

UNIVERSIDADE DE BRASÍLIA - UnB
INSTITUTO DE CIÊNCIAS BIOLÓGICAS - IB
DEPARTAMENTO DE BIOLOGIA CELULAR - CEL
LABORATÓRIO DE BIOLOGIA DO GENE - LaBioGene

**IDENTIFICAÇÃO E CARACTERIZAÇÃO
DE UMA PROTEÍNA COM MOTIVOS**

ZINC FINGER de Trypanosoma cruzi

ALESSANDRA R. E. DE OLIVEIRA

Brasília, DF

2006

UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOLOGIA CELULAR
LABORATÓRIO DE BIOLOGIA DO GENE

IDENTIFICAÇÃO E CARACTERIZAÇÃO
DE UMA PROTEÍNA COM MOTIVOS
ZINC FINGER de Trypanosoma cruzi

ALESSANDRA R. E. DE OLIVEIRA

Tese apresentada ao Departamento de Biologia Celular do
Instituto de Ciências Biológicas da Universidade de Brasília
como requisito parcial para obtenção do título de Doutor em
Biologia Molecular

Orientadora: Dr^a Beatriz Dolabela de Lima

Brasília, DF

2006

“Nobody is so wise that he doesn’t have something to learn. Nobody is so simple that he doesn’t have something to teach!” (Unknowledge author)

**Dedico esta tese ao meu Deus, ao meu
esposo Mauro e aos meus pais Carlinho e
Cleisy.**

AGRADECIMENTOS

À Profa.Dra. Beatriz Dolabela de Lima (Bia), pela orientação e oportunidade de realização deste trabalho, além da amizade, apoio e confiança depositada em mim.

Ao Prof.Dr. Cézar Martins de Sá, pelas discussões e sugestões durante o trabalho, e por quem eu também tenho uma grande amizade e admiração.

Ao Prof.Dr. Carlos Roberto Félix pelo apoio ao meu ingresso na pós-graduação em Biologia Molecular assim que cheguei a Brasília.

Às Profas. Dras Lídia Maria Pepe de Moraes e Sônia Freitas, pelos incentivos, compreensão e excelentes conselhos durante a minha qualificação.

Às Profas. Dras Lenise, Cynthia, Loreny e Marlene pela amizade, compreensão e incentivo.

Às Profas. Dras Sílvia, Rosália e Célia por terem me proporcionado a oportunidade do primeiro contato com a Biologia Molecular na UFG. Também aos amigos daquela época: Wesley, Valéria, Andréia, Chris, Thais, Renata, Letícia, Adriano e o inesquecível Wilsinho!

Aos amigos e colegas do Laboratório de Biologia do Gene: Lilica, Dany, Nany, Ana Paula, Lalá e Celsinho pela grande família que nos tornamos em todos aqueles momentos difíceis e também os de muita alegria.

Aos ex-colegas do Laboratório de Biologia do Gene: Sílvia, Eduardo, Fabian, Camila, Ana, Érika e Davi.

Aos amigos do laboratório de Microbiologia: Inaiara, Heide, Alex, Rosana, Kalinka e Emerson.

Aos amigos do Laboratório de Biologia Molecular: Chrisinha, Alessandra (Ildinete), Fábio, Alessandra Dantas, Bruno Daher, Túlio, João Ricardo, Viviane, Livônios, Diorge, Fabrício, Patvet, Patgirl, Loíse, Saulo, Nádia, Menino Antônio, Luane, Indra, Luna, Henrique, Maria José, Guilherme, Ivana, Lorena, Tatiane Iembo, Carmela, Rosana, Maria, Daniel, André, e aqueles que

por acaso não lembrava neste momento, pela ajuda e amizade demonstrada durante a realização deste trabalho.

Aos amigos do laboratório de Microscopia Eletrônica, especialmente ao Prof Dr. Bergmann, Hugo, Jose e Diana.

Aos amigos do laboratório de Farmacologia Molecular: Rutinha, Ranieri, Gustavo e Guilherme.

Aos amigos do Cenargen-Embrapa, especialmente Ana (microscopia), Eduardo e Danielle Cordeiro.

À secretária e amiga Ana, sempre disposta a resolver os problemas dos alunos da pós-graduação em Biologia Molecular.

Aos funcionários do laboratório de Biologia do Gene: Liliam, dona Hilda e Tânia pelo apoio técnico durante a realização deste trabalho. Em especial a Liliam por ter me ensinado tanto a trabalhar tão corretamente com cultura de células e pela grande amizade que fizemos nestes 5 anos de UnB.

À tia Deusáida pelo carinho e alegria que sempre esteve estampada em seu rosto para todas as horas.

Aos amigos da FIPLAC e em especial a Nicinha, Dr. Antônio Carlos e Regina.

Aos colegas e amigos da NOVO Nordisk Produção Farmacêutica do Brasil por me encorajarem e torcerem tanto para finalização desta tese, em especial: Ângela, Karina, Jorge, Marina, Lília e Almerinda.

Aos amigos de outros departamentos e laboratórios pela ajuda, amizade e incentivo.

Aos tios Cleila, Adalberto, Cleide e Wálter pelo apoio no começo de tudo, sem o qual nada disso teria sido possível!

Aos tios Olguinha, Climeny, Bill e Luciano pela amizade de sempre e conversas gostosas em Ipameri e Goiânia.

Aos tios Arnaldo e Clemei e aos meus primos e amigos Tereza Cristina e Douglas que me acolheram em sua casa quando decidi vir para Brasília, pelo carinho e preocupação comigo durante este período de convivência.

Aos primos Bia e Junior, Maroca e tio Disson pelo grande apoio e amizade.

Aos meus pimpolhos primos Fê, Pipo e Dudu que me trouxeram tanta alegria pela pureza de crianças alegres.

Às grandes amigas Leilah e Mércia pela linda amizade que fizemos.

Às irmãs do Educandário Espírito Santo pela amizade, carinho e por me proporcionarem um lugarzinho aconchegante e tão sossegado pertinho da UnB, do qual nunca vou esquecer.

Ao meu amigo e esposo Mauro que esteve tão presente durante esta tese me dando apoio pessoal, profissional e tanto amor! Por ser um presente de Deus na minha vida!

Aos meus pais, Carlinho e Cleisy que, mesmo distantes, sempre me deram o carinho e os incentivos necessários e por me amarem incondicionalmente.

Aos meus irmãos, Alessandro e Alexandre que são meu orgulho, por sermos tão unidos apesar da distância física.

À minha linda e tão querida sobrinha Vitória que mesmo distante está sempre em meu coração e minhas orações.

A todos que, de uma forma ou de outra colaboraram para a realização deste trabalho.

À Capes pela concessão de auxílio financeiro.

ÍNDICE GERAL

Índice das Figuras.....	i
Abreviaturas.....	ii
Resumo.....	iv
Abstract.....	v
INTRODUÇÃO.....	01
I – Estado da arte.....	01
1 – <i>Trypanosoma cruzi</i> e Doença de Chagas.....	01
2 – Biologia Molecular dos Tripanossomatídeos.....	02
3 – Proteínas de ligação a ácidos nucleicos.....	06
3.1 – <i>Helix-turn-helix</i> e homeodomínio.....	07
3.2 – <i>Leucine Zipper</i>	08
3.3 – <i>Zinc finger</i>	08
II – Relevância e Justificativa.....	33
III – Objetivos.....	36
MATERIAIS E MÉTODOS.....	37
I - Linhagens de células utilizadas.....	37
II – Plasmídeos.....	38
III - Meios de cultura.....	38
IV - Reagentes e Soluções.....	40
V - Métodos utilizados.....	45
RESULTADOS.....	54
I - Identificação de um cluster contendo três genes <i>in tandem</i> codificando para proteínas zinc finger CCHC em <i>T. cruzi</i>	54
II – Análise do gene <i>TcZFP8</i> e da sua proteína predita.....	59
III - Análise da expressão do gene <i>TcZFP8</i> em células de <i>T. cruzi</i> por Northern blot.....	63
IV – Expressão do gene <i>TcZFP8</i> como proteína de fusão com a glutathione S transferase de <i>Shistosoma japonicum</i> e a purificação da proteína recombinante.....	63
V - Análise da interação da <i>TcZFP8</i> com ácidos nucleicos por SELEX e <i>gel shift</i>	67
VI - Citolocalização da proteína <i>TcZFP8</i> em células de <i>T. cruzi</i> por intermédio da fusão com a proteína verde fluorescente.....	72

VII - Expressão em <i>E. coli</i> da TcZFP8 com fusão a His-tag, purificação da proteína recombinante e produção de anticorpo policlonal contra esta proteína.....	76
VIII - Análise da expressão e localização da proteína TcZFP8 em lisados de células de <i>T. cruzi</i> por <i>Western blot</i>	79
IX - Imunocitocalização da proteína TcZFP8 em formas epimastigotas de <i>T. cruzi</i>	79
DISCUSSÃO.....	82
CONCLUSÃO.....	90
BIBLIOGRAFIA.....	91
ANEXO.....	100

ÍNDICE DAS FIGURAS

Figura 01 – Ciclo de vida do <i>Trypanosoma cruzi</i>	03
Figura 02 – Motivo Zinc Finger	10
Figura 03 – Motivo Zinc Finger da proteína TFIIIA de <i>Xenopus</i>	11
Figura 04 – A proteína Zif 268	12
Figura 05 – O contato de Zif 268 com o DNA	13
Figura 06 – Estrutura dos domínios dos receptores nucleares	17
Figura 07 – Domínio de ligação ao DNA (DBD)	18
Figura 08 – Domínio de ligação ao ligante (LBD)	19
Figura 09 – Transdução de sinal de hormônios e outros ligantes	21
Figura 10 – Estrutura do Epf	24
Figura 11 – Mecanismo de crescimento tumoral por Epf	25
Figura 12 – Estrutura dos domínios <i>ring finger</i>	26
Figura 13 – Esquema da região genômica de 17 kb da banda XX de <i>T. cruzi</i>	35
Figura 14 – Comparação de regiões homólogas de tripanosomatídeos	55
Figura 15 – Árvore filogenética de proteínas <i>zinc finger</i>	58
Figura 16 – Seqüência nucleotídica do gene <i>TcZFP8</i> e sua proteína predita	60
Figura 17 – Alinhamento dos motivos CCHC da <i>TcZFP8</i>	61
Figura 18 – Esquema geral	62
Figura 19 – Análise de <i>Northern blot</i> de RNA total de <i>T. cruzi</i>	64
Figura 20 – Mapa genético do plasmídeo pEXZ	65
Figura 21 – Análise em gel de agarose dos plasmídeos pEXZ 18, 19 e 23	66
Figura 22 – Análise da expressão de <i>TcZFP8</i> fusionada a GST	68
Figura 23 – Expressão e purificação da GST e GST- <i>TcZFP8</i>	69
Figura 24 – Oligonucleotídeos para o experimento de SELEX	70
Figura 25 – Estratégia do experimento de SELEX	71
Figura 26 – Alinhamento das seqüências do SELEX e <i>gel shift</i>	73
Figura 27 – Mapa genético do plasmídeo p33GZ	74
Figura 28 – Análise em gel de agarose dos plasmídeos p33GZ	75
Figura 29 – Mapa genético do plasmídeo pETZ	77
Figura 30 – Purificação da proteína <i>TcZFP8</i> – His tag	78
Figura 31 – Análise da expressão de <i>TcZFP8</i> por <i>Western Blot</i> de lisados de <i>T. cruzi</i>	80
Figura 32 – Citolocalização por imunofluorescência da proteína <i>TcZFP8</i> em <i>T. cruzi</i>	81
Figura 33 – Esquema da região genômica contendo os três genes <i>zinc finger</i>	83

