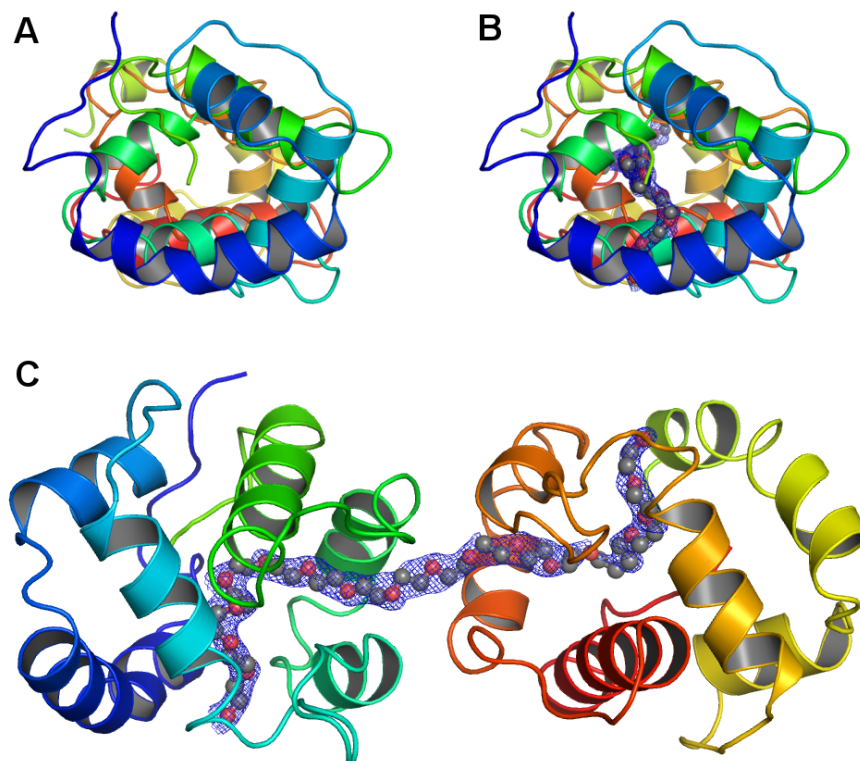


Figura AIII-1.2: Representação em “cartoons” do dímero da AegOBP1. A) visão superior mostrando o canal formado pelas cadeias A e B. B) Mesma orientação da figura A com a longa molécula de PEG atravessando completamente os dímeros. C) Visão lateral do dímero formado pelas cadeias A e B da AegOBP1 com PEG atravessando completamente os dímeros. Mapas de densidade eletrônica (2Fo-Fc) contornado a 1 sigma ao longo da molécula de PEG

Uma intensa rede de interações de hidrogênio principalmente entre oxigênios de resíduos de ácido glutâmico com nitrogênios de argininas ou lisinas mantém o C-terminal preso à estrutura. A diminuição do pH, semelhante a que ocorre próximo à membrana de ligação aos receptores odoríferos, poderia causar a protonação dos resíduos dessa região, liberando o C-terminal e o conteúdo presente dentro do bolsão de ligação do substrato.

2- Estudos Estruturais de uma Glyoxalase II de *Salmonella typhimurium*.

Metilglioxal (MG) é um composto cito tóxico e mutagênico formado primariamente como produto do metabolismo de carboidratos e lipídios. Em *E. coli*, a síntese endógena de MG ocorre quando existe um desequilíbrio durante a gliocólise para a formação de piruvato a partir de gliceraldeído 3-fosfato (PGAL). MG é capaz de modificar bases do DNA e resíduos de proteínas conduzindo a degradação dessas moléculas seguida de apoptose celular. O maior meio de proteção contra o MG é a sua reação espontânea com glutathione, seguido da detoxificação via sistema glioxalase. Esse sistema está presente em diferentes organismos, e consiste de duas enzimas: glioxalase 1 (lactoil glutathione metilglioxalase liase – GLX1) e glioxalase 2 (hidroxi acil glutathione hidrolase – GLX2). Dentro do metabolismo da glioxalase os α -ceto aldeídos aromáticos e alifáticos como o metil glioxal reagem espontaneamente com a glutathione para formar tio-hemioacetal. A GLX1 catalisa a isomerização desse tio-hemioacetal para produzir derivados S-(2-hidroxiacil)-glutathione. A GLX2 hidrolisa esses derivados S-(2-hidroxiacil)-glutathione para regenerar a glutathione e liberar hidroxi-ácidos que não apresentam efeito deletério no organismo.

GLX1 é uma metalo-enzima bem descrita, que requer um íon metálico divalente para sua atividade. Considerando-se a sua similaridade seqüencial e estrutural, a família GLX1 pode ser dividida em enzimas Zn(II) e Ni(II) dependentes. Estudos estruturais mostram que a seletividade dos metais é finamente regulada por elementos presentes na estrutura da proteína. GLX2 também é uma metalo-enzima, porém com algumas diferenças não completamente elucidadas. Estudos bioquímicos e estruturais mostram que dois íons metálicos divalentes são necessários para a completa atividade da enzima. Nos estudos descritos até o presente momento formas GLX2 com presença de Ferro, Zinco e Manganês, em seus sítios metálicos têm sido reportadas.

GLX2 é um dos muitos membros da família das metalo-beta-lactamases, apresentando o enovelamento característico $\alpha\beta/\beta\alpha$, capaz de ligar até dois íons metálicos em seu sítio ativos. Metalo-beta-lactamases da mesma família da GLX2

são caracterizadas como enzimas Zn dependentes, com alguns exemplos de enzimas Fe dependentes. Esse mecanismo de seletividade não é completamente compreendido até o presente momento, porém uma rede formada por resíduos de aminoácidos ao redor do sítio ativo parece estar diretamente ligada a essa seletividade, formando uma estrutura denominada segunda esfera de coordenação.

GLX2 existe como várias iso-enzimas nos eucariotos, e sua atividade tem sido demonstrada tanto no citosol como no interior das mitocôndrias, sugerindo um papel fisiológico mais importante do que a GLX1.

Diferentes patógenos como *Salmonella* e *Brucella*, quando são fagocitados por macrófagos, utilizam-se do sistema glioxalase como defesa ao mecanismo de destruição celular gerado pelo fagolisossomo.

Após o sequenciamento do genoma da *Salmonella enterica* sorotipo *typhimurium* foi observado que essa bactéria possui em seu genoma somente uma janela aberta de leitura para a enzima GLX2, enquanto a GLX1 apresenta três genes codificados nos genoma. Essa característica não é observada na maioria das bactérias aonde um único gene para cada enzima é comumente encontrado. A fim de elucidar o mecanismo de seletividade da enzima para os diferentes metais, bem como para aumentar o conhecimento dessa classe de enzimas, foi então realizado um estudos bioquímico e estrutural da GLX2 de *S. typhimurium*.

A determinação da estrutura cristalográfica da enzima Glioxalase II (GloB) foi realizada através de colaboração de nosso laboratório com a doutoranda Valéria Campos Bermudez e seu orientador o prof. Dr. Alejandro Vila, do Instituto de Biología Molecular e Celular de Rosario, da Universidade Nacional de Rosário, Argentina.

Resumidamente o trabalho realizado em nosso laboratório envolveu a cristalização da proteína através da metodologia de difusão de vapor; coleta e processamento de dados de difração de raios-X, resolução e refinamento da estrutura até a resolução de 1,45 Å. Os parâmetros e as estatísticas de coleta e processamento de dados são resumidos na tabela AIII-2.1.

Tabela AIII-2.1: Estatísticas de coleta dados, processamento e refinamento.

Conjunto de dados		Anodo rotatório	LNLS
Grupo espacial		P2 ₁	P2 ₁
Cela unitária	<i>a, b, e c</i> (Å)	44.13, 58.37, 54.76	43.37, 57.17, 53.9
	β (°)	112.08	111.90
Comprimento de onda (Å)		1.541	1.427
Resolução (Å)		50.76 (1.67)	14.4 (1.45)
R_{merge} (%) [#]		5,7 (32)	7,3 (31,6)
I/σ (I)		9.2 (2.3)	6.7 (1.9)
Completeza (%)		98.3 (88.9)	99.9 (99.9)
Redundância		4.8 (4.5)	5.4 (5.2)
Nº. de reflexões observadas		142464	235043
Nº. de reflexões únicas		29723	43405
Refinamento			
Número de átomos			2670
$R_{\text{work}}/R_{\text{free}}$ (%)			15.7/17.8
Ângulos de ligações (°)			0.930
r.m.s.d.	Comprimento de ligações (Å)		0.005
	Mais favoráveis		91.4/202
Ramachandran plot (%/No.)	Região adicional		8.1/18
	Região generosamente permitida		0.5/1

Números entre parênteses representam à camada de resolução mais externa (1.53 - 1.45 Å)

Glyoxalase II (GLX2) é uma enzima hidrolítica que faz parte do sistema glioxalase, que é responsável pela detoxificação de vários compostos citotóxicos utilizando glutatona como um dos substratos. A estrutura determinada apresenta enovelamento característico da família das metallo-beta-lactamases (figura AIII-2.1), com um motivo conservado capaz de ligar até dois metais (figuras AIII-2.2 A e B). Maior parte dos membros dessa família é constituída como hidrolases zinco dependentes. Todavia, várias GLX2 de eucariotos têm sido caracterizadas com diferentes taxas de ferro, zinco e manganês em seus centros metálicos. A GLX2 de *S. tiphymurium* recombinante apresenta atividades similares, quando expressas em meio mínimo suplementado com os metais específicos. Todavia os experimentos de NMR e EPR mostram que a GLX2 de *S. tiphymurium* expressa em meio de cultivo LB apresenta predominantemente ferro em seu centro metálico, com proporções menores de manganês e zinco. Essa característica indica certa flexibilidade do sítio de acomodar os metais acima citados sem significativas quedas na eficiência da enzima.