ABREVIATURAS

1x - concentração de trabalho
5x - cinco vezes concentrado
10x - dez vezes concentrado
Amp - Ampicilina
ATP - Adenosina Trifosfato
bisacrilamida - N,N' - metileno bisacrilamida
BrEt - Brometo de Etídeo
BSA - Albumina sérica bovina
CIP - fosfatase de intestino de vitela
Da - Daltons
DTT - ditioneitol
EDTA - Ácido Etileno-Diamino-Tetracético
His tag - seqüência de 6 resíduos de Histidina
IPTG - isopropil- β -D-galactosídeo
Kb - quilo bases - 1.000 pares de bases
Kda - quilo dáltons - 1.000 dáltons
mA - miliampere
mM - milimolar
ng - nanograma
Ni-NTA - *nickel-nitrilotriacetic acid*
p/v - peso/volume
pb - pares de bases
pH - potencial hidrogeniônico
RNase - ribonuclease
SAP - *shrimp alkaline phosphatase*
SDS - dodecilsulfato de sódio
SDS-PAGE - *sodium dodecil sulphate - polyacrylamide gel electrophoresis*
TCA - ácido tricloroacético
TEMED - N, N, N', N' - tetrametiletlenodiamina
Tris-base - Tris(hidroximetil)aminometano
X-gal - 5-bromo-4-cloro-3-indolil- β -D-galactosídeo
SELEX - *systematic evolution of ligands by exponential enrichment*
EMSA - *electrophoresis mobility assay*
rpm - rotações por minuto
min - minuto
nm - nanômetro
 μ g - micrograma
 μ l - microlitro
V - Volts
 μ F - microfaraday
Å - angstrom

RESUMO

Proteínas *zinc finger* são compostas por domínios compactados contendo α -hélices e folhas β unidos e estabilizados por um ou dois átomos de zinco. Arranjos repetidos de *zinc fingers* são comumente utilizados para reconhecimento de ácidos nucleicos. Dentre outras atividades, eles estão envolvidos em replicação, transcrição e reparo de DNA. No nucleocapsídeo do vírus HIV tipo 1 foi identificada uma proteína contendo o motivo *zinc finger* CX₂CX₄HX₄C, estando esta proteína envolvida em várias etapas do ciclo de vida viral, incluindo a propriedade de encapsidação do RNA viral. Em tripanosomatídeos, somente poucas proteínas contendo o motivo *zinc finger* já foram identificadas até o presente momento.

Em um fragmento genômico de 17 kb da banda XX de *T. cruzi*, nós identificamos três genes *in tandem* codificando para proteínas zinc finger do tipo CX₂CX₂HX₄C. Nós também demonstramos que genes similares estão presentes em *T. brucei* e *L. major* como três definidos grupos monofiléticos entre esses tripanosomatídeos. Em *T. cruzi*, *TcZFP8* corresponde a um novo gene codificando para uma proteína com oito motivos zinc finger. O homólogo de *TcZFP8* em *T. brucei* é aparentemente ausente, enquanto um candidato foi identificado em *L. major*. A clonagem molecular e a expressão heteróloga de *TcZFP8* foi realizada para produção de anticorpos e procedimentos como imunocitocalização, SELEX e EMSA. Análise por *Western blotting* revelou a presença dessa proteína nas três formas do parasita. Análises usando extratos protéicos nucleares e citoplasmáticos de *T. cruzi* mostraram que essa proteína está presente na porção nuclear. Esse resultado foi confirmado através de análise de microscopia por imunofluorescência indireta. Experimento de SELEX demonstrou quatro diferentes populações com uma região interna rica em C e/ou G, porém sem seqüências consenso específicas. Análises preliminares de EMSA de uma das quatro populações selecionadas revelaram evidências de que *TcZFP8* possa ter afinidade de ligação à fita simples de DNA.

ABSTRACT

Zinc fingers are compact protein domains composed of a α -helix and a β -sheet held together by a zinc ion. Tandem arrays of zinc fingers are commonly used to recognize nucleic acids. Among other activities, they are involved in the processes of replication, transcription, and DNA repair. The nucleocapsid protein of HIV-1 contains a zinc finger motif CX₂CX₄HX₄C that contributes to multiple steps of the viral life cycle, including the proper encapsidation of HIV RNA. In trypanosomatids, only a few of the proteins that contain such fingers were identified.

In a 17-kb genomic fragment of *Trypanosoma cruzi* chromosome XX we identified three tandemly linked genes coding for CX₂CX₄HX₄C zinc finger proteins. We also showed that similar genes are present in *Trypanosoma brucei* and *Leishmania major* sharing three monophyletic groups among these trypanosomatids. In *T. cruzi*, *TcZFP8* corresponds to a novel gene coding for a protein containing eight zinc finger motifs. Homologous of *TcZFP8* in *T. brucei* is apparently absent, while one candidate in *L. major* was identified. Molecular cloning of gene *TcZFP8* and heterologous expression were performed in *Escherichia coli*. The purified recombinant protein His₆x-TcZFP8 was used to produce antibody in rabbits and GST-TcZFP8 in SELEX and EMSA procedures. Using Western blot analysis, we observed the presence of this protein in all three forms of the parasite: amastigote, trypomastigote and epimastigote. Analysis using cytoplasm and nuclear cell extracts showed that this protein is present in the nuclear extracts and indirect immunofluorescence microscopy analysis confirmed the nuclear localization of the TcZFP8. SELEX experiment showed four different populations rich in C and/or G nucleotides, but with none consensus sequence. Preliminary EMSA from one population gave evidence that TcZFP8 has affinity to bind to single-stranded DNA.

INTRODUÇÃO

I - Estado da arte

1 - *Trypanosoma cruzi* e Doença de Chagas

Trypanosoma cruzi, o agente causador da doença de Chagas em humanos, foi primeiramente descrito por Carlos Chagas em 1909. Este protozoário é flagelado e pertencente à ordem Kinetoplastida, família Trypanosomatidae e gênero *Trypanosoma* apresentando importantes variações morfológicas e diferentes padrões de infectividade e patogenicidade.

Chagas (1909) identificou inicialmente o *T. cruzi* em insetos da espécie *Panstrongylus megistus*, coletados em Lassance, estado de Minas Gerais, Brasil. Outras espécies também atuam como vetores da transmissão da doença de Chagas como *Triatoma infectans* e *Triatoma Sordida* bem como outras espécies (não somente a *Homo sapiens*) podem apresentar *T. cruzi* e outros tripanossomatídeos em diferentes estágios morfológicos (De Sousa, 1999). O *Trypanosoma cruzi* apresenta ao longo do seu ciclo de vida três formas básicas (epimastigota, tripomastigota e amastigota) classificadas em função de sua morfologia, da maneira como o flagelo emerge do corpo celular e da posição do núcleo em relação ao cinetoplasto (mitocôndria modificada). Esses diferentes estágios celulares são determinados pela passagem do parasita pelos hospedeiros invertebrado e vertebrado (Brenner, 1965).

O ciclo de vida do *T. cruzi* consiste das etapas descritas a seguir. Ao picar o homem, o triatomíneo infectado defeca próximo à picada e, ao coçar a região irritada, o indivíduo leva formas tripomastigotas metacíclicas, contidas nas fezes do vetor, para o interior da pele. Estes tripomastigotas caem na circulação sanguínea e, entrando em contato com as células do sistema mononuclear fagocitário (SMF), interagem com lisossomos. Após o parasita escapar para o citosol, este se diferencia na forma amastigota (intracelular). A partir desse momento, ocorrerão sucessivas divisões (em torno de nove), tornando a célula repleta de parasitas. Depois, ocorre nova diferenciação (amastigota→tripomastigota sanguíneo), liberando flagelados na circulação, podendo assim, serem seguidos dois caminhos: (i) os tripomastigotas invadem novas células do hospedeiro

(SMF, células musculares: lisas, esqueléticas, cardíacas, nervosas, etc.), reiniciando o ciclo neste hospedeiro, ou (ii) o sangue contaminado é sugado pelo triatomíneo, infectando este inseto com tripomastigotas sanguíneos, que posteriormente se converterão em epimastigotas, dividir-se-ão e transformar-se-ão em tripomastigotas metacíclicas, readquirindo a capacidade infectiva (de Souza, 1984) (Figura 1).

2 - Biologia Molecular dos Tripanossomatídeos

Os tripanossomatídeos possuem alguns aspectos nucleares distintos, que os diferem dos eucariotos superiores. Seus cromossomas não condensam durante a divisão celular (Vickerman & Preston, 1970) e a compactação da cromatina no núcleo é menos pronunciada quando comparada com a cromatina nuclear de coração de ratos (Hecker & Gander, 1985). Suas histonas diferem significativamente dos outros eucariotos em relação à composição e à seqüência de aminoácidos (Hecker & Gander, 1989; Burri *et al.*, 1993). A análise citogenética dos cromossomas de tripanossomatídeos é dificultada pelo fato dos mesmos não condensarem durante a divisão celular. Alternativamente, o desenvolvimento da técnica de eletroforese em campo pulsante (PFGE) tem permitido a separação e visualização dos cromossomas desses parasitas (Wagner & So, 1990). Essa técnica tem sido utilizada para tentar se determinar o cariótipo molecular de alguns tripanossomatídeos, dentre os quais, o da cepa CL Brener de *T. cruzi*, onde já foi estimado o tamanho do seu genoma em 87 Mpb (Degrave *et al.*, 1997).

Apesar de não estar bem definido o número de cromossomas de *T. cruzi* foram identificadas 20 bandas cromossomais distintas, variando entre 0,45 Mb a 3,5 Mb sendo numeradas de 1 a 20 em géis de baixa resolução (Henriksson *et al.*, 1996). Os tripanossomatídeos apresentam, além do DNA nuclear, um DNA mitocondrial (kDNA) que é distinto em sua estrutura e em seu modo de replicação.

O kDNA consiste em uma rede de milhares de DNAs circulares topologicamente interligados formando uma estrutura condensada em forma de disco denominado cinetoplasto. Cada célula contém um cinetoplasto localizado na matriz de sua única mitocôndria, situada adjacente ao corpo basal do flagelo. A rede contém dois tipos de moléculas de DNA: os maxi-círculos presentes em algumas cópias (20 a 50) e os mini-círculos representados por milhares de cópias (5.000-10.000). Esse kDNA corresponde a 30% do DNA celular (Shapiro & Englund, 1995; de Souza, 1999).

Os mini-círculos têm tamanho que variam de 0,5 a 2,9 Kb e codificam pequenos RNAs denominado de RNAs guia (gRNA) que estão envolvidos no processo de editoração (Shapiro & Englund, 1995) Neste processo ocorrem deleções e adições de uridinas em alguns RNAs precursores mitocondriais mediado pelos gRNAs, estabelecendo a fase de leitura dos mesmos produzindo, então, mRNAs funcionais (Hajduk & Sabatini, 1996).

Os maxi-círculos são genética e estruturalmente similares ao DNA mitocondrial de outros eucariotos e em diferentes espécies apresentam tamanho de 20 a 40 kb. Eles codificam RNAs ribossomais, gRNAs e algumas proteínas mitocondriais envolvidas no transporte de elétrons e na fosforilação oxidativa (Simpson *et al.*, 1987; Shapiro & Englund, 1995).

Pays (1993) estimou que o genoma dos tripanossomatídeos estivesse entre 10^7 - 10^8 pb variando de uma espécie a outra, sendo cerca de 10 vezes maior do que o genoma de *Escherichia coli* e 2-3 vezes maior que o genoma da levedura.

Recentemente foi realizado o sequenciamento do genoma do protozoário *Trypanosoma cruzi* revelando ser o seu genoma diplóide, contendo 22.570 proteínas codificadas por genes dos quais 12.570 são pares de alelos. Cerca de 50% do genoma deste parasita consiste de seqüências repetidas, tais como, retrotransposons e genes codificando para uma grande família de moléculas de superfícies, incluindo trans-sialidases, mucinas, GP63 e uma nova família (>1.300 cópias) de genes que codificam para uma proteína de superfície associada à mucina (MASP) (El-Sayed *et al.*, 2005).

Berriman *et al.* (2005) estimaram o tamanho do genoma do *Trypanosoma brucei* em 26 Mb contendo 9.068 genes preditos, incluindo aproximadamente 900 pseudogenes e aproximadamente 1700 genes específicos do *T. brucei*. Arranjos subteloméricos cromossomais revelaram um arquivo de 806 glicoproteínas de superfícies variantes (VSG). Essas VSG são usadas pelo parasita para invadir o sistema imune do hospedeiro mamífero. Muitas VSG são pseudogenes, cujo caminho pode ser usado para gerar a expressão de genes mosaicos por recombinação ectópica.

O tamanho do genoma haplóide de *Leishmania major* foi estimado 32-38 Mb onde 36 cromossomas foram seqüenciados predizendo 911 RNA genes, 39 pseudogenes e 8.272 genes codificando proteínas, das quais 36% acredita-se já poder ser elucidada suas supostas funções. Esses genes incluem genes envolvidos em interação patógeno-hospedeiro, tais como enzimas proteolíticas e uma extensiva maquinaria para síntese de complexo de glicoconjugados de superfície (Ivens *et al.*, 2005).

Os mecanismos de transcrição e de tradução em tripanossomatídeos são pouco conhecidos, mas ocorrem mecanismos moleculares distintos aos demais eucariotos, como por exemplo, *trans splicing* e transcrição policistrônica. Alguns genes cromossomais estão dispostos *in tandem* e parecem ser transcritos como um RNA policistrônico. Além disto, acreditava-se que introns não estavam presentes no genoma dos tripanossomatídeos, o que descartava a ocorrência do processo de *cis splicing*. Foi demonstrado entretanto que seqüências de genes codificando uma poli(A)polimerase (PAP) em *T. cruzi* e *T. brucei* sofrem tanto *trans splicing* quanto *cis splicing* (Mair *et al.*, 2000).

Os RNAs mensageiros dos tripanossomatídeos apresentam uma seqüência líder (SL) adicionada pós transcricionalmente no seu terminal 5' no processo denominado de *trans splicing*. Os substratos para este processo são um pequeno RNA não poliadenilado, o medRNA, de tamanho variando de 100 a 140 nucleotídeos, que contém um exon na região 5' contendo a seqüência líder (SL) de tamanho em torno de 39 nucleotídeos, e o pré-mRNA que apresenta o sítio acceptor da SL na região 5' não traduzida sendo, então, gerados os RNAs mensageiros (Nilsen, 1994).

Na transcrição policistrônica, os RNAs mensageiros individuais parecem ser produzidos a partir de um único promotor, como um RNA precursor que é processado tanto por *trans splicing* quanto por poliadenilação, gerando transcritos independentes, ou seja, mRNA maduros. Os tripanossomatídeos não apresentam um sinal de poliadenilação conservado como em outros eucariotos e este parece ocorrer a uma distância conservada do sítio de *trans splicing* posterior em transcritos policistrônicos (Vanhamme & Pays, 1995).

Inúmeras proteínas que supostamente se ligam a RNA foram também identificadas nos genomas seqüenciados do *T. cruzi*, *T. brucei* e *L. major*, enfatizando-se, portanto, a ativa regulação da expressão gênica destes parasitas a nível pós-transcricional (Ivens *et al.*, 2005; El-Sayed *et al.*, 2005).

Diferentemente dos demais eucariotos, os tripanossomatídeos possuem genes codificando proteínas transcritos pela RNA polimerase I como, por exemplo, em *T. brucei*, os genes de VSGs (*variant surface glycoprotein*) e da PARP (*procyclic acidic repetitive protein*) (Kooter & Borst, 1984).

3 - Proteínas de Ligação a Ácidos Nucléicos

A expressão gênica em resposta a uma ampla variedade de estímulos extracelulares e/ ou intracelulares pode ser regulada nos níveis transcricional, pós-transcricional, traducional e pós-traducional. Assim, vários fatores protéicos controlam esses níveis, sendo pontos chaves de regulação da expressão gênica e dentre eles a iniciação, a elongação e a terminação da transcrição; *cis/trans splicing* e poliadenilação; estabilidade do pré-mRNA e do mRNA maduro e do e transporte para o citoplasma; tradução; modificações pós-traducionais e estabilidade da proteína (Graham, 1995).

Deste modo, o controle da expressão gênica dá-se em vários níveis, dentre os quais o transcricional. Neste nível algumas proteínas apresentam motivos específicos de ligação a ácidos nucléicos e têm sido descritas como fatores de transcrição, atuando como peças chaves moleculares

desse processo e conseqüentemente controlando vários eventos celulares, dentre eles, o crescimento, a diferenciação, a proliferação e apoptose.

O estudo de proteínas regulatórias reconhecadora de seqüências específicas de DNA revela em sua maioria que o contato ocorre no sulco maior do DNA, envolvendo ligações com hidrogênio, ligações iônicas e interações hidrofóbicas. Estudos de cristalografia por difração de raios X e por ressonância magnética nuclear (RMN) têm revelado que essas proteínas contêm domínios contendo motivos de ligação ao DNA. Essas são formadas tanto por α -hélice quanto folhas β -pregueadas onde se ligam ao sulco maior (Harrison, 1991; Pabo & Sauer, 1992).

Essas famílias de proteínas têm sido identificadas e classificadas em diferentes estruturas familiares como: homeodomínio, *zinc finger*, *leucine zipper* e *helix-turn-helix*. Os membros das famílias homeodomínio e *zinc finger* são os mais identificados como fatores de transcrição desde insetos até humanos. Mutações ou expressões aberrantes de genes codificando essas proteínas resultam em transformações neoplásicas em tipos diferentes de células, dentre as quais células pancreáticas. Dessa maneira essas proteínas são fortes candidatas ao papel de guardiãs do crescimento e da diferenciação normal dessas células (Urrutia, 1997).

3.1 – Motivo HTH e Homeodomínio

O *helix-turn-helix* (HTH) foi o primeiro motivo protéico contendo domínios de ligação ao DNA a ser reconhecido, tendo sido identificado amplamente em proteínas de ligação a DNA em procariotos. As primeiras estruturas cristalográficas de proteínas de ligação a DNA descritas foram a proteína Cro do fago λ , a proteína CAP de *E. coli* e o domínio de ligação ao DNA do repressor λ . Comparações dessas três proteínas revelaram um motivo conservado de reconhecimento consistindo em uma α -hélice, uma volta e uma segunda α -hélice (Alberts *et al.*, 1997).

O homeodomínio é um motivo de ligação a DNA presente em famílias de proteínas regulatórias em eucariotos. Comparações de seqüências de aminoácidos sugerem que o homeodomínio contém o motivo HTH descoberto primeiramente em *Drosophila* (os genes

homeóticos seletores conhecidos por *Antennapedia* e *Engrailed*) e subsequentemente em esponjas, plantas, fungos e humanos. O homeodomínio permite o reconhecimento de um grupo de genes alvos por proteínas homeodomínio, que primariamente regulam genes e atuam como fator de transcrição regulando genes alvos em um padrão espacial e temporal preciso (Pabo & Sauer, 1992; Gehring *et al.*, 1994)

3.2-Motivo *Leucine zipper*

O *leucine zipper* tem um papel importante na diferenciação e no desenvolvimento. Esse motivo dimerizado reconhece o DNA e geralmente a porção da proteína responsável pela dimerização é diferente da responsável pela ligação ao DNA, apresentando, portanto, dois subdomínios distintos. Estudos bioquímicos revelam que as duas α -hélices, uma de cada monômero, permanecem juntas formando uma pequena estrutura com leucinas. A estrutura completa desse motivo *leucine zipper* contém domínios conhecidos como bCC, também chamados de bZip (*basic leucine zipper*), estando presentes em proteínas de leveduras (GCN4), em proteínas de eucariotos superiores (como a *cAMP response element binding* ou CREB), Fos e Jun (produtos de oncogenes) (Harrison, 1991; Pabo & Sauer, 1992).

O bZip está envolvido na regulação da transcrição do DNA. Ele juntamente com outros fatores reconhece regiões promotoras e *enhancers* contribuindo para que a RNA polimerase se ligue eficientemente ao DNA e inicie a transcrição (Cranz *et al.*, 2004).

3.3 – Motivo *Zinc finger*

Zinc finger é um dos motivos estruturais mais abundantes envolvidos na interação ácidos nucléicos e proteínas. As proteínas *zinc finger* estão envolvidas em vários aspectos da regulação de genes eucarióticos. A primeira proteína contendo motivo *zinc finger* descrita foi o fator de transcrição TFIIIA de *Xenopus*. Homólogos de *zinc finger* ocorrem em proteínas que induzem diferenciação e crescimento, em proto-oncogenes, em fatores gerais de transcrição, em genes que

regulam o desenvolvimento em *Drosophila* e em genes regulatórios de organismos eucarióticos inferiores. Todos os eucariotos até então estudados, desde leveduras até humanos, possuem proteínas contendo o domínio *zinc finger*, onde além dos aminoácidos há também um ou mais átomos de zinco como um de seus componentes estruturais (Pavletich & Pabo, 1991; Pabo & Sauer, 1992).

Estudos de proteínas contendo motivo *zinc finger* revelam o papel do átomo de zinco na estabilidade das mesmas. Um exemplo é a proteína δ_3 identificada em reovírus, que na ausência do íon, a proteína apresenta-se com conformação alterada e estabilidade reduzida, tornando-se alvo de degradação por proteases celulares (Mabrouk & Lemay, 1994). Esses motivos são formados por um átomo de zinco coordenado por cisteínas e histidinas.

Estudos têm revelado a existência de vários grupos de proteínas contendo motivo *zinc finger*, dentre os quais: $Zn(C_2H_2)$; $[Zn(C_4)]_2$; $Zn_2(C_6)$ e $Zn(C_4)$; $Zn_2(C_3HC_4)$ – *ring finger* motif; $Zn(C_3H)$; $Zn(CCHC)$.