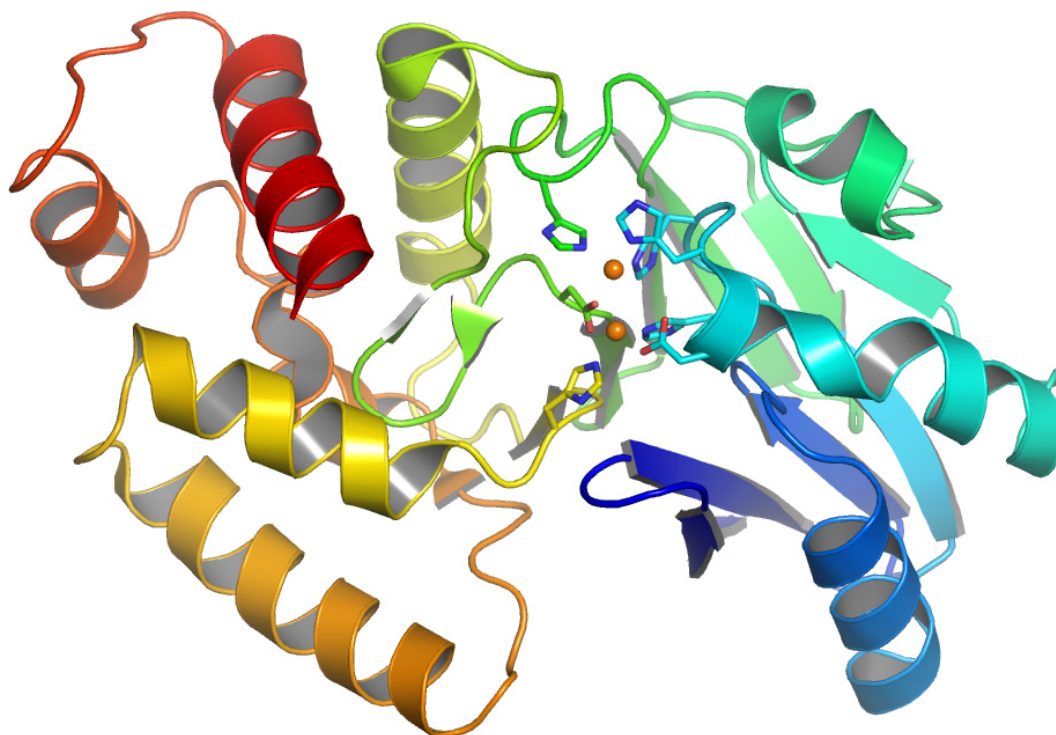


Figura AIII-2.1: Representação em “cartoons” da estrutura cristalográfica da glicoxalase II de *S. tihymurium*. Esferas representam o ferro, ligado ao sítio de ligação aos metais (“sticks”).

Uma inspeção detalhada na segunda esfera de ligação ao redor do sítio metálico mostra novas características desse sítio de ligação que pode direcionar a seleção metal ligado ao sítio. Um resíduo D11, na segunda esfera de ligação, fortemente coordenado à H165 aumenta a rigidez desse sítio, favorecendo assim a entrada o íon ferro nessa posição do sítio metálico.

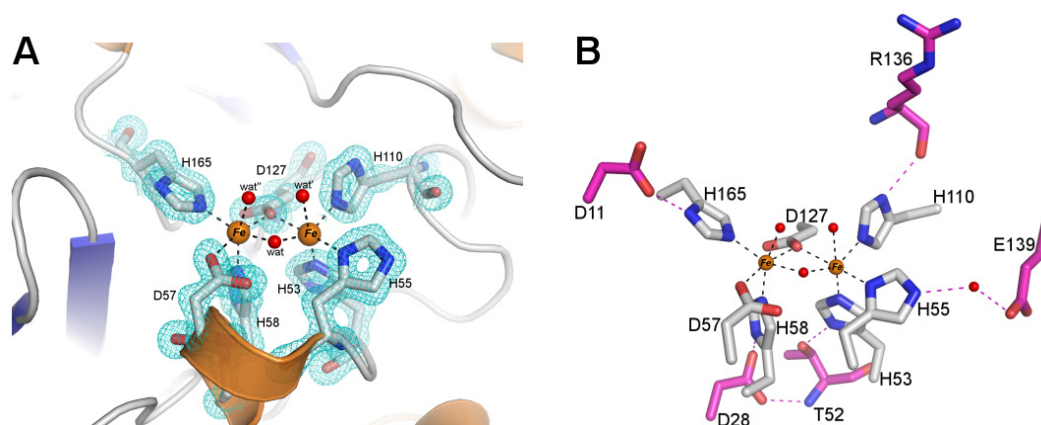


Figura AIII-2.2: Detalhe do sítio de ligação aos metais e da segunda esfera de coordenação. **A)** Mapas de densidade eletrônica (2fo-fc) contornados a 1σ ao redor dos resíduos que formam o sítio de ligação ao substrato. **B)** Detalhe da segunda esfera de coordenação (“sticks” rosa) coordenando o sítio de ligação ao metal (cinza). Duas águas completam a coordenação dos metais.

Apêndice IV

Artigo publicado na revista “Biochemistry”.

Dentro desse apêndice encontra-se o artigo aceito para publicação na revista “*Biochemistry*”, intitulado “***Biochemical and Structural Characterization of Salmonella typhimurium glyoxalase II: New insights in metal ion select***”. Esse trabalho foi realizado em colaboração com a doutoranda Valéria Campos Bermudez e o prof. Dr. Alejandro Vila, Instituto de Biología Molecular y Celular de Rosario – Argentina.

Biochemical and Structural Characterization of *Salmonella typhimurium* Glyoxalase II: New Insights into Metal Ion Selectivity^{†,‡}

Valeria A. Campos-Bermudez,[§] Ney Ribeiro Leite,^{||} Renata Krog,^{||} Antonio J. Costa-Filho,[⊥] Fernando C. Soncini,[§] Glaucius Oliva,^{||} and Alejandro J. Vila^{*,§}

Instituto de Biologia Molecular y Celular de Rosario, IBR-CONICET and Area Biofisica, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Rosario S2002LRK, Argentina, and Centro de Biotecnología Molecular e Estructural and Grupo de Biofísica Molecular Sérgio Mascarenhas, Instituto de Física de São Carlos, Universidade de São Paulo, CP 369, 13560-970 São Carlos, SP, Brazil

Received April 17, 2007; Revised Manuscript Received July 30, 2007

ABSTRACT: Glyoxalase II is a hydrolytic enzyme part of the glyoxalase system, responsible for detoxifying several cytotoxic compounds employing glutathione. Glyoxalase II belongs to the superfamily of metallo- β -lactamases, with a conserved motif able to bind up to two metal ions in their active sites, generally zinc. Instead, several eukaryotic glyoxalases II have been characterized with different ratios of iron, zinc, and manganese ions. We have expressed a gene coding for a putative member of this enzyme superfamily from *Salmonella typhimurium* that we demonstrate, on the basis of its activity, to be a glyoxalase II, named GloB. Recombinant GloB expressed in *Escherichia coli* was purified with variable amounts of iron, zinc, and manganese. All forms display similar activities, as can be shown from protein expression in minimal medium supplemented with specific metal ions. The crystal structure of GloB solved at 1.4 Å shows a protein fold and active site similar to those of its eukaryotic homologues. NMR and EPR experiments also reveal a conserved electronic structure at the metal site. GloB is therefore able to accommodate these different metal ions and to carry out the hydrolytic reaction with similar efficiencies in all cases. The metal promiscuity of this enzyme (in contrast to other members of the same superfamily) can be accounted for by the presence of a conserved Asp residue acting as a second-shell ligand that is expected to increase the hardness of the metal binding site, therefore favoring iron uptake in glyoxalases II.

Methylglyoxal is a cytotoxic and mutagenic compound that is formed primarily as a byproduct of carbohydrate and lipid metabolism (1, 2). In *Escherichia coli*, endogenous methylglyoxal synthesis occurs when there is an imbalance

between the rates of the lower segment of glycolysis, from glyceraldehyde 3-phosphate (PGAL)¹ to pyruvate, and the pathways generating dihydroxyacetone phosphate and PGAL (3, 4). Methylglyoxal is able to modify DNA and protein residues, leading to the degradation of these essential components and subsequent apoptosis of the cells (5). The major means of protection against this harmful compound relies on its spontaneous reaction with glutathione, followed by detoxification through the glyoxalase system (Figure 1). The glyoxalase system consists of two enzymes, glyoxalase I (GLX1; EC 4.4.1.5, lactoylglutathione methylglyoxal lyase) and glyoxalase II (GLX2; EC 3.1.2.6, hydroxyacylglutathione hydrolase) (1, 3). Aromatic and aliphatic α -keto aldehydes such as methylglyoxal react spontaneously with glutathione to form thiohemiacetals, and GLX1 catalyzes the isomerization of this thiohemiacetal to produce *S*-(2-hydroxyacyl)-glutathione derivatives (6–8). GLX2 hydrolyzes *S*-(2-hydroxyacyl)glutathione derivatives to regenerate glutathione, thus releasing harmless hydroxyacids (Figure 1) (3).

The glyoxalase system is present in different organisms, and so far, a series of studies on their constituent enzymes have revealed conserved structural features. GLX1 is a well-described metalloenzyme that requires one divalent metal ion in its active site essential for its activity (6–8). Despite a large sequence and structural similarity among the GLX1 family from different sources, they can be divided into

[†] This work was supported by grants from ANPCyT and HHMI to A.J.V. V.A.C.-B. is the recipient of a doctoral fellowship from CONICET. The Bruker Avance II 600 MHz NMR spectrometer was purchased with funds from ANPCyT (Grant No. PME2003-0026) and CONICET. A.J.V. and F.C.S. are staff members from CONICET. A.J.V. is an International Research Scholar of the Howard Hughes Medical Institute. The Centro de Biotecnología Molecular Estructural and NRL are supported by FAPESP. A.J.C.-F. thanks PRONEX/FAPESP/CNPq (Grant No. 03/09859-2) and CNPq (Grant No. 307102/2006-8) for financial support.

[‡] Coordinates for the structure discussed in this manuscript have been deposited in the Protein Data Bank under PDB code 2QED.

* To whom correspondence should be addressed. Phone: +54-341-4351235 ext 108. Fax: +54-341-4390465. E-mail: vila@ibr.gov.ar.

[§] Universidad Nacional de Rosario.

^{||} Centro de Biotecnología Molecular e Estructural, Universidade de São Paulo.

[⊥] Grupo de Biofísica Molecular Sérgio Mascarenhas, Universidade de São Paulo.

¹ Abbreviations: GLX1, glyoxalase I; GLX2, glyoxalase II; PGAL, glyceraldehyde 3-phosphate; IPTG, isopropyl β -D-thiogalactoside; SLG, *S*-D-lactoylglutathione; DTNB, 5,5'-dithiobis(2-nitrobenzoate); MOPS, 3-(*N*-morpholino)propanesulfonic acid; GST, glutathione-*S*-transferase; PEG, polyethylene glycol; MES, 2-(*N*-morpholino)ethanesulfonic acid; EDTA, ethylenediaminetetraacetic acid; EPR, electronic paramagnetic resonance; NMR, nuclear magnetic resonance; rmsd, root mean square deviation.