3.3.1- Motivo $Zn(C_2H_2)$

Este motivo foi primeiramente descrito em estudos de transcrição de genes ribossomais 5S em *Xenopus*. Uma seqüência essencial para a transcrição de genes 5S pela RNA polimerase III foi identificada pela deleção de partes do gene, provando ser crítica para a transcrição eficiente. Essa região interna de controle contendo 45 nucleotídeos é reconhecida por uma proteína de 40 kDa chamada TFIIIA. Uma molécula de TFIIIA liga-se ao gene de RNA 5S formando um complexo que se liga seqüencialmente a TFIIBC, TFIIB e RNA polimerase III. Os fatores C e B também iniciam a transcrição dos genes tRNA, enquanto TFIIIA é específico para os genes de RNA 5S (Miller *et al.*, 1985; Romaniuk, 1985; Pavletich & Pabo, 1991). O TFIIIA contém uma folha β e uma α -hélice (Figuras 2 e 3). Duas cisteínas estão na região folha β e duas histidinas na α -hélice. O átomo de zinco coordenado pelas cadeias laterais desses aminoácidos estabiliza essa estrutura secundária formando um domínio globular compactado. A Zif268 também é uma proteína que participa da

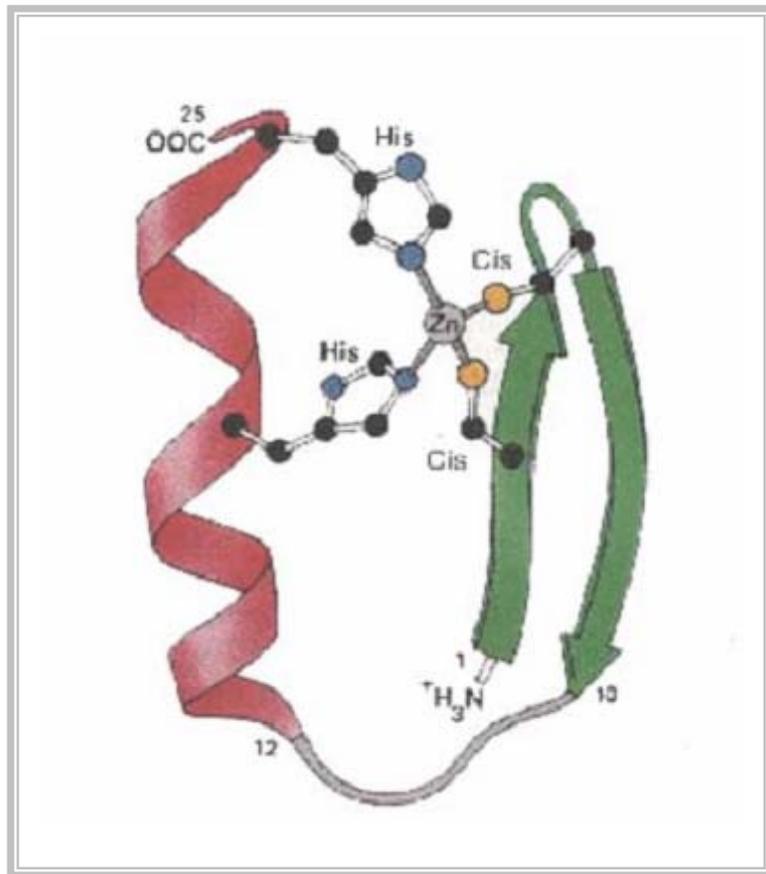


Figura 2 - Esquema de um motivo *zinc finger*. Um grampo β antiparalelo (verde) e a alça (cinza) são seguidos de uma hélice (vermelho) que tem um papel importante no reconhecimento do DNA. Um íon zinco liga-se a duas histidinas da α -hélice e duas cisteínas da folha β (Alberts *et al.*, 1997).

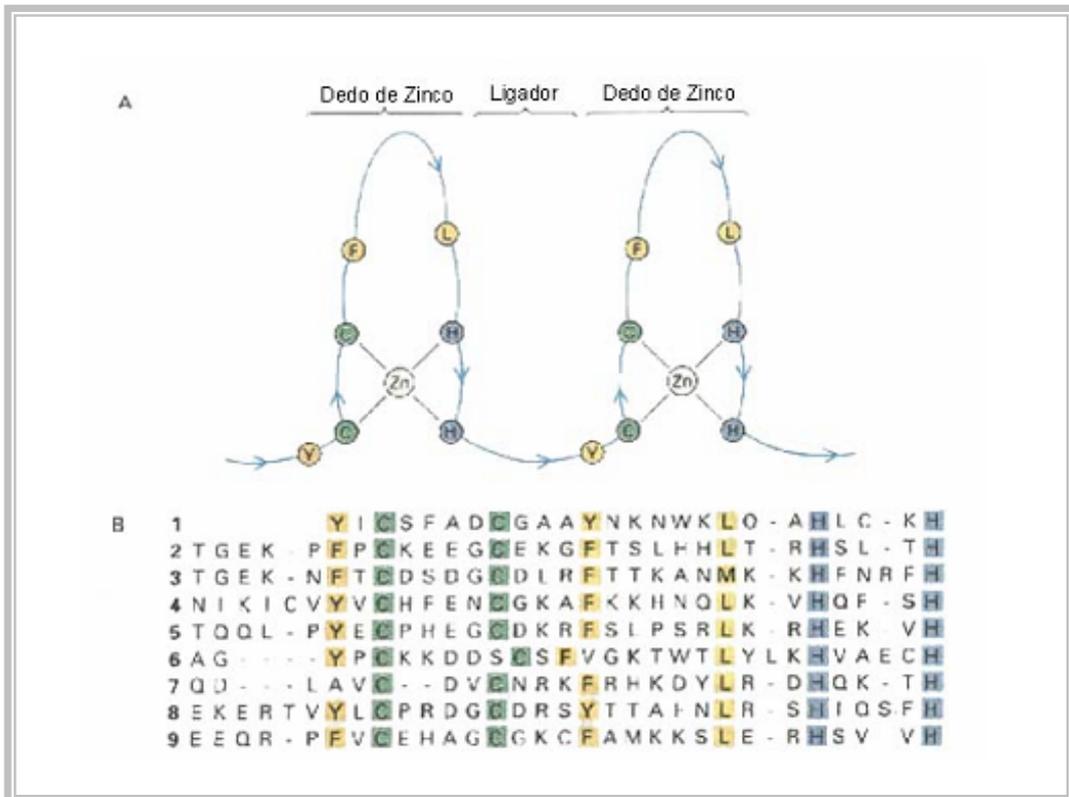


Figura 3 – A *Zinc finger* do fator de transcrição TFIIIA de *Xenopus*. (A) Esquema mostrando um par de *zinc fingers* unidos por uma alça. (B) A seqüência de aminoácidos do fator de transcrição TFIIIA de *Xenopus* mostrando nove seqüências em *zinc finger*. O zinco é ligado a um par de cisteínas (verde) e um par de histidinas (azul). Três aminoácidos hidrófobos (amarelo) em cada repetição também são conservados (Alberts *et al.*, 1997).

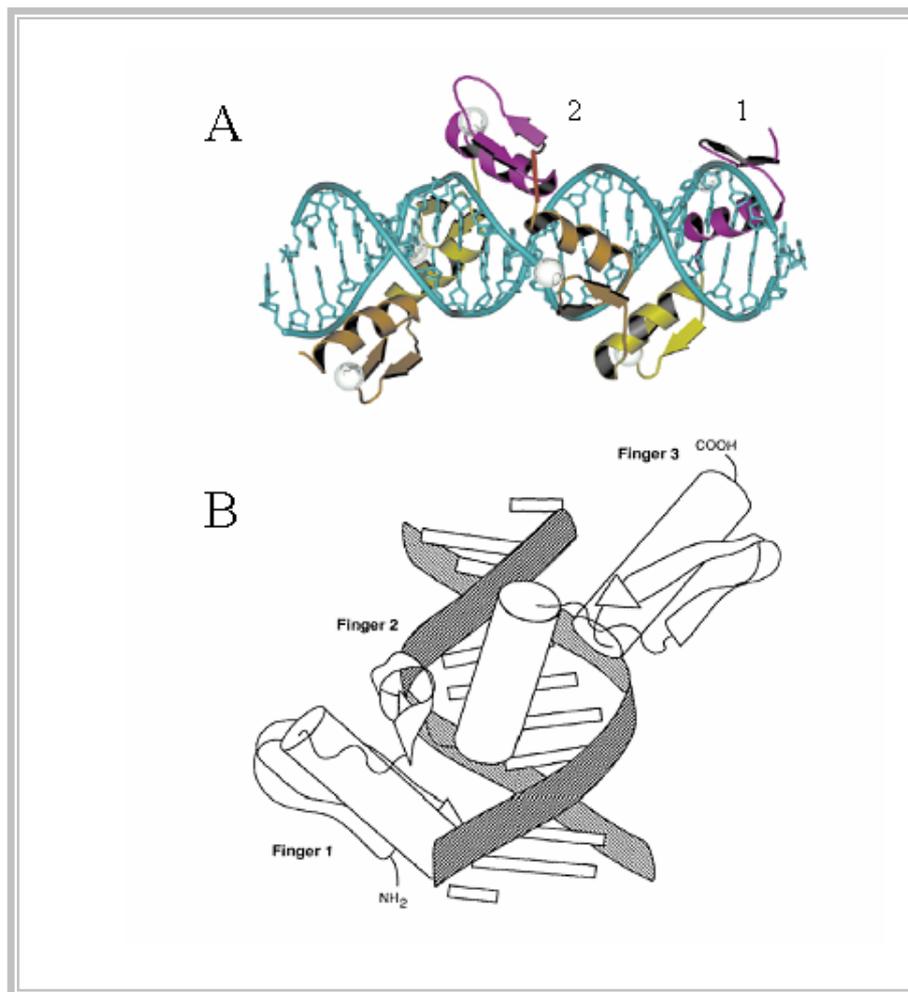


Figura 4 – A proteína Zif 268.(A) -Modelo de duas moléculas zif 268 (1 e 2) ligando *in tandem* a sítios específicos no DNA. As *zinc fingers* de 1 a 3 estão ligadas da direita para a esquerda e coloridas de lilás, amarelo e alaranjado respectivamente. A linha em vermelho indica um *gap* para que o *link* possa ser desenhado. (B) -Esquema mostrando o complexo zif268 – DNA em arranjo. As α -helices são mostradas como cilindros; as folhas β são mostradas em setas; o átomo de zinco foi omitido neste esquema. Os três *zinc fingers* fazem contato com o sulco maior do DNA. Cada *zinc finger* faz contato com três pares de bases no DNA (Pabo & Sauer, 1992; Peisach e Pabo, 2003).

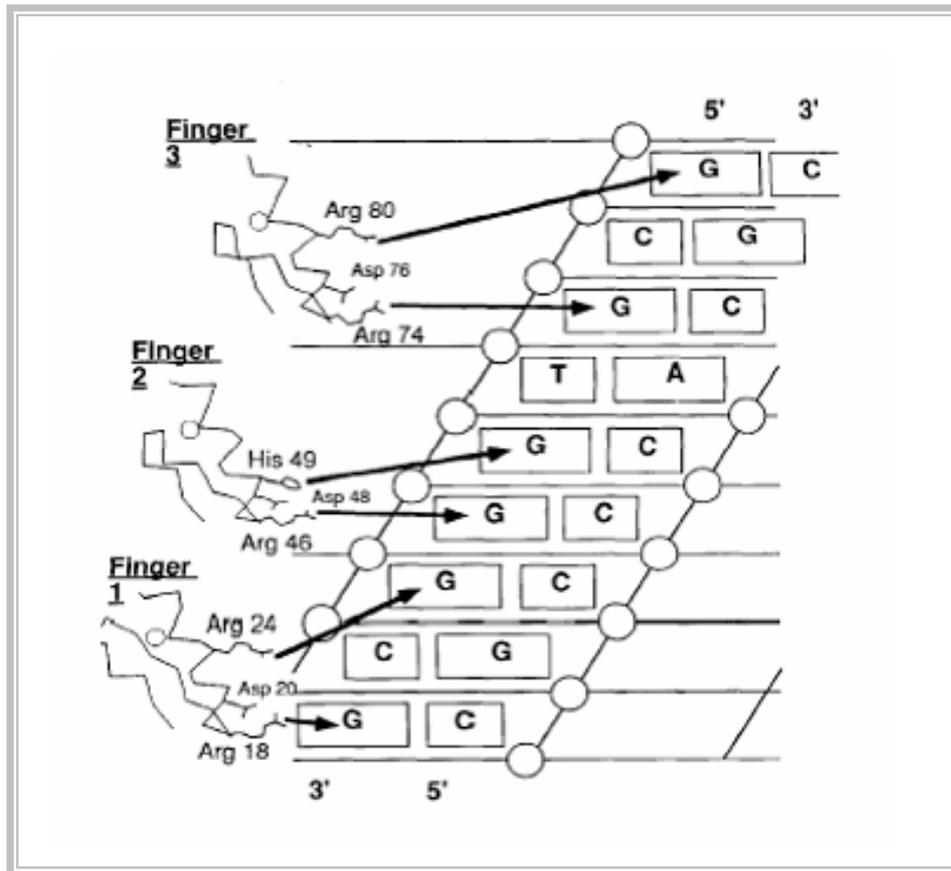


Figura 5 – Esquema resumindo a base do contato da *Zinc finger* Zif 268 com o DNA. O DNA é representado como uma projeção cilíndrica, as setas indicam o contato com o sulco maior. O contato envolve uma região rica em guaninas no sítio de ligação (Pabo & Sauer, 1992).

transcrição de genes do início do desenvolvimento de camundongos que apresenta três *zinc finger* $\text{Cis}_2\text{-His}_2$ cada qual fazendo contato com três pares de bases do DNA (Figura 4). O local no DNA onde se estabelece a interação é rico em aminoácidos carregados positiva e negativamente (da proteína) e guaninas (do ácido nucléico) (Figura 5) (Pavletich & Pabo, 1991; Pabo & Sauer, 1992 e Peisach & Pabo, 2003).

Em *Trypanosoma cruzi* foi identificado um elemento retrotransposon non-LTR (*long terminal repeat*) presente em alto número de cópias, contendo várias fases abertas de leitura, dentre as quais: uma denominada ORF 1 homóloga a uma família de enzimas de reparo de DNA (APE), uma ORF 2 homóloga à enzima transcriptase reversa presente em retrovírus e uma ORF 3 contendo o motivo *zinc finger* $\text{Cis}_2\text{-His}_2$ (Martin *et al.*, 1995).

Dezesseis proteínas relacionadas ao processo de editoração já foram identificadas em tripanossomatídeos. As mesmas fazem parte de um complexo protéico denominado editossoma (Worthey *et al.*, 2003). Em *Leishmania chagasi* o editossoma é denominado complexo L sendo também composto por dezesseis proteínas estabilizadas através de interações umas com as outras. Três proteínas pertencentes ao complexo L contêm o motivo *zinc finger* CCHH e são chamadas LC-1, LC-4 (correspondente a TbMP63 de *T. brucei*) e LC-7b. Experimentos com a proteína LC-4 modificada por técnicas de mutação em seus dois motivos (ZnF-1 e ZnF-2), substituindo os resíduos de cisteínas por resíduos de glicinas resultam em um crescimento defeituoso (mutação no motivo ZnF-1) e na quebra parcial do complexo-L (tanto para ZnF-1 quanto ZnF-2). Através de experimentos de crosslinking e co-imunoprecipitação foi demonstrada a interação de LC-4 com uma REL1 (RNA ligase) e LC-3. Esses resultados sugerem que os dois motivos CCHH presentes na proteína LC-4 desempenham um papel na interação proteína-proteína e indicam que o sub-complexo LC-4 pode ser necessária para formação ou estabilização do complexo L (Kang *et al.*, 2004).

Em *Trypanosoma brucei* foi identificada uma proteína denominada BandIII contendo dois motivos CCHH envolvidos no processo de editoração. A depleção desta proteína leva a uma deficiência na atividade endonucleásica dirigida pelo gRNA em clivar ambos os sítios de

deleção e inserção de uracilas (Huang *et al.*, 2002). Outras quatro proteínas de *T. brucei* contendo este motivo CCHH já foram identificadas: TbMP81, TbMP63, TbMP42 e TbMP18. Anticorpos monoclonais específicos para TbMP63 e TbMP42 imunoprecipitam *in vitro* componentes do complexo de editoração de RNA de extratos protéicos de *T. brucei*, sugerindo o papel destas proteínas neste mecanismo (Drozd *et al.*, 2001; Panigrahi *et al.*, 2001). Experimentos recentes com silenciamento do gene que codifica para TbMP42 de *T. brucei* sugeriram o papel dessa proteína com atividades endo e exoribonuclease para o ciclo de reação da editoração (Brecht *et al.*, 2005).

3.3.2- Motivo [Zn(C₄)]₂

Esse motivo faz parte da estrutura dos domínios dos receptores nucleares. Os receptores nucleares pertencem a uma das classes mais abundantes de reguladores transcricionais em metazoários. Eles regulam diversas funções, como homeostase, reprodução, desenvolvimento e metabolismo. Um grande número de receptores foi identificado através da análise de seqüência de similaridade com receptores conhecidos, porém, ainda não se conhece o ligante natural de alguns deles, sendo então, estes denominados receptores órfãos (Robinson-Rechavi *et al.*, 2003). Dentre os receptores dos quais se conhecem os ligantes naturais, inclui-se a família dos receptores da tireóide, hormônios esteróides, retinóides e vitamina D. Os receptores nucleares hormonais atuam como fatores de transcrição por interagirem diretamente como monômeros homodímeros ou heterodímeros a elementos responsivos presentes no DNA de seus genes alvo. Além de agirem em suas próprias vias de sinalização, muitos receptores fazem *cross-talking* com outras vias de sinalização, muitas vezes coordenadas por receptores de membrana. O efeito dos receptores nucleares na transcrição é mediado pelo recrutamento de co-reguladores. Um grupo de receptores se liga a co-repressores reprimindo a expressão de um gene alvo na ausência de seu ligante. Co-repressores são encontrados como complexos formados por multicomponentes contendo atividade histona deacetilase. A deacetilação induz a compactação da cromatina e a repressão transcricional.

Após a interação com o ligante, os receptores sofrem uma mudança conformacional que permite o recrutamento de um complexo múltiplo co-ativador. Muitas dessas proteínas são fatores de remodelamento de cromatina ou possuem atividade histona acetilase, além de interagir diretamente com a maquinaria transcricional básica, auxiliando no recrutamento dos fatores gerais de transcrição. As atividades de acetilação de histonas e remodelamento da cromatina causam descompactação da cromatina e promovem a ativação transcricional (Pascual & Aranda, 2001).

3.3.2.1. Domínio Estrutural e Função dos Receptores Nucleares

Os receptores nucleares possuem uma organização estrutural em comum: uma região N-terminal variável (A/B), um domínio C de ligação ao DNA (DBD), uma região de conexão D e uma região E conservada, contendo o domínio de ligação ao ligante (LBD). Alguns receptores também possuem uma região F (Figura 6). A região N-terminal (domínio A/B) é altamente variável e contém uma região ativa de trans-ativação (AF-1); os domínios A/B possuem tamanhos variados entre 50 e 500 aminoácidos, e sua estrutura tridimensional não é conhecida. A região C é a mais conservada e contém o P-box e o B-box, que são pequenos motivos envolvidos na dimerização dos receptores nucleares e sendo também responsáveis pela ligação ao DNA, especificamente em seqüências contendo tipicamente o consenso AGGTCA. Esta dimerização inclui homodímeros e heterodímeros. A estrutura tridimensional do domínio de ligação ao DNA (DBD) de um grande número de receptores nucleares já foi resolvida por cristalografia e a mesma contém dois motivos *zinc finger* altamente conservados: $CX_2CX_{13}CX_2C$ e $CX_5CX_9CX_2C$ onde cada quatro cisteínas são coordenadas por um átomo de zinco. A Figura 7 mostra um diagrama contendo o DBD dos receptores nucleares. Entre o domínio de ligação ao DNA e o domínio de ligação ao ligante, existe uma região conservada denominada D, que se comporta com uma alça flexível entre os domínios C e E, e contém um sinal de localização nuclear (NLS), que faz uma sobreposição com o domínio C (Figura 6). Um domínio denominado E, compreende o domínio de ligação ao ligante (LBD), cuja estrutura secundária compreende, em geral, 12 α -hélices bem conservadas em sua seqüência primária (Figura 8).

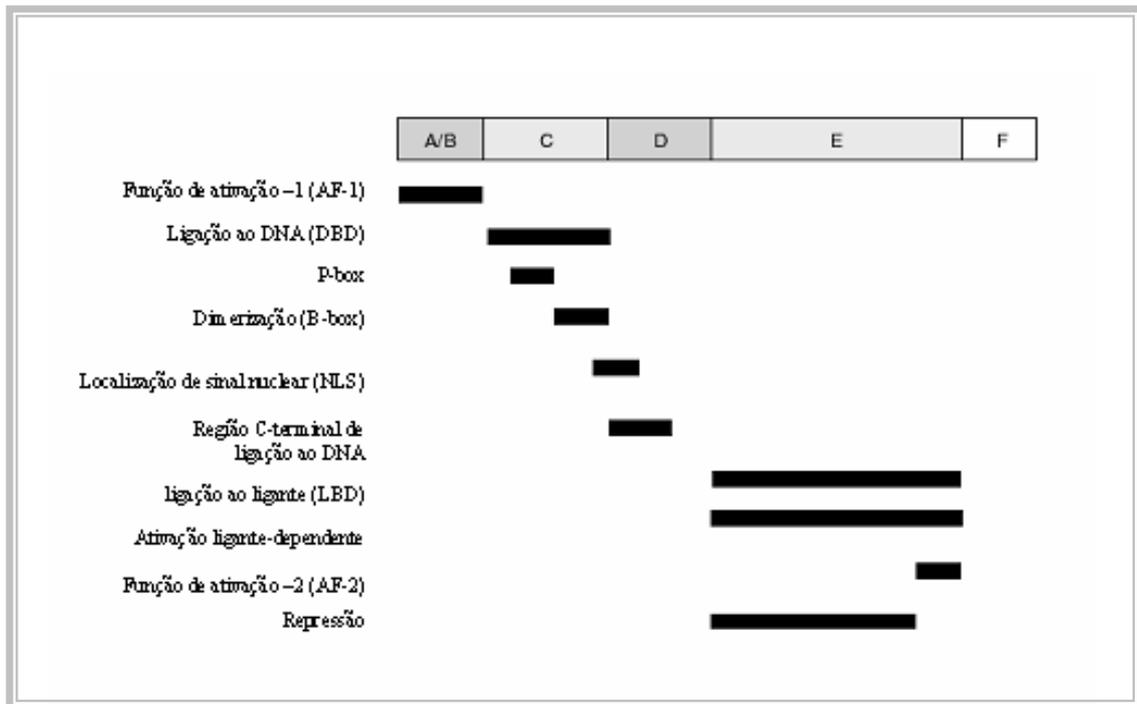


Figura 6 – Estrutura dos Domínios dos Receptores Nucleares. As barras em preto indicam a região de cada domínio dentro da estrutura completa dos receptores nucleares esquematizada em retângulos de A- F (Lazar *et al.*, 2002).