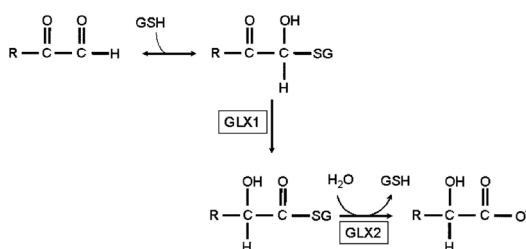


FIGURE 1: Mechanism of the glyoxalase system. The major means of protection against methylglyoxal is the spontaneous reaction of this compound with glutathione to form the hemithiolacetal, followed by detoxification by the glyoxalase system, leading to the production of D-lactate.

Zn(II)-dependent and Ni(II)-dependent enzymes. It has been shown that the metal ion selectivity within different GLX1 family members is finely controlled by the protein structure (6–9). GLX2 is also a metalloenzyme, but with quite different (and less explored) features in this respect. The first crystal structure available for a GLX2 was that of the human variant, reported as a dizinc enzyme, despite the observation of a nonnegligible iron content in the same sample (10). Later, other eukaryotic GLX2 enzymes were isolated and characterized in their zinc, iron, and manganese forms (11–14), since they are usually expressed in *E. coli* as mixed metal forms. While all these metal derivatives seem to be equally active, the *in vivo* metal requirements of these enzymes are still obscure.

GLX2 is one of many members characterized to date of the metallo- β -lactamase superfamily (15, 16). Proteins belonging to this superfamily are characterized by the presence of an $\alpha\beta/\beta\alpha$ fold and a conserved motif able to bind up to two metal ions in their active sites (17, 18). Metallo- β -lactamases and other hydrolases from the same superfamily (such as phosphodiesterase, acyl-homoserine-lactone lactonase, methylparathion hydrolase, and alkylsulfatase) (19–24) are zinc-dependent enzymes, while the terminal oxidase rubredoxin:oxygen oxidoreductase from *Desulfovibrio gigas* (ROO) is a diiron enzyme (25). In this respect, GLX2, despite exhibiting the same coordination sphere displayed by other hydrolases of the superfamily, has been an outstanding exception regarding the metal ion dependence.

GLX2 exists as multiple isoenzymes in eukaryotes. GLX2 activity has been found both in the cytosol and in the mitochondria (26–28), suggesting a more complex physiological role than that of GLX1. In contrast, most bacteria sequenced up to date harbor a single gene copy of each of these enzymes (<http://biocyc.org>). *Salmonella enterica* serovar Typhimurium constitutes an exception to this common rule (29). While it harbors only one open reading frame coding for a GLX2 enzyme (STM0261 or *gloB*), three putative GLX1 enzymes are encoded in this pathogen's genome, including STM0235 (*yaeR*) and STM1435 (*gloA*), both present also in other *S. enterica* serovars. It also includes STM3117, a gene located in a putative pathogenicity island present only in some *Salmonella* serovars, such as Choleraesuis and Enteritidis, that encodes a virulence factor (30). Different intracellular pathogens, such as *Salmonella* and *Brucella*, when internalized by macrophages, employ this detoxifying system as a defense mechanism to the accumula-

tion of methylglyoxal generated in the phagolysosome (31, 32).

We report here the biochemical, structural, and biophysical characterization of GloB, the GLX2 from the pathogen *Salmonella typhimurium*. We have found that this bacterial enzyme can also employ different metal cofactors such as Fe(II), Fe(III), Zn(II), and Mn(II) to give rise to active forms, though with a clear predominance favoring iron ions. The iron form shares some spectroscopic features reported for eukaryotic homologues. On the basis of the high-resolution crystal structure of this enzyme, we also propose a rationale for the selectivity for iron of GLX2, based on the presence of conserved amino acid residues at the second shell around the metal binding sites. This preference for the iron form of the *Salmonella* enzyme, when compared to the recently characterized GLX2 from *E. coli* enzyme (33) in its di-Zn(II) form, could be related to its role in pathogenicity.

EXPERIMENTAL PROCEDURES

General Information. S-D-Lactoylglutathione (SLG) was purchased from Sigma-Aldrich. All chromatographic steps were performed on an Amersham Biosciences liquid chromatography system operating at 4 °C. Metal standards were purchased from Fisher Scientific and were diluted with distilled water. All other chemicals used in this study were purchased commercially and were of the highest quality available.

Gene Amplification and Cloning. STM0261 (*gloB*) was PCR amplified from *S. enterica* serovar Typhimurium 14028s chromosome as previously described (34), using the primers 0261NTF (5'-ATCGGATCCATGAATCTTAACAG-TATTCCCGCG-3') and 0261CTR (5'-ATGCCTCGAGAA-GCTTTCAGAACGTGTCTTTCTTTGAC-3'), which include a *Bam*HI and an *Xho*I restriction site at the 5' and 3' ends of the fragment. The amplified 700 bp *gloB* PCR fragment was purified from a 2% agarose gel using a Qiaex (Qiagen) band purification kit and cloned between the *Bam*HI and *Xho*I restriction sites in the pKP cloning vector (a derivative of the pBluescript II vector, Stratagene). The *gloB* sequence was determined from the resulting plasmid pKP-*gloB*. *gloB* was then cloned into a pET32a(+) expression vector (Novagen), generating plasmid pET32-*gloB*. The plasmid was finally introduced into XL1-blue *E. coli* cells by electroporation using a Bio-Rad apparatus following the manufacturer's recommendations. This cloning strategy resulted in the addition of six extra residues to the N-terminus of GloB.

Protein Overexpression and Purification. The GLX2 (GloB) from *S. typhimurium* recombinantly produced in *E. coli* was used for this study. Overexpression and purification were performed using the pET32-*gloB* vector. This system yields high-level expression of the recombinant GloB protein sequence fused by its N-terminal end to the cleavable Trx-Tag thioredoxin protein and the His-Tag sequence.

Optimal induction conditions for the expression of GloB protein were achieved using isopropyl β -D-thiogalactoside (IPTG) or lactose as the inductor agent, and different temperatures of induction were attempted. Optimal overexpression of the fusion protein was achieved by inducing a BL21-(DE3)pLysS' *E. coli* culture at OD₆₀₀ 0.4–0.6 with 0.5 mM IPTG and growing it for 4 h at 30 °C.

In a typical protein preparation, a 2 L BL21-(DE3)pLysS⁺ *E. coli* culture was grown in LB medium (Sigma) and induced as described above. The bacteria were harvested by centrifugation and resuspended in 50 mM Tris-HCl, pH 8.0, 1 mM phenylmethanesulfonyl fluoride, 0.01 mg/mL DNase, and 5 mM MgCl₂. Sonication was performed four times for 30 s followed by ultracentrifugation at 100000g. The bacterial lysate was applied to a column of Ni-agarose (Qiagen). After washing with 50 mM Tris-HCl, pH 8.0, 300 mM NaCl, and 20 mM imidazole, the fusion protein was eluted with 50 mM Tris-HCl, pH 8.0, 300 mM NaCl, and 250 mM imidazole. The fusion protein was diafiltered using an Amicon concentrator to exchange to the buffer needed for the enterokinase digestion (50 mM Tris-HCl, pH 8.0, 10 mM CaCl₂). The sample was then incubated with 50 units of enterokinase protease (EKmax, Invitrogen) for 2 h at 20 °C, according to the manufacturer's instructions. After the digestion the sample was diluted 1:3 in 50 mM Tris-HCl, pH 8.0, and loaded onto the Ni-agarose column. Trx-Tag thioredoxin protein plus the His-Tag was bound to the resin, whereas the pure GloB passed through it. Fractions with GLX2 activity were pooled and dialyzed against 10 mM 4-morpholinepropanesulfonic acid (MOPS) and 0.2 M NaCl at pH 7.2.

Minimal medium M9 was employed to produce selectively metal-enriched enzyme species. The minimal medium contained 4 g/L D-(+)-glucose (Sigma), 12.8 g/L Na₂HPO₄·7H₂O, 3 g/L KH₂PO₄, 0.5 g/L NaCl, 1.0 g/L ammonium sulfate, 10 μM CaCl₂, and 1 mM MgSO₄. Supplemental metal ions were added from stock solutions of Fe(NH₄)SO₄, MnCl₂, or ZnCl₂ to reach a final concentration in the culture of 100 μM. The metal ions were added in the moment of the induction, when the growth cultures reached an OD₆₀₀ of 0.6–0.8.

Protein Concentration. Enzyme concentrations were determined by measuring the $A_{280\text{ nm}}$ and using the extinction coefficient $\epsilon_{280\text{ nm}} = 28\,030\text{ M}^{-1}\text{ cm}^{-1}$. This parameter was calculated from amino acid composition information by applying a modified Edelhoch method (35).

Steady-State Kinetics. Steady-state kinetic parameters were determined by measuring the initial rates of formation of glutathione due to the hydrolysis of SLG in the presence of 200 μM 5,5'-dithiobis(2-nitrobenzoate) (DTNB) at 412 nm ($\epsilon_{412\text{ nm}} = 13\,600\text{ M}^{-1}\text{ cm}^{-1}$) as described in ref 36. The concentration of SLG was varied from 18 to 1000 μM, and the amount of enzyme was 100–200 nM. The measurements were performed in a reaction volume of 300 μL of 10 mM MOPS, 0.2 M NaCl, pH 7.2, buffer, at 30 °C in a Jasco 550 UV-vis spectrophotometer. In the case of the enzyme obtained from minimal medium the measurements were performed with the addition of a 20 μM concentration of the corresponding metal ion (iron, zinc, or manganese) according to the metal form of the enzyme that was being measured. The measurements were performed at least in triplicate. Kinetic constants were determined by fitting the data of initial rates to the Michaelis-Menten equation.

Metal Analysis. The metal content of the GloB samples was measured using atomic absorption spectroscopy in a Metrolab 250 AA instrument. Purified enzymes were diluted with 10 mM MOPS, pH 7.2, to a concentration of 30 μM and analyzed for zinc, manganese, and iron. The metal

content data presented in this paper represent an average from at least three preparations of each growth condition.

Protein Crystallization and X-ray Diffraction Data Collection. GloB purified from expression in rich growth medium was concentrated to 8.0 mg/mL and screened against Crystal Screen I and II kits (Hampton Research) with the hanging-drop vapor diffusion method at 4 and 18 °C employing equal volumes of the reservoir solution and protein. Small crystals were observed after two weeks at 4 °C in 0.1 M MES buffer, pH 6.5, 12% PEG 20 000. The crystals were cryoprotected by addition of 20% ethylene glycol to the reservoir solution.

An initial diffraction data set was collected to 1.7 Å resolution on a MAR345dtb image-plate detector using Cu K α radiation generated by a Rigaku Ultra-X 18 rotating anode operating at 90 mA and 50 kV and focused using Osmic mirrors. Other crystals from the same crystallization batch were used to collect a 1.45 Å resolution diffraction data set at beamline MX1 in the Brazilian National Light Source (LNLS) (37, 38). Diffraction data were integrated and reduced with MOSFLM (39), and intensities were scaled with SCALA from the CCP4 suite (40). Data collection statistics are summarized in Table 3.

Structure Solution, Refinement, and Validation. The molecular replacement was performed with PHASER (41) using the 1.7 Å data set and a search model prepared with Chainsaw (40) based on the crystal structure of *Arabidopsis thaliana* GLX2 (Protein Data Bank code 1XM8) (13). The *S. typhimurium* structure was modeled and rebuilt using COOT (42), and refinement was carried out against the 1.45 Å resolution data set using REFMAC5 (43) and PHENIX-REFINE (44). The structure was validated using PROCHECK (45), and the final data refinement statistics are shown in Table 3. All pictures were created in PyMOL (46). The coordinates and structure factors have been deposited in the Protein Data Bank (accession number 2QED).

¹H NMR Spectroscopy. NMR spectra were recorded on a Bruker Avance II 600 spectrometer operating at 600.13 MHz at different temperatures, as indicated. ¹H NMR spectra were recorded under conditions to optimize detection of the fast relaxing paramagnetic resonances, using either the super-WEFT pulse sequence (47, 48) or water presaturation. Spectra were acquired over large spectral widths with acquisition times ranging from 16 to 80 ms and intermediate delays from 2 to 35 ms. 1D experiments with solvent presaturation were used to record isotropically shifted signals closer to the diamagnetic envelope. The samples in D₂O were made by performing three or more dilution/concentration cycles in an Amicon centrifugal filter device to a final concentration of 1.3 mM. The samples were then loaded into Wilmad 5 mm tubes for NMR.

Electron Paramagnetic Resonance Spectroscopy. X-band (9.5 GHz) EPR spectra were measured on a Bruker ELEXSYS E580 system (Bruker BioSpin, Germany) at 4.7 and 31 K. The temperature was controlled with an Oxford ITC503 cryogenic system. EPR samples containing a convenient amount of protein were frozen by immersion in liquid nitrogen and then placed in the spectrometer rectangular cavity. All EPR data were corrected by subtracting a baseline corresponding to the EPR signal of the buffer. Other acquisition conditions: modulation amplitude, 1 mT; modu-

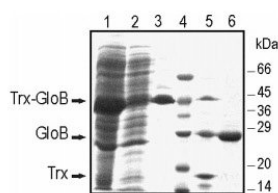


FIGURE 2: SDS-PAGE of the purification steps of GloB. The different lanes correspond to (1) the soluble fraction of the induced bacterial extract, (2) the flow-through after the wash step in a Ni-agarose affinity column, (3) elution of the 45 kDa fusion protein from this resin, (4) molecular mass standards (Sigma), (5) digestion of the fusion protein by enterokinase, and (6) purified GloB.

lation frequency, 100 kHz; microwave power, 4 mW. A typical EPR sample was 1 mM in 10 mM MOPS, pH 7.2, buffer.

RESULTS

Cloning of the *gloB* Gene and Protein Expression. The STM0261 gene (*gloB*) encodes a 28 kDa protein with high homology with GLX2. This gene was specifically amplified from the *S. enterica* serovar Typhimurium 14028s genome and finally cloned in an expression vector. Different expression and purification strategies were attempted. Overexpression of GloB in a pET22b plasmid gave reasonable protein yields. All attempts to purify GloB by using a glutathione-agarose resin, as reported for the yeast enzyme (49), were unsuccessful. Interestingly, this approach could not be applied either to the *E. coli* protein (33). Then we resorted to other strategies in which GloB was expressed as a fusion protein. When GloB was produced as an N-terminal or C-terminal fusion protein to GST, good levels of protein expression were obtained, but cleavage of the fusion protein with a protease gave low yields. Finally, GloB was cloned in a pET32a(+) plasmid, and the protein was overexpressed in a BL21-(DE3)pLysS' *E. coli* culture as a fusion protein to Trx-Tag plus His-Tag (Figure 2). After sonication and ultracentrifugation, the soluble fraction of the culture was loaded onto a Ni-agarose column. Followed by the wash step, the protein was eluted with a buffer containing 250 mM imidazole. The fusion protein was then cleaved with enterokinase (EKmax). To separate GloB from the Trx-Tag plus His-Tag, the digestion sample was loaded again onto the Ni-agarose column. In this case, Trx-Tag plus His-Tag and some of the eventually nondigested fusion were retained on the resin, whereas GloB passed directly through the column. Once obtained, the pure protein was dialyzed against MOPS buffer and quantified using the molar extinction coefficient at 280 nm of $28\,030\text{ M}^{-1}\text{ cm}^{-1}$. This successful purification procedure resulted in an average yield of 70 mg of GloB/2 L of culture.

Due to the high sequence homology between GloB and other GLX2 enzymes from different organisms and the presence of the typical metal binding site of the metallo- β -lactamase superfamily members, we assayed the hydrolytic activity of GloB against SLG. GloB was able to hydrolyze SLG, confirming its predicted identity as GLX2. This activity was impaired with the addition of EDTA to the reaction medium (data not shown).

Metal Content and Catalytic Activity of GloB. Previous studies on GLX2 have reported different metal contents for

Table 1: Metal Content of GLX-2 Samples: Comparison of GLX2 Metal Contents Reported in Previous Studies with That Obtained in This Job in Similar Work Conditions

isoform	Zn	Fe	Mn	total metal content
human GLX2-2 ^a	1.5	0.7	not tested	2.2
<i>Arabidopsis</i> GLX2-2 ^b	0.4	0.8	0.30	1.5
<i>Arabidopsis</i> GLX2-5 ^c	0.58	0.61	ND ^e	1.19
<i>Salmonella</i> GLX2 ^d	0.21 ± 0.01	0.64 ± 0.02	0.30 ± 0.05	1.15 ± 0.28

^a From ref 10. ^b From ref 14, grown in ZY medium without extra addition of metal. ^c From ref 13, without extra added metal ion. ^d This work, grown in LB medium. ^e ND = not detected.

the recombinant enzymes (Table 1). In general, the wild-type forms of the recombinant GLX2 isoenzymes characterized up to now (GLX2 from *Homo sapiens* and *A. thaliana*) were isolated with significant but varying amounts of iron, zinc, and manganese bound to their active sites (10–14).

To study the metal content of a prokaryotic GLX2 recombinantly produced in *E. coli*, we overexpressed GloB either in a rich medium or in minimal medium supplemented with different metal ions. The identity of the metal ions bound to the protein obtained under different growth conditions was determined by atomic absorption measurements. In the case of the protein purified from LB medium, the metal content was 0.21 zinc, 0.64 iron, and 0.30 manganese per protein. These results are compared in Table 1 with those already reported for the human and *Arabidopsis* enzymes. It is clear that, as reported for the eukaryotic GLX2's, GloB is overexpressed with iron, zinc, and manganese bound. A metal occupancy lower than 2 may be attributed to a loss of metal during purification or to a limited metal availability during overexpression.

The kinetic parameters for the enzyme purified from a rich expression medium using SLG as the substrate are summarized in Table 2. The reported values are comparable to those obtained for GLX2 from other organisms. In an attempt to elucidate whether one of these metal ions is essential for the catalytic activity of GloB, we overexpressed the protein in minimal medium and supplemented it with different metal ions in the induction step. In this case, we were able to obtain different metal forms of GloB as described in Table 2. The catalytic efficiencies (k_{cat}/K_M) for the four forms of the protein were within the same order of magnitude and very close to those reported for other GLX2 family members, with a higher catalytic performance of the Mn variant. These assays revealed that GloB is able to bind the three different metal ions into its active site, yielding active forms in all cases.

Spectroscopic Studies on GloB. (1) Electron Paramagnetic Resonance. EPR spectra of GloB isolated from rich medium at 4.7 K (Figure 3A, upper trace) and 31 K (inset in Figure 3A) indicate the presence of four different sets of lines: (1) a resonance at $g_{eff} \approx 9.1$, (2) resonances at $g_{eff} \approx 4.3$, (3) a six-line pattern with hyperfine splitting of ca. 8 mT and centered around $g_{eff} \approx 2.0$, and (4) a broad resonance at $g_{eff} < 2.0$. Simulations of the EPR spectrum in Figure 3A using the software EasySpin (50) allowed the assignment of those sets of resonances as described below.

The first two sets ($g_{eff} \approx 9.1$ and 4.3) constitute the main components in the EPR spectra and are characteristic of magnetically isolated high-spin Fe(III) in a rhombic environment (51–53). Those lines arise from the ground-state m_S

Table 2: Steady-State Kinetic Constants and Metal Content of GloB Samples^a

	k_{cat} (s ⁻¹)	K_M (μM)	$k_{\text{cat}}/K_M \times 10^5$ (s ⁻¹ /M)	Fe/enzyme	Mn/enzyme	Zn/enzyme	total metal content
rich	168.8 ± 8.5	408 ± 50	3.5 ± 0.6	0.64 ± 0.02	0.30 ± 0.05	0.21 ± 0.01	1.15 ± 0.28
Zn-M9	209.9 ± 6.2	295 ± 27	7.1 ± 0.6	ND	ND	1.40 ± 0.02	1.40 ± 0.02
Fe-M9	149.9 ± 12.9	463 ± 10	3.2 ± 0.3	0.65 ± 0.01	ND	0.07 ± 0.01	0.72 ± 0.11
Mn-M9	394.9 ± 11.2	241 ± 18	16.4 ± 1.7	ND	1.10 ± 0.03	ND	1.10 ± 0.03

^a Both the metal content and the steady-state kinetics were determined as described in the Materials and Methods. Values are the average standard deviation obtained from independent measurements on at least three protein preparations. “rich” corresponds to the enzyme obtained from LB medium. “Zn/Fe/Mn-M9” corresponds to the enzyme obtained from minimal medium supplemented with the different metal ions. ND = not detected.