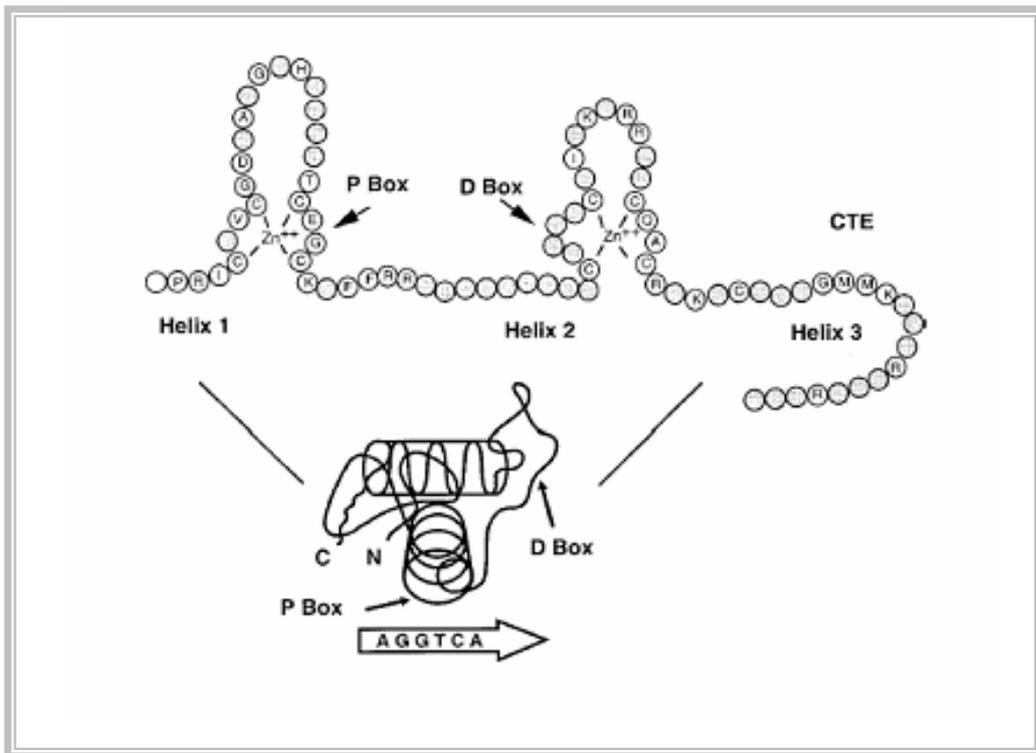


Figura 7 - Domínio de Ligação ao DNA (DBD) dos receptores nucleares. Diagrama contendo dois *zinc fingers* e uma extensão C-terminal (CTE). Nas *zinc fingers*, quatro resíduos conservados do aminoácido cisteína são coordenados por um íon zinco. Outros resíduos conservados de aminoácidos são mostrados e designados com a letra do aminoácido correspondente. A hélice 1 contém os resíduos do P-Box envolvidos na discriminação de um elemento responsivo. Resíduos de aminoácidos de uma segunda *zinc finger* são marcados como D-Box formando a interface de dimerização (modificado de Aranda & Pascual, 2001).

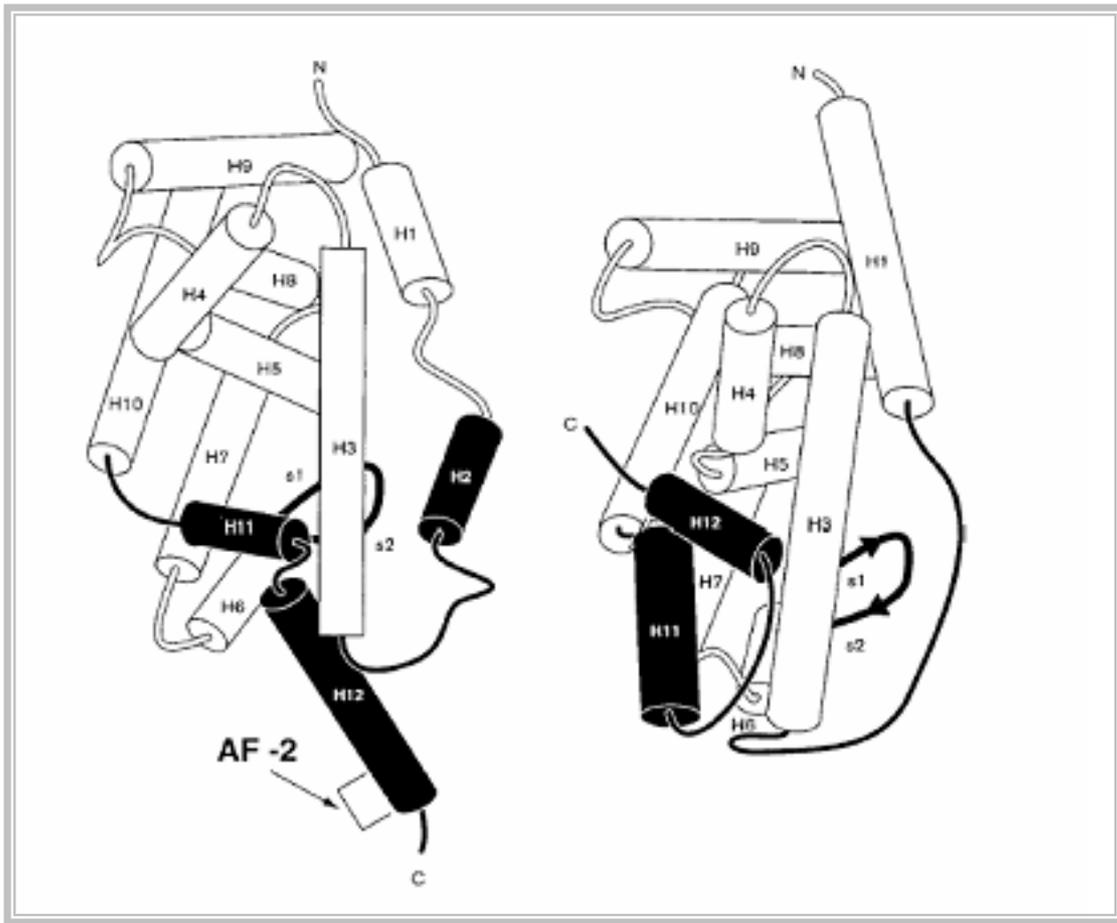


Figura 8 – Domínio de Ligação ao Ligante (LBD) dos receptores nucleares. À esquerda, a estrutura em cristal do LBD do RXR α (receptor do ácido retinóico X retinóide) não ligado. À direita, a estrutura em cristal do LBD do RAR γ (receptor do ácido retinóico) ligado. Os cilindros representam as α -hélices e são numerados de 1 a 12. Quando o receptor está ligado (à direita) ele se apresenta de forma mais compactada, mostrando uma mudança conformacional. Note as duas posições diferentes da hélice 12 (C-terminal) contendo o núcleo do domínio AF-2. (Aranda & Pascual, 2001).

O domínio E é responsável por muitas funções de ativação ou repressão, como exemplo, a função de transativação (AF2). Ele também pode conter outro NLS. O domínio F não apresenta ainda função conhecida (Robinson-Rechavi *et al.*, 2003; Pascual & Aranda, 2001; Lazar *et al.*, 2002).

Quando os receptores hormonais nucleares estão ativados (ligados a seus respectivos hormônios), os mesmos se ligam a seqüências específicas do DNA chamados elementos de resposta hormonal (HREs: *hormone response elements*). Como efeito, genes são transcritos e traduzidos em proteínas gerando diferentes respostas dentro da célula (Figura 9). A ligação ao HRE é mediada pelo domínio central C de um receptor hormonal. Este domínio é compreendido pelo motivo $Zn(Cys_4)_2$ (Lazar *et al.*, 2002).

3.3.3 – Motivo Zn_2C_6 e Motivo $Zn(C_4)$

O fator de transcrição GAL4 em leveduras (normalmente responsável por ativar a transcrição de genes que codificam enzimas que convertem galactose à glicose neste microrganismo) contém seis cisteínas e dois átomos de zinco ($Cis_2-Zn-Cis_2-Zn-Cis_2$). As cadeias laterais e aminoácidos cisteínas coordenam em duas curtas α -hélices, unidas por uma volta oposta ao átomo de zinco. O dímero binuclear Zn reconhece 17 pares de bases no DNA (Pabo & Sauer, 1992).

Estudos extensivos sobre metabolismo do nitrogênio e sua regulação têm sido realizados em *Saccharomyces cerevisiae*, *Aspergillus nidulans* e *Neurospora crassa*. Experimentos têm demonstrado que a grande maioria das proteínas envolvidas com o controle transcricional de genes estruturais relacionados ao metabolismo do nitrogênio em fungos possuem os motivos Zn_2/Cis_6 ou Zn/Cis_4 (GATA-binding). Um exemplo é a proteína Nit2 de *Neurospora crassa* cujo motivo Zn/Cis_4 (GATA-binding) é o responsável pelo reconhecimento de elementos localizados a jusante de alguns genes estruturais relacionados à regulação positiva do metabolismo do nitrogênio, dentre os quais: *nit3* (nitrato redutase), *alc* (alantoicase) e *lao* (L-amino-oxidase). Em *Saccharomyces cerevisiae* a