 Table 3: Crystallographic Data Collection and Refinement Statistics^a

data set	home source	LNLS
space group	$P2_1$	$P2_1$
cell dimens	a, b, c (Å)	44.13, 58.37, 43.37, 57.17, 54.76, 53.9
	β (deg)	112.08, 111.90
wavelength (Å)		1.541, 1.427
resolution (Å)		50.76 (1.67), 14.4 (1.45)
R_{merge}^b (%) [#]		5.7 (32), 7.3 (31.6)
$I/\sigma(I)$		9.2 (2.3), 6.7 (1.9)
completeness (%)		98.3 (88.9), 99.9 (99.9)
redundancy		4.8 (4.5), 5.4 (5.2)
no. of obsd reflns		142 464, 235 043
no. of unique reflns		29 723, 43 386
	Refinement	
no. of atoms		2670
$R_{\text{work}}/R_{\text{free}}$ (%)		15.7/17.8
rmsd	bond angles (deg)	0.930
	bond lengths (Å)	0.005
Ramachandran analysis (%/no.)	most favored	91.4/202
	additional allowed	8.1/18
	general allowed	0.5/1

^a Numbers in parentheses represent the highest resolution bin (1.53–1.45 Å). ^b $R_{\text{merge}} = \sum_h \sum_l |I_{hl} - \langle I_h \rangle| / \sum_h \sum_l \langle I_h \rangle$, where I_{hl} is the intensity of the l th observation of reflection h and $\langle I_h \rangle$ are the mean intensities of the h reflections. R_{merge} is computed over all l observations of h reflections.

$= \pm 1/2$ ($g_{\text{eff}} \approx 9.1$) and from the $m_S = \pm 3/2$ ($g_{\text{eff}} \approx 4.3$) states. The former states are populated only at very low temperatures, explaining why this resonance is not observed at 31 K (inset in Figure 3A). Simulation of such components using the spin Hamiltonian $H = \beta g_0 \mathbf{B} \cdot \mathbf{S} + \mathbf{S} \cdot \mathbf{D} \cdot \mathbf{S}$, where β is the Bohr magneton, \mathbf{B} is the applied magnetic field, and \mathbf{D} is the zero-field tensor used for $S > 1/2$ systems, resulted in a good agreement between experimental and calculated spectra (Figure 3A, lower trace) when the following parameters were used: $g_0 = 2.0$, $S = 5/2$, $D = 0.33 \text{ cm}^{-1}$, and $E = 0.0645 \text{ cm}^{-1}$ ($E/D = 0.195$). Strains in D (ΔD) and E (ΔE) values were also needed to achieve a better fit of the experimental data ($\Delta D = 0.18 \text{ cm}^{-1}$ and $\Delta E = 0.033 \text{ cm}^{-1}$). The agreement between calculated and experimental spectra was very sensitive to variations in these values, which enables us to estimate errors around $\sim 10\%$ for each parameter. The parameter values determined from our simulation are very close to those obtained by Marasinghe et al. (13) for mitochondrial glyoxalase II, thus suggesting similar local environments in both cases. The D value determined from our simulations (0.33 cm^{-1}) is also consistent with values obtained for other non-heme iron proteins (54) and indicates coordination of the Fe(III) ion to a ligand system primarily formed by oxygen and nitrogen atoms that makes D

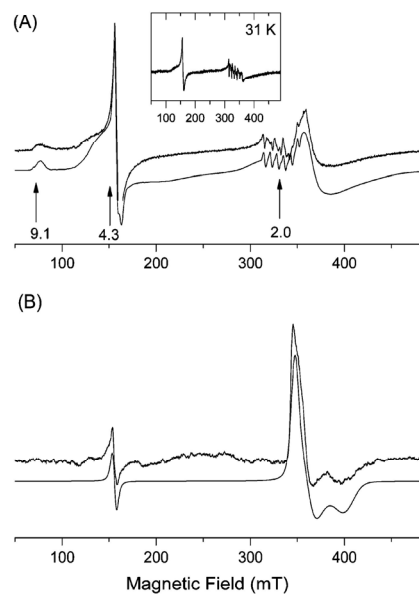


FIGURE 3: Experimental (upper traces) and calculated (lower traces) X-band EPR spectra from (A) 1 mM GloB and (B) 1 mM iron-only GloB. The inset in panel A shows the GloB spectrum measured at 31 K. The arrows indicate approximate effective g values.

approximately 0.5 cm^{-1} (51, 55). We also attempted to quantify the paramagnetic species on the basis of the integration of their calculated EPR spectra (Figure 3A, lower trace) with corrections for variations in g factors (56). We should however keep in mind that this is not very precise in determining spin concentrations in cases where $S > 1/2$ and multiple Kramer’s doublets are populated (small D value). Using this strategy, we estimated that resonances around 9.1 and 4.3 account for ca. 80% of the total iron content detected in this EPR spectrum. On the basis of the heterogeneous metal content of this sample, these signals can be tentatively attributed to an Fe(III)/Zn(II) center, as previously observed for other metalloproteins (13, 14, 57) (see below).

The third set of resonances comprises a sextet (region around $g \approx 2$ in Figure 3A) and was simulated assuming an apparent electron spin of $1/2$ ($g_{\text{eff}} = 1.99$) interacting with a nuclear spin $I = 5/2$ (calculated isotropic hyperfine interaction $A = 72 \text{ G}$), characteristic of a mononuclear Mn(II) center.

Finally, the broad signal at $g < 2$ has been previously assigned to an antiferromagnetically coupled center constituted by a high-spin ($S = 5/2$) Fe(III) ion and high-spin ($S = 2$) Fe(II) ion (13), and similar coupled Fe(III)/Fe(II)

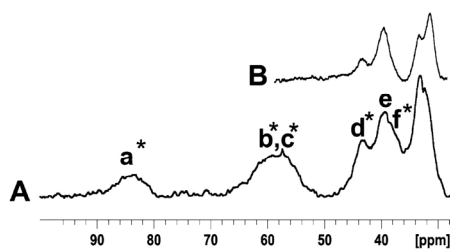


FIGURE 4: ^1H NMR spectra of GloB in (A) H_2O and (B) D_2O . The spectra were recorded at 600 MHz, pH 7.2, and 298 K in 10 mM MOPS buffer, 0.2 M NaCl.

binuclear sites have been observed for *A. thaliana* GLX2 (14) and calcineurin (57). This broad component is not observed in the EPR spectrum at 31 K (inset in Figure 3A), indicating a strong temperature dependence of this signal as expected for dinuclear coupled centers. Simulation of the spectrum assuming an $S = 1/2$ spin system allowed us to obtain effective g values of 1.87, 1.81, and 1.60 for this binuclear center (Figure 3A, lower trace). These g values are in good agreement with those determined for mitochondrial glyoxalase II (13). The use of a more detailed spin Hamiltonian, which included individual zero-field splitting for each ion and an exchange coupling, did not yield a unique set of parameters as also pointed out by Marasinghe et al. (13). Bearing in mind the limitations involved in quantifying the spin concentration in the present case, we determined that the broad signal at $g < 2$ accounts for roughly 20% of the total iron contributing to the EPR spectrum in Figure 3A.

The EPR spectrum of a sample of iron-only GloB (obtained from expression in minimal medium, as described above) at 4.7 K (Figure 3B, upper trace) is dominated by the signal at $g_{\text{eff}} < 2$ (associated with a coupled Fe(III)/Fe(II) center), with a concomitant decrease of the signal at $g_{\text{eff}} \approx 4.3$ corresponding to a magnetically isolated Fe(III) center. Simulation of this spectrum (Figure 3B, lower trace) led to similar effective g values (1.90, 1.83, 1.62) for the Fe(III)/Fe(II) species as compared to the simulation shown in Figure 3A (lower trace). However, line widths are somewhat narrower for the iron-only enzyme than observed for the multi-ion sample, thus leading to a better resolved spectrum (Figure 3B, lower trace). The resonance corresponding to mononuclear Fe(III) ($g_{\text{eff}} \approx 4.3$) was satisfactorily simulated with the same set of parameters employed for the previous sample. The component at 9.1 could not be detected on the basis of the low signal-to-noise ratio. The double integral of the calculated spectrum revealed that the coupled dinuclear Fe(III)/Fe(II) center represents ca. 90% of the total iron content, confirming a full metal upload in the protein expressed in minimal medium.

(2) *Nuclear Magnetic Resonance.* The ^1H NMR spectrum of GloB recorded under conditions tailored to optimize detection of the fast relaxing signals (48) close to a paramagnetic metal ion shows several hyperfine-shifted signals in the downfield region, up to 85 ppm (Figure 4). In this experiment, we expect to observe resonances only from iron ligands, since Mn(II) induces a considerable line broadening and yields no observable paramagnetic NMR spectrum.

The two most downfield shifted resonances, located at 84 (signal a) and 58 ppm, are absent when the spectrum is recorded in a D_2O solution. The signal envelope located at 58 ppm accounts for more than one proton, on the basis of its intensity and the fact that two partially overlapping signals can be better resolved at lower temperatures, and we label them as resonances b and c, accordingly. The signal envelope from 50 to 30 ppm includes several partially overlapping signals that are clearly sharper, but could not be resolved into individual components by changing the temperature. However, the intensity of resonances d and f at 43 and 38 ppm is decreased in the D_2O spectrum. These signals can be assigned to His metal ligands, suggesting that there are five His iron ligands. This is compatible with the presence of Fe(III) or Fe(II) ions in the two binding sites that in total bear five His ligands. This spectrum shows more signals than that reported by Marashinge et al. for the mitochondrial GLX2 from *Arabidopsis* (13). The narrow lines in the *Arabidopsis* enzyme have been interpreted as arising from a coupled Fe(III)/Fe(II) system. The present spectrum, despite displaying broader resonances, can only be compatible with a coupled Fe(III)/Fe(II) center. A magnetically isolated Fe(III) site (such as expected in an Fe(III)/Zn(II) center) would give rise to NMR resonances even broader (48, 58), and an uncoupled Fe(II) site renders narrower signals (59, 60). The NMR lines and chemical shifts are similar to those observed for purple acid phosphatases (58, 61, 62) that were shown to exhibit a weak antiferromagnetic coupling ($J \leq 10 \text{ cm}^{-1}$).

The NMR spectrum of iron-only GloB (obtained as described above) was identical to the one reported in Figure 4, but with a significantly higher signal-to-noise ratio. This confirms that these resonances correspond exclusively to the coupled Fe(III)/Fe(II) form, in agreement with the EPR spectrum, which revealed an increase of the contribution of this species from 20% to 90% on going from the heterogeneous sample to the iron-only protein.

Crystal Structure of GloB. (1) Overall Structure. The structure of GloB was solved by molecular replacement using the structural data of the mitochondrial GLX2 from *A. thaliana* (13). The top solutions of the rotation and translation functions had a Z score of 16.2 and 16.0, respectively. The final structure was refined to an R_{work} of 15.7% and an R_{free} of 17.8% at 1.45 Å resolution (Table 3). The asymmetric unit contains 1 protein molecule, and the final model contains all residues present in the amino acid sequence, 475 water molecules, and 2 metal ions. The electron density was well defined for the whole molecule except by a short loop near the C-terminus (Leu234–Arg240), with B factors around 20.3 \AA^2 against an average value of 14.5 \AA^2 , indicating some flexibility of this region compared to the rest of the protein.

As expected from the sequence homology, GloB shows the same protein fold observed in the *Arabidopsis* and human enzymes (10, 13), with a root-mean-square deviation (rmsd) of 1.0 and 0.8 Å in the $\text{C}\alpha$ positions, respectively (Figure 5A). The structure is composed by two tightly interacting domains. The N-terminal domain, which includes the first 165 residues, has the typical $\alpha\beta/\beta\alpha$ four-layer sandwich metallo- β -lactamase fold (15, 17). The smaller C-terminal domain is situated at one edge of the main core and consists of five helices. This domain is typical of the GLX2 family and is considered to be essential for substrate binding (10, 13, 63).

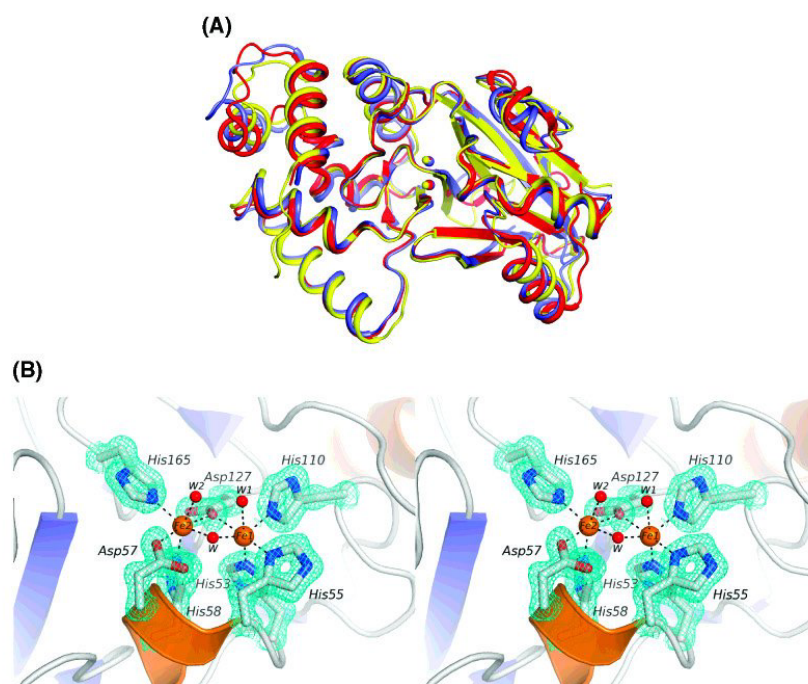


FIGURE 5: (A) Superposition of the structures of different GLX2's. The backbone traces of different glyoxalases II are shown: GloB (2QED, in blue; this work), the *Arabidopsis* mitochondrial isoform of GLX2 (1XM8, in red; 13), and human isoenzyme (1QH5, in yellow; 10). (B) Stereoview of the stick representation and the electron density map of the metal binding site of GloB. A H-bond network situated at the metal binding site. Fe1 is coordinated by N ϵ 2@His53, N δ 1@His55, and N ϵ 2@His110, while Fe2 is bound to O δ 1@Asp57, N ϵ 2@His58, and N ϵ 2@His165. The two metals are bridged by O δ 1@Asp127 and a water/OH molecule (W). Two water molecules (W1 and W2) are situated on the top of the metal site interacting with Fe1 and Fe2. The electron density map around the amino acids is a $[2F_o - F_c]$ map contoured at 1.5σ .

(2) *Metal Site Structure.* Seven protein residues and one water molecule directly interact with the two metal ions (Figure 5B). The two metals are separated by a distance of 3.3 Å and are bridged by a water molecule and the O δ 1 atom of Asp127. The identity of the metal ions cannot be unequivocally assessed from the electron density map. However, Zn(II)/Zn(II), Zn(II)/Mn(II), and Zn(II)/Fe(II) dinuclear centers are possibly less favorable since refinement attempts assuming these dinuclear centers gave rise to a strong negative signal around the density map of Zn(II) in $F_o - F_c$ even though the Zn atom occupancy is decreased. On the other hand, assumption of a diiron, a dimanganese, or a hybrid Fe/Mn center can account for the observed electron density maps; thus, none of these possibilities can be excluded. Therefore, since the metal content measurements show a predominance of iron in the active site, the final structure was refined by assuming the presence of two iron ions in its dinuclear active site. Analysis of the side chain atoms enrolled in the metal coordination showed the average B factor to be around 6.4 \AA^2 ; this value is similar to those observed for Fe1 and Fe2, which were 6.97 and 6.00 Å^2 , respectively. The small difference between Fe1/Fe2 B factors does not provide any evidence for differential occupancies for the two sites.

According to this final model, Fe1 is coordinated by His53 N ϵ 2, His55 N δ 1, and His110 N ϵ 2, while Fe2 is bound to Asp57 O δ 1, His58 N ϵ 2, and His165 N ϵ 2. All His ligands are bound through the N ϵ 2 atom, except in the case of His55,

in which the N δ 1 is the donor atom. This agrees with the His binding mode observed along the metallo- β -lactamase superfamily (15).

During the refinement, a weak electron density was located close to the active site. Different attempts to refine this density with a glutathione molecule were performed, but the low occupancy does not allow us to define the position nor the identity of this exogenous ligand. Therefore, the structure was refined with only two solvent molecules as apical metal ligands, one bound to each iron atom. The position of these water molecules is the same as occupied by the oxygen atoms of a cacodylate ion found in the structure of human GLX2 (1QH3) (10) and for a water molecule bound to Zn1 in the *Arabidopsis* GLX2 structure (1XM8) (13).

The position of the His side chains is determined by second-shell interactions, such as the hydrogen bonds subtended by the nonbinding N atom in the imidazole rings. Figure 6 shows second-shell interactions observed in the present structure. The carboxylate moiety of Asp28 is oriented in such a way that its O δ 2 is within H-bonding distance of the N δ 1 atom of His58 (2.71 Å), while its O δ 1 atom is located 2.71 Å from the backbone nitrogen of Thr52. Since Thr52 acts as a second-shell ligand of His53, with its O γ 1 as a H-bond acceptor from the His53 N δ 1 atom, Asp28 seems to play an essential role in building the structure of the metal site by orienting one His ligand from each of the two metal binding sites. His110 forms a H-bond through its N δ 1 and the main chain carbonyl oxygen from Arg136.

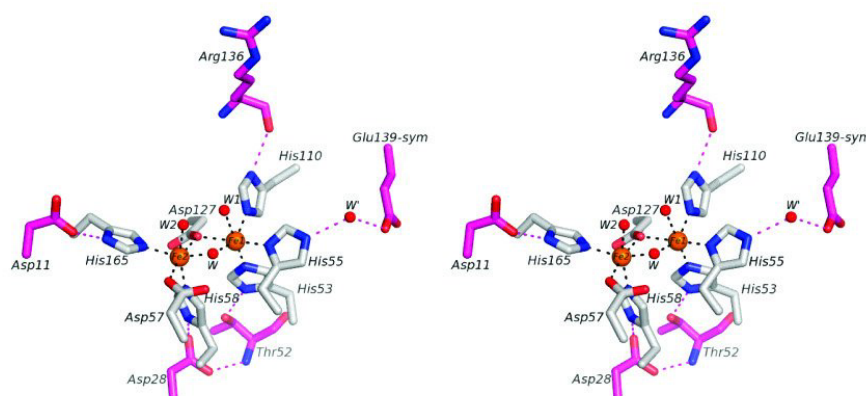


FIGURE 6: Schematic stereoview of the second-shell amino acid ligands of the metal binding site of GloB. The metal ions are shown as golden spheres. The metal ligands are colored in CPK mode, while carbon atoms from second-shell ligands are shown in purple. Metal–ligand bonds are shown as black dashed lines, while H-bonding interactions are shown as purple dashed lines.

Table 4: Differences in the Second-Shell Ligand Set among Structurally Characterized MBL Superfamily Members^a

PDB coordinates	metal ion	His ligand	metal–ligand distance (Å)	second-sphere ligand (distance, Å)
2QED	Fe2	Ne2@His165	2.08	Oδ1@Asp11 (2.80)
1QH5	Zn2	Ne2@His173	2.13	Oδ1@Asp11 (3.02)
1XM8	Fe2	Ne2@His169	2.00	Oδ1@Asp11 (2.89)
2BR6	Zn2	Ne2@His235	2.20	O@Pro34 (2.66)
1Y44	Zn2	Ne2@His269	2.29	Oγ1@Thr233 (2.97)
1P9E	Zn2	Ne2@His302	2.18	O@Ser96 (2.80)

^a The PDB files correspond to the following enzymes: 2QED (GloB from *S. typhimurium*; this work), 1QH5 (human GLX2; 10), 1XM8 (*A. thaliana* GLX2; 13), 2BR6 (homoserine lactone hydrolase; 23), 1Y44 (tRNaseZ; 65), 1P9E (methylparathion hydrolase; 22). The residue numbering of each protein has been maintained. The compared His ligand is homologous for all examined enzymes. The second-sphere ligands are residues forming a H-bond with the noncoordinating His Nδ1 atom.

Instead, His55 H-bonds to a solvent molecule that on its turn interacts with Oε2 from Glu139. Finally, His165 is oriented by an interaction with the Oδ1 from Asp11. Asp11 plays the same role in all studied GLX2's and seems to be unique to these enzymes, since it is not conserved in other members of the metallo-β-lactamase superfamily (Table 4 and Table 1 in the Supporting Information).

(3) *Substrate Binding Site.* The crystal structure of human GLX2 complexed with glutathione and *S*[(hydroxybromophenyl)carbamoyl]glutathione showed that an arginine residue, two lysine residues, and two tyrosine residues (Arg249, Lys252, Lys143, Tyr175, and Tyr145 in the human GLX2 sequence (10)) are within 13 Å of Zn1 and interact with these ligands. These residues involved in substrate binding in human GLX2 are conserved in GloB, except for Tyr145 and Lys143, which are substituted by a Phe and Arg residue, respectively (Phe138 and Arg136 in the GloB sequence). In the structure of the complexed human enzyme, the phenolic OH@Tyr145 interacts with a carbonyl group from glutathione. This interaction cannot be possible when this residue is replaced by a Phe in GloB. Multiple sequence alignment of different glyoxalases II reveals that an aromatic residue is fully conserved in this position in all enzymes. GLX2's from mammals are unique in harboring a Tyr residue in this position, while the rest (including prokaryotic, fungal,

plant, and yeast enzymes) possess a Phe residue. We therefore conclude that this interaction is not essential for substrate recognition, yet the residue substitution among different species remains an interesting issue to solve.