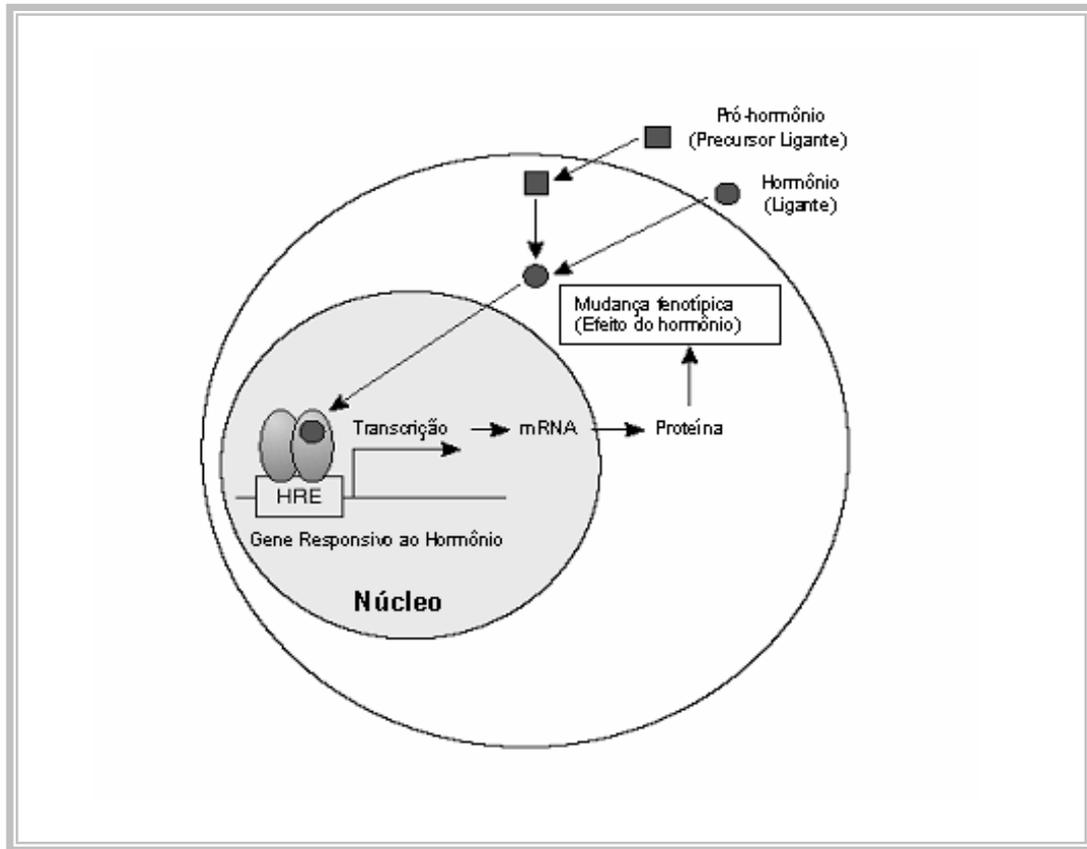


Figura 9 – Transdução de sinal de hormônios e outros ligantes que atuam via receptores nucleares. HRE, elemento de resposta ao hormônio (Lazar *et al.*, 2002).

proteína ArgnII contendo o motivo Zn_2/Cis_6 participa do controle da expressão de enzimas envolvidas em síntese e degradação da arginina (Marzluf, 1997).

Várias outras proteínas contendo o motivo Zn/Cis_4 (GATA-binding) também têm sido identificadas atuando com fatores de transcrição em diferentes organismos, dentre os quais: nematóides, plantas e vertebrados (Harrison, 1991; Lowry & Atchey, 2000).

3.3.4- Motivo $Zn_2C_3HC_4$ – *ring finger*

O motivo *ring finger* está presente em várias proteínas de origens diferentes. Neste grupo estão incluídas as proteínas Rad5 de leveduras envolvidas em reparo de DNA e a proteína humana RAG1 essencial para o rearranjo de imunoglobulinas. Proteínas contendo estes motivos têm sido relacionadas como reguladoras de vários processos celulares, dentre os quais a transcrição envolvendo interações proteína-proteína, transporte de RNA, transdução de sinal e ubiquitinação de várias proteínas (Borden, 2000).

A estrutura de três proteínas contendo este motivo já foi determinada: o peptídeo 63 do vírus tipo 1 da herpes equina, a proteína humana RAG1 e uma proteína da leucemia promielocítica humana. Este motivo é coordenado por dois átomos de zinco e a estrutura secundária consiste de duas folhas β e uma α hélice formando dois *loops* simétricos (Roth & Leon, 2000).

A proteína humana TRM11, membro do chamado tripartido motivo (TRIM) é composta por um motivo *ring finger*, um motivo E3 (ubiquitina ligase) e um motivo B-box (*coiled-coil*). Ensaio de duplo híbrido em levedura demonstraram que a proteína TRM11 interage com a proteína denominada humanina, um peptídeo neuroativo, envolvido na supressão da Doença de Alzheimer. Esse peptídeo foi detectado em cérebro humano, cólon e testículo de ratos. Experimentos demonstram que a deleção de cisteínas do motivo *ring finger* ou tratamento inibindo o proteossoma diminuem significativamente o nível intracelular de humanina. Esses resultados sugerem que TRM11 pode ter um papel na regulação do nível intracelular da humanina (Niikura, 2003).

Horie *et al.* (2003) identificaram várias regiões a jusante do gene para o receptor do hormônio estrogênio que incluem os elementos responsivos ao estrogênio. Uma dessas regiões codifica uma proteína *ring finger* denominada Epf (*estrogen responsive finger protein*). A estrutura de Epf é formada por um domínio *ring finger*, dois B-Boxes, um domínio de α -hélice *coiled coiled* e um domínio C-terminal SPRY (Figura 10). Epf é predominantemente expressa em tecidos alvos ao estrogênio, incluindo glândulas mamárias, útero e osteoblastos, e também em câncer de mama. Epf é essencial para o crescimento de órgãos das fêmeas, como o útero, uma vez que camundongos fêmeas deficientes no gene Epf não desenvolvem este órgão (Ikeda *et al.*, 2000).

Recentes avanços na pesquisa molecular revelaram que um grande número de proteínas *ring finger* funciona como proteínas-ubiquitinas ligases ou E3 sinalizando o caminho de ubiquitinação. A ubiquitinação regula uma variedade de funções celulares, frequentemente por mediar a degradação seletiva de proteínas regulatórias através do proteassoma. A proteína Epf é uma candidata a funcionar como uma ligase E3. Experimentos de duplo híbrido em leveduras revelam que Epf interage com uma proteína denominada 14-3-3 σ . Esta proteína participa do controle do ciclo celular, sequestrando cdc2 para o citoplasma e conseqüentemente induzindo a célula a parar em G2. Experimentos de *pulse-chase* confirmaram que a degradação da proteína 14-3-3 σ é potencializada em células MCF7 transfectadas com o gene Epf. A degradação de 14-3-3 σ é proteassoma dependente, uma vez que experimentos utilizando o inibidor de proteassoma MG132 resultou no aumento da proteína 14-3-3 σ ligada a Epf. Esses resultados sugerem que o Epf pode prover uma proliferação ilimitada de células cancerígenas na mama por acelerar a destruição da proteína 14-3-3 σ neste órgão (Figura 11) (Horie *et al.*, 2003).

Foi demonstrado que várias outras proteínas contendo o motivo *ring finger* são conhecidas por serem responsáveis por muitos tumores malignos (Figura 12). Por exemplo, a proteína PML (*promyelocytic protein*) é responsável pela leucemia promielocítica aguda quando ocorre uma fusão com o receptor do ácido retinóico (RAR α) (Jensen *et al.*, 2001). Uma diminuição no gene supressor

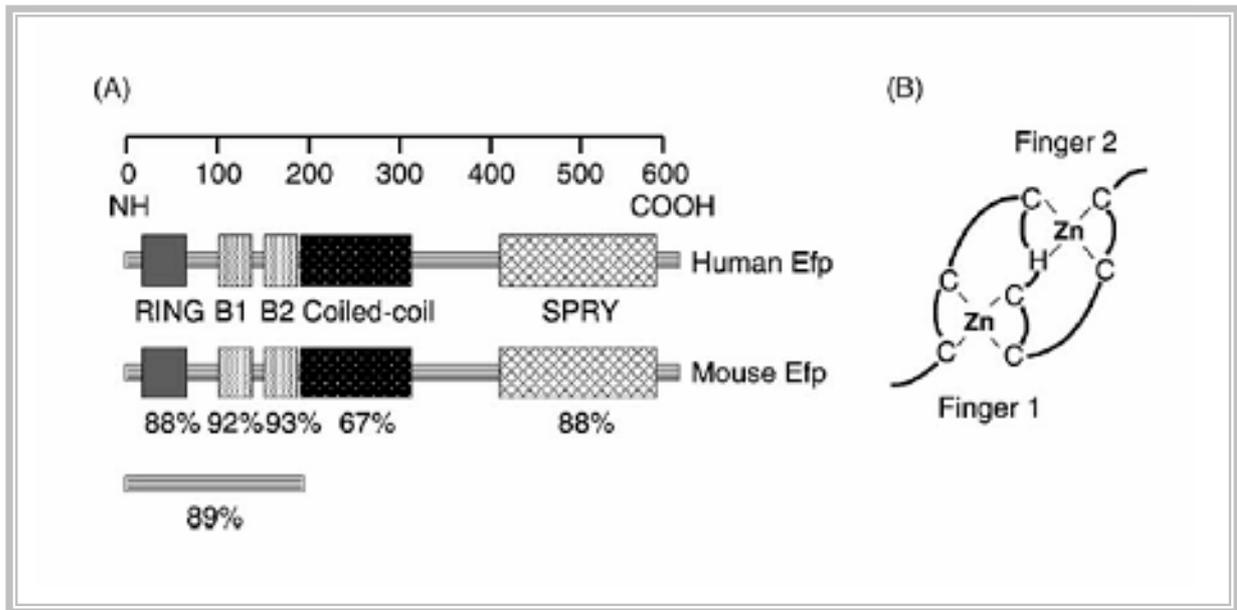


Figura 10 - Estrutura da proteína Efp: (A) Estrutura do domínio da Efp humana e de camundongo. As caixas indicam os vários domínios de Efp. Embaixo das caixas representando os domínios é mostrada a percentagem de similaridade entre o Efp humano e o Efp de camundongo; (B) Estrutura do motivo *ring finger*. Quatro cisteínas são coordenadas por um átomo de zinco juntamente com três outras cisteínas e uma histidina também coordenadas por um átomo de zinco (Modificado de Horie *et al.*, 2003).