DISCUSSION

In this work we have accomplished a structural and biochemical characterization of GloB, the GLX2 from *S. typhimurium*. This enzyme recombinantly produced in *E. coli* was isolated with significant amounts of iron, manganese, and zinc bound to its active site, with a clear preference toward iron. Recombinant GloB displayed similar catalytic efficiencies with different ratios of active site bound iron, manganese, and zinc. This feature is indicative of certain flexibility of the active site of GloB to accommodate any of these metal ions and to carry out the hydrolytic reaction with similar efficiency in all cases. The Mn-only variant displays a somehow larger catalytic efficiency and k_{cat} values. Interestingly, the recently characterized GLX2 from *E. coli*, originally isolated in the dizinc form, also shows a high activity in the reconstituted Mn(II) form (33).

EPR spectroscopy reveals the presence of a magnetically uncoupled Fe(III) center, a magnetically spin coupled Fe(III)/Fe(II) center, and an uncoupled Mn(II) center, thus indicating that a range of metal–metal combinations are incorporated during the overexpression. NMR data are consistent with the presence of an Fe(III)/Fe(II) center with a weak antiferromagnetic coupling. This is in agreement with the crystal structure that shows a bridged dimetallic center. The EPR and NMR spectra of the Fe-only sample obtained by expression in minimal medium supplemented with iron reveal that this enzyme harbors exclusively a coupled Fe(III)/Fe(II) dinuclear site.

The binding of different metal ions may be a consequence of the overexpression of this protein in *E. coli* and may not represent the physiologically active metal-loaded species. One possible explanation for this observation is that a limited metal availability during overexpression forces the nascent polypeptide to bind other metals in addition to the native one. Another hypothesis could be that GLX2 employs different metal uptake machinery on its native environment that is lacking in the heterologous host. This hypothesis could

Biochemistry

explain the behavior observed in the eukaryotic proteins, but may not hold for a *Salmonella* protein overexpressed in *E. coli*. Finally, overexpression of different members of the metallo- β -lactamase superfamily yields exclusively Zn(II) variants, GLX2 being the only exception. This strongly suggests that there must be other features related to the enzyme itself to explain the preference for iron and the metal ion promiscuity.

The immediate environment of metal sites in metalloproteins determines the selectivity in the metal ion uptake (64). A detailed inspection of the enzyme metal site compared to that in other systems reveals that the metal binding site of GLX2 is identical to that found in different members of the superfamily that have been characterized as Zn-only proteins, such as phosphodiesterase (19), acyl-homoserine-lactone lactonase (20, 23, 24), methylparathion hydrolase, and alkylsulfatase (21). Instead, some differences arise in analysis of the second-shell amino acid ligands (Table 4, Figure 6, and Table 1 in the Supporting Information). Asp11 forms a strong II-bond with His165, a Site2 ligand. This residue is conserved in all known GLX2's, but is replaced by Thr or Ser residues in other superfamily members. An Asp residue in this position is expected to increase the hardness of the metal binding site (favoring iron uptake), as well as the negative charge of the ligand set, favoring binding of trivalent cations to the active site. We therefore propose that this second-shell amino acid ligand favors iron uptake in GLX2. Interestingly, both GLX1 and GLX2 enzymes display a subtle control of metal ion recognition. However, the metal ion requirements (either Ni(II) or Zn(II)) are strict among GLX1's, while GLX2's seem to show no preference regarding the different metal cations to render a preferentially more active enzyme.

The present study shows that bacterial GLX2's are active as diiron enzymes. The isolation of *E. coli* GLX2 (which displays a 78% sequence identity with GloB) as a dizinc enzyme is reported in a recent work. The study of the iron variant of the *E. coli* enzyme, which has not been characterized yet, would provide considerable insight into the understanding of metal dependence among prokaryotic glyoxalases II. Although the role of GloB in *Salmonella* pathogenesis has not yet been studied, a recent report showed that STM3117 (coding for a GLX1) contributes to the replication of the pathogen inside macrophages (30). Thus, the glyoxalase system is expected to play an important role in the survival of the pathogen under these conditions. Fe(II) concentrations above 5–10 μ M have been reported (31) to provide an appropriate environment for the metal uptake event of this enzyme.

These observations, together with the presence of two other homologous genes coding for GLX1 isoenzymes in *Salmonella* allow us to postulate that proper methylglyoxal processing would play a major role in *Salmonella*'s survival, probably both inside and outside the host. It is possible that these three enzymes would display differential substrate selectivity. Since it is expected that GloB will second STM3117 in the methylglyoxal pathway, a direct correlation of the *Salmonella* GLX2 in virulence can be predicted.

ACKNOWLEDGMENT

We thank Dr. R. Girolami for atomic absorption measurements.

Characterization of *Salmonella typhimurium* Glyoxalase II I

SUPPORTING INFORMATION AVAILABLE

A table listing the first- and second-shell ligands commonly found in some representative members of the metallo- β -lactamase superfamily. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

- Thornalley, P. J. (1993) The glyoxalase system in health and disease, *Mol. Aspects Med.* 14, 287–371.
- Thornalley, P. J. (1998) Glutathione-dependent detoxification of alpha-oxoaldehydes by the glyoxalase system: involvement in disease mechanisms and antiproliferative activity of glyoxalase I inhibitors, *Chem.-Biol. Interact.* 111–112, 137–151.
- Ferguson, G. P., Totemeyer, S., MacLean, M. J., and Booth, I. R. (1998) Methylglyoxal production in bacteria: suicide or survival? *Arch. Microbiol.* 170, 209–218.
- Kalapos, M. P. (1999) Methylglyoxal in living organisms: chemistry, biochemistry, toxicology and biological implications, *Toxicol. Lett.* 110, 145–175.
- Thornalley, P. J. (1996) Pharmacology of methylglyoxal: formation, modification of proteins and nucleic acids, and enzymatic detoxification—a role in pathogenesis and antiproliferative chemotherapy, *Gen. Pharmacol.* 27, 565–573.
- Clugston, S. L., Barnard, J. F., Kinach, R., Miedema, D., Ruman, R., Daub, E., and Honek, J. F. (1998) Overproduction and characterization of a dimeric non-zinc glyoxalase I from *Escherichia coli*: evidence for optimal activation by nickel ions, *Biochemistry* 37, 8754–8763.
- Sukdeo, N., Clugston, S. L., Daub, E., and Honek, J. F. (2004) Distinct classes of glyoxalase I: metal specificity of the *Yersinia pestis*, *Pseudomonas aeruginosa* and *Neisseria meningitidis* enzymes, *Biochem. J.* 384, 111–117.
- Creighton, D. J., and Hamilton, D. S. (2001) Brief history of glyoxalase I and what we have learned about metal ion-dependent, enzyme-catalyzed isomerizations, *Arch. Biochem. Biophys.* 387, 1–10.
- Armstrong, R. N. (2000) Mechanistic diversity in a metalloenzyme superfamily, *Biochemistry* 39, 13625–13632.
- Cameron, A. D., Ridderstrom, M., Olin, B., and Mannervik, B. (1999) Crystal structure of human glyoxalase II and its complex with a glutathione thioester substrate analogue, *Struct. Folding Des.* 7, 1067–1078.
- Zang, T. M., Hollman, D. A., Crawford, P. A., Crowder, M. W., and Makaroff, C. A. (2001) Arabidopsis glyoxalase II contains a zinc/iron binuclear metal center that is essential for substrate binding and catalysis, *J. Biol. Chem.* 276, 4788–4795.
- Wenzel, N. F., Carenbauer, A. L., Pfister, M. P., Schilling, O., Meyer-Klaucke, W., Makaroff, C. A., and Crowder, M. W. (2004) The binding of iron and zinc to glyoxalase II occurs exclusively as di-metal centers and is unique within the metallo-beta-lactamase family, *J. Biol. Inorg. Chem.* 9, 429–438.
- Marasinghe, G. P., Sander, I. M., Bennett, B., Periyannan, G., Yang, K. W., Makaroff, C. A., and Crowder, M. W. (2005) Structural studies on a mitochondrial glyoxalase II, *J. Biol. Chem.* 280, 40668–40675.
- Schilling, O., Wenzel, N., Naylor, M., Vogel, A., Crowder, M., Makaroff, C., and Meyer-Klaucke, W. (2003) Flexible metal binding of the metallo-beta-lactamase domain: glyoxalase II incorporates iron, manganese, and zinc in vivo, *Biochemistry* 42, 11777–11786.
- Daiyasu, H., Osaka, K., Ishino, Y., and Toh, H. (2001) Expansion of the zinc metallo-hydrolase family of the beta-lactamase fold, *FEBS Lett.* 503, 1–6.
- Gomes, C. M., Frazao, C., Xavier, A. V., LeGall, J., and Teixeira, M. (2002) Functional control of the binuclear metal site in the metallo-beta-lactamase-like fold by subtle amino acid replacements, *Protein Sci.* 11, 707–712.
- Carfi, A., Pares, S., Duee, E., Galleni, M., Duee, C., Frère, J. M., and Dideberg, O. (1995) The 3-D structure of a zinc metallo-beta-lactamase from *Bacillus cereus* reveals a new type of protein fold, *EMBO J.* 14, 4914–4921.
- Crowder, M. W., Spencer, J., and Vila, A. J. (2006) Metallo-beta-lactamases: Novel Weaponry for Antibiotic Resistance in Bacteria, *Acc. Chem. Res.* 39, 721–728.

19. Vogel, A., Schilling, O., Niecke, M., Bettmer, J., and Meyer-Klaucke, W. (2002) ElaC encodes a novel binuclear zinc phosphodiesterase, *J. Biol. Chem.* **277**, 29078–29085.
20. Thomas, P. W., Stone, E. M., Costello, A. L., Tierney, D. L., and Fast, W. (2005) The quorum-quenching lactonase from *Bacillus thuringiensis* is a metalloprotein, *Biochemistry* **44**, 7559–7569.
21. Hagelueken, G., Adams, T. M., Wiehlmann, L., Widow, U., Kolmar, H., Tummeler, B., Heinz, D. W., and Schubert, W. D. (2006) The crystal structure of SdsA1, an alkylsulfatase from *Pseudomonas aeruginosa*, defines a third class of sulfatases, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 7631–7636.
22. Dong, Y. J., Bartlam, M., Sun, L., Zhou, Y. F., Zhang, Z. P., Zhang, C. G., Rao, Z., and Zhang, X. E. (2005) Crystal structure of methyl parathion hydrolase from *Pseudomonas* sp. WBC-3, *J. Mol. Biol.* **353**, 655–663.
23. Kim, M. H., Choi, W. C., Kang, H. O., Lee, J. S., Kang, B. S., Kim, K. J., Derewenda, Z. S., Oh, T. K., Lee, C. H., and Lee, J. K. (2005) The molecular structure and catalytic mechanism of a quorum-quenching N-acyl-L-homoserine lactone hydrolase, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 17606–17611.
24. Liu, D., Lepore, B. W., Petsko, G. A., Thomas, P. W., Stone, E. M., Fast, W., and Ringe, D. (2005) Three-dimensional structure of the quorum-quenching N-acyl homoserine lactone hydrolase from *Bacillus thuringiensis*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 11882–11887.
25. Frazao, C., Silva, G., Gomes, C. M., Matias, P., Coelho, R., Sieker, L., Macedo, S., Liu, M. Y., Oliveira, S., Teixeira, M., Xavier, A. V., Rodrigues-Pousada, C., Carrondo, M. A., and Le, Gall, J. (2000) Structure of a dioxygen reduction enzyme from *Desulfovibrio gigas*, *Nat. Struct. Biol.* **7**, 1041–1045.
26. Bito, A., Haider, M., Hadler, I., and Breitenbach, M. (1997) Identification and phenotypic analysis of two glyoxalase II encoding genes from *Saccharomyces cerevisiae*, GLO2 and GLO4, and intracellular localization of the corresponding proteins, *J. Biol. Chem.* **272**, 21509–21519.
27. Maiti, M. K., Krishnasamy, S., Owen, H. A., and Makaroff, C. A. (1997) Molecular characterization of glyoxalase II from *Arabidopsis thaliana*, *Plant Mol. Biol.* **35**, 471–481.
28. Cordell, P. A., Futers, T. S., Grant, P. J., and Pease, R. J. (2004) The Human hydroxyacylglutathione hydrolase (HAGH) gene encodes both cytosolic and mitochondrial forms of glyoxalase II, *J. Biol. Chem.* **279**, 28653–28661.
29. McClelland, M., Sanderson, K. E., Spieth, J., Clifton, S. W., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F., Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Grewal, N., Mulvaney, E., Ryan, E., Sun, H., Florea, L., Miller, W., Stoneking, T., Nhan, M., Waterston, R., and Wilson, R. K. (2001) Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2, *Nature* **413**, 852–856.
30. Shi, L., Adkins, J. N., Coleman, J. R., Schepmoes, A. A., Dohinkova, A., Mottaz, H. M., Norbeck, A. D., Purvine, S. O., Manes, N. P., Smallwood, H. S., Wang, H., Forbes, J., Gros, P., Uzzau, S., Rodland, K. D., Heffron, F., Smith, R. D., and Squier, T. C. (2006) Proteomic analysis of *Salmonella enterica* serovar typhimurium isolated from RAW 264.7 macrophages: identification of a novel protein that contributes to the replication of serovar typhimurium inside macrophages, *J. Biol. Chem.* **281**, 29131–29140.
31. Eriksson, S., Lucchini, S., Thompson, A., Rhen, M., and Hinton, J. C. (2003) Unravelling the biology of macrophage infection by gene expression profiling of intracellular *Salmonella enterica*, *Mol. Microbiol.* **47**, 103–118.
32. Eskra, L., Canavessi, A., Carey, M., and Splitter, G. (2001) *Brucella abortus* genes identified following constitutive growth and macrophage infection, *Infect. Immun.* **69**, 7736–7742.
33. O'young, J., Sukdeo, N., and Honek, J. F. (2007) *Escherichia coli* glyoxalase II is a binuclear zinc-dependent metalloenzyme, *Arch. Biochem. Biophys.* **459**, 20–26.
34. Checa, S. K., Epariz, M., Audero, M. E., Botta, P. E., Spinelli, S. V., and Soncini, F. C. (2007) Bacterial sensing of and resistance to gold salts, *Mol. Microbiol.* **63**, 1307–1318.
35. Gill, S. C., and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data, *Anal. Biochem.* **182**, 319–326.
36. Ridderstrom, M., Saccucci, F., Hellman, U., Bergman, T., Principato, G., and Mannervik, B. (1996) Molecular cloning, heterologous expression, and characterization of human glyoxalase II, *J. Biol. Chem.* **271**, 319–323.
37. Polikarpov, I., Perles, L. A., de Oliveira, R. T., Oliva, G., Castellano, E. E., Garratt, R. C., and Craievich, A. (1998) Set-up and Experimental Parameters of the Protein Crystallography Beamline at the Brazilian National Synchrotron Laboratory, *J. Synchrotron Radiat.* **5**, 72–76.
38. Polikarpov, I., Oliva, G., Castellano, E. E., Garratt, R. C., Armada, P., Leite, A., and Craievich, A. (1997) The protein crystallography beamline at LNLS, the Brazilian National Synchrotron Light Source, *Nucl. Instrum. Methods, A* **405**, 159–164.
39. Leslie, A. G. W. (1992) Recent changes to the MOSFLM package for processing film and image data, *Jt. CCP4 ESF-EADBM Newsl. Protein Crystallogr.* **26**, 27–33.
40. Collaborative Computational Project Number 4 (1994) The CCP4 Suite: Programs for Protein Crystallography, *Acta Crystallogr., D: Biol. Crystallogr.* **50**, 760–763.
41. Read, R. J. (2001) Pushing the boundaries of molecular replacement with maximum likelihood, *Acta Crystallogr., D: Biol. Crystallogr.* **57**, 1373–1382.
42. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr., D: Biol. Crystallogr.* **60**, 2126–2132.
43. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr., D: Biol. Crystallogr.* **53**, 240–255.
44. Afonine, P. V., Grosse-Kunstleve, R. W., and Adams, P. D. (2005) *Acta Crystallogr., D: Biol. Crystallogr.* **61**, 850–855.
45. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr.* **26**, 283–291.
46. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA.
47. Inubushi, T., and Becker, E. D. (1983) Efficient detection of paramagnetically shifted NMR resonances by optimizing the WEFT pulse sequence, *J. Magn. Reson.* **51**, 128–133.
48. Bertini, I., Turano, P., and Vila, A. J. (1993) NMR of paramagnetic metalloproteins, *Chem. Rev.* **93**, 2833–2932.
49. Bito, A., Haider, M., Briza, P., Strasser, P., and Breitenbach, M. (1999) Heterologous expression, purification, and kinetic comparison of the cytoplasmic and mitochondrial glyoxalase II enzymes, Glo2p and Glo4p, from *Saccharomyces cerevisiae*, *Protein Expr. Purif.* **17**, 456–464.
50. Stoll, S., and Schweiger, A. (2006) Easy Spin, a Comprehensive Software Package for Spectral Simulation and Analysis in EPR, *J. Magn. Reson.* **178**, 42–55.
51. Wickman, H. H., Klein, M. P., and Shirley, D. A. (1965) Paramagnetic Resonance of Fe³⁺ in Polycrystalline Ferrichrome A, *J. Chem. Phys.* **43**, 2113–2117.
52. Dowsing, R. D., and Gibson, J. F. (1969) Electron Spin Resonance of High-Spin d⁵ Systems, *J. Chem. Phys.* **50**, 294–303.
53. Aasa, R. (1970) Powder Line Shape in the Electron Paramagnetic Resonance Spectra of High-Spin Ferric Complexes, *J. Chem. Phys.* **52**, 3919–3930.
54. Pinkowitz, R. A., and Aisen, P. (1972) Zero-field splittings of ***iron** complexes of transferrins, *J. Biol. Chem.* **247**, 7830–7834.
55. Aasa, R., Carlsson, K. E., Reyes, L. S. A., and Vanngard, T. (1966) Fine and hyperfine structure in electron spin resonance spectrum of Fe³⁺ in ethylenediaminetetraacetic acid (EDTA), *Arch. Kemi* **25**, 285–291.
56. Pilbrow, J. R. (1990) *Transition Ion Electron Paramagnetic Resonance*, Clarendon Press, Oxford.
57. Yu, L., Haddy, A., and Rusnak, F. (1995) Evidence that Calcineurin accommodates an Active Site Binuclear Metal Center, *J. Am. Chem. Soc.* **117**, 10147–10148.
58. Lauffer, R. B., Antanaitis, B. C., Aisen, P., and Que, L., Jr. (1983) 1H-NMR studies of porcine uteroferrin, *J. Biol. Chem.* **258**, 14212–14218.
59. Sorkin, D. L., and Miller, A. F. (2000) Amino acid-specific isotopic labeling and active site NMR studies of iron(II)- and iron(III)-superoxide dismutase from *Escherichia coli*, *J. Biomol. NMR* **17**, 311–322.
60. Sorkin, D. L., and Miller, A. F. (1997) Spectroscopic measurement of a long-predicted active site pK in iron-superoxide dismutase from *Escherichia coli*, *Biochemistry* **36**, 4916–4924.
61. Wang, Z., Ming, L. J., Que, L., Jr., Vincent, J. B., Crowder, M. W., and Averill, B. A. (1992) 1H-NMR and NOE studies of the purple acid phosphatases from porcine uterus and bovine spleen, *Biochemistry* **31**, 5263–5268.

Biochemistry

PAGE EST: 10.1

Characterization of *Salmonella typhimurium* Glyoxalase II K

62. Scarrow, R. C., Pyrz, J. W., and Que, L., Jr. (1990) NMR studies of the dinuclear iron site in reduced uteroferrin and its oxoanion complexes, *J. Am. Chem. Soc.* *112*, 657–665.
63. Park, H. S., Nam, S. H., Lee, J. K., Yoon, C. N., Mannervik, B., Benkovic, S. J., and Kim, H. S. (2006) Design and evolution of new catalytic activity with an existing protein scaffold, *Science* *311*, 535–538.
64. Dudev, T., and Lim, C. (2003) Principles governing Mg, Ca, and Zn binding and selectivity in proteins, *Chem. Rev.* *103*, 773–788.
65. de la Sierra-Gallay, I. L., Pellegrini, O., and Condon, C. (2005) Structural basis for substrate binding, cleavage and allostery in the tRNA maturase RNase Z, *Nature* *433*, 657–661.

BI7007245