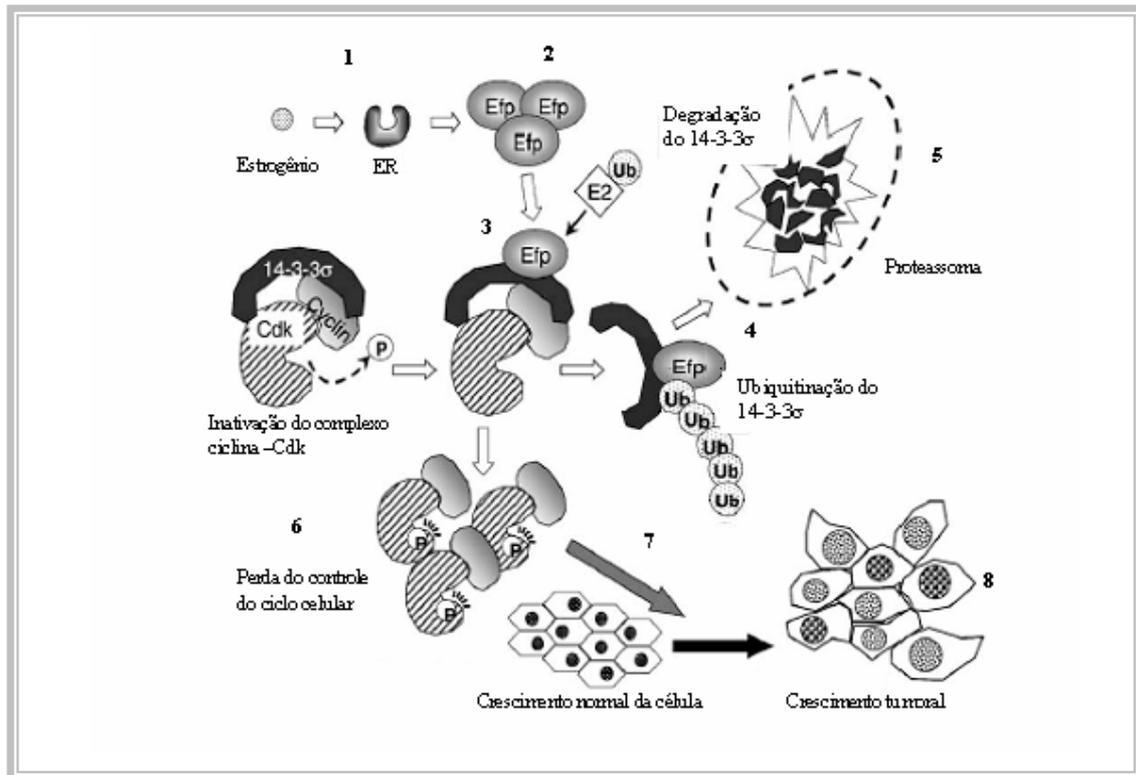


Figura 11 – Mecanismo do crescimento tumoral desencadeado pela proteína *ring finger* Efp em resposta ao estrogênio. (1) Presença do hormônio estrogênio dentro da célula, (2) O hormônio esteróide estrogênio ativa Efp, (3) Em resposta ao estrogênio, Efp interage com a proteína 14-3-3 σ (responsável pelo seqüestro de ciclina/cdc2 no citoplasma, levando a célula a parar em G2 da interfase), (4) A proteína Efp atua como uma E3 ubiquitina ligase, adicionando uma cauda de ubiquitina a proteína 14-3-3 σ resultando na proteólise da mesma, (5) A proteína 14-3-3 σ ubiquitinada é degradada via proteassoma, (6) A degradação da proteína 14-3-3 σ impede que a célula pare em G2, ocorrendo a perda do controle do ciclo celular, (7) Indução de crescimento de células tumorais, (8) Crescimento tumoral (Modificado de Horie *et al.*, 2003)

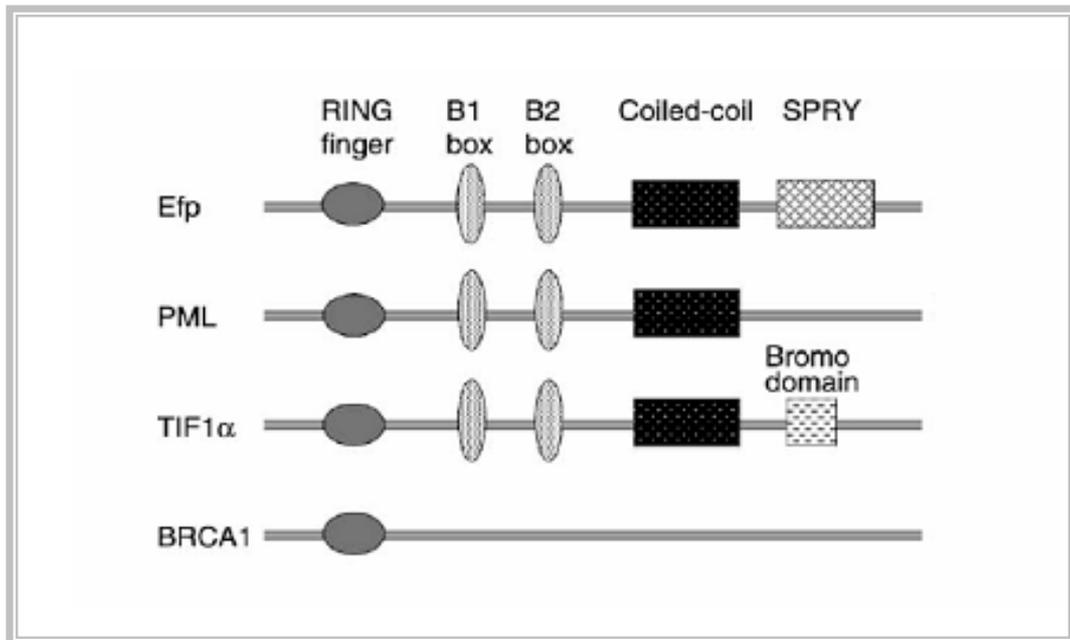


Figura 12 – Estrutura dos domínios de vários membros da família *ring finger*. Efp: Elemento responsivo ao estrogênio relacionado à promoção de crescimento celular; PML, responsável pela leucemia promielocítica aguda; TIF1 α , regulador transcricional relacionado a tumor de fígado e BRCA1 um supressor de tumor relacionado ao câncer de seio e ovário (modificado de Horie *et al.*, 2003).

de tumor BRCA1 resulta em instabilidade cromossômica levando ao desenvolvimento de um câncer familiar de mama e ovário (Ruffner *et al.*, 2001).

Garcia-Estrada *et al.* (2003) identificaram em *Leishmania infantum* uma proteína denominada LORIEN. Essa proteína contém um domínio conhecido por SP-RING ou Miz, pertencente ao motivo RING-finger, encontrado em proteínas que atuam com E3 ligase em conjunto com pequeno modificador relacionado a ubiquitinas (SUMO). SUMO é altamente conservado desde leveduras até humanos. Em contraste com o sistema de ubiquitina, cujo objetivo geralmente é marcar as proteínas para degradação, o processo conhecido como *SUMOylation* parece modular a localização ou estabilidade das proteínas alvo.

3.3.5 – Motivo Zn (C₃H)

O motivo *zinc finger* CCCH tem sido encontrado em diversas proteínas que se ligam a RNA, estando envolvido em vários aspectos no controle da homeostase e diferenciação celular. Duas proteínas, denominadas tbZFP1 e tbZFP2 contendo cada uma um motivo CCCH, foram identificadas em *T. brucei*, sendo implicadas na regulação da morfogênese e diferenciação deste parasita. Experimentos com interferência de RNA demonstram que o gene *TbZP2* é essencial para que ocorra a diferenciação das formas sanguíneas para procíclica em *Trypanosoma brucei* (Hendriks *et al.*, 2001). A proteína TbZFP1 é altamente expressa durante a diferenciação da forma sanguínea para a procíclica e experimentos de perturbação genética desse gene compromete o reposicionamento do genoma mitocondrial, um evento específico ocorrido durante a diferenciação deste parasita (Hendriks *et al.*, 2005).

Membros da família ZFP de *T. brucei* foram também identificados em *T. cruzi*: ZFP1A, ZFP1B, ZFP2A e ZFP2B. Além do domínio *zinc finger*, essas proteínas ZFPs contém outros domínios de ligação a proteínas (WW e PPXY) sugerindo a ligação de umas com as outras. Experimentos de duplo híbrido em leveduras revelaram que as interações entre ZFP1A e ZFP2A, ZFP 1A e ZFP2B, ZFP1B e ZFP2B (Caro *et al.*, 2005). Hendriks *et al.* (2003) identificaram em *T.*

brucei uma proteína denominada TbCPSF30. Essa proteína contém cinco motivos CCCH e é homóloga à proteína CPSF (fator específico de clivagem e poliadenilação), um dos componentes da maquinaria requerida na formação do 3' final em leveduras e mamíferos. Experimentos com interferência de RNA demonstram que este gene é essencial para as formas sanguíneas e procíclicas de *Trypanosoma brucei*. O RNAi de TbCPSF30 resulta no acúmulo de um mRNA TbCPSF30 aberrante e na depleção da proteína tbCPSF30. A depleção dessa proteína gera o acúmulo de mRNAs de tubulinas não processados, implicando, portanto, na participação de TbCPSF30 no processamento policistrônico de RNA. Este é o primeiro estudo que identifica e caracteriza uma proteína CPSP em *T. brucei* e que demonstra o seu envolvimento no processamento policistrônico de RNAs (Hendriks *et al.*, 2003).

Morking *et al.* (2004) identificaram em *T. cruzi* duas proteínas denominadas TcZFP1 e TcZFP2 contendo cada uma delas um motivo CCCH. Experimentos de *gel shift* mostraram a capacidade da proteína tcZFP1 em interagir especificamente com oligorribonucleotídeos poli(C). Proteínas que se ligam a poli(C) estão normalmente envolvidas com as atividades de estabilidade de mRNA e regulação traducional em eucariotos.

3.3.6 – Motivo Zn (CCHC)

O reovírus apresenta no seu genoma o gene S4 que codifica uma proteína de 41 kDa conhecida por δ_3 . Esta proteína desempenha vários papéis durante a replicação viral, tais como o processamento correto de proteínas virais formando o complexo δ 3- μ 1C, que constitui o maior componente de capsídeos maduros do reovírus, na inibição de síntese de RNA e proteínas de células hospedeiras, no estabelecimento da infecção viral, na estimulação da tradução do RNA viral e na inibição da ação do interferon (esta proteína contém o domínio CX₂CX₁₂HX₃C) (Mabrouk & Lemay, 1994).

As proteínas de nucleocapsídeos virais dos lentivírus (NCps) contêm um ou dois domínios CX₂CX₄HX₄C altamente conservados e homólogos aos NCp7 do vírus da imunodeficiência humana

HIV-1. Esta proteína está envolvida em várias etapas chave do ciclo de vida viral (empacotamento do RNA, transcrição reversa e integração) através de interações com ácidos nucleicos de fita simples (RNA e também com outras proteínas virais) (Rosbash *et al.*, 1997; Ramboarina *et al.*, 1999). A encapsidação do genoma do HIV-1 durante a montagem do retrovírus é mediada entre os domínios do nucleocapsídeo e um segmento genômico de 110 nucleotídeos conhecido como *Psi-site*. O *Psi-site* contém quatro *stem-loops* (SL1, SL2, SL3 e SL4) alvo de interação destas proteínas (Amarasinghe *et al.*, 2001)

Experimentos também demonstram o papel de cada um dos dois motivos CCHC presentes no nucleocapsídeo viral tipo 1 de HIV na ligação a ácidos nucleicos. Tem sido demonstrado que estes nucleocapsídeos também funcionam como chaperoninas anelando a ácidos nucleicos e atuando na dimerização do genoma e maturação das partículas virais. Ensaio de anelamento com proteínas contendo este motivo mutado têm sido realizados para identificar quais são os resíduos necessários para o desempenho da atividade de chaperonina (Heath *et al.*, 2003).

Os dois domínios repetidos CCHC de NCp7 possuem estrutura similar. Em experimentos realizando mutações, como a troca de uma histidina por uma cisteína, demonstrou que ocorre uma modificação na região dos *zinc fingers* resultando em um vírus não infeccioso. Já a substituição de CCHC por CCHH afetou a integridade estrutural dessa proteína, resultando numa drástica alteração no reconhecimento dos ácidos nucleicos pelo vírus, bem como outras proteínas celulares (transcriptase reversa e integrase) (Ramboarina *et al.*, 1999). Experimentos posteriores com NCp7 indicaram que primeira zinc finger CCHC do nucleocapsídeo em sua conformação intacta é a responsável pela ótima interação com a transcriptase reversa, mantendo portanto a infectividade do HIV (Ramboarina *et al.*, 2004).

O motivo CCHC também está presente em NCps de vírus mosaico de couve-flor e no produto do gene *Xpo* regulado no desenvolvimento em *Xenopus*. Além destes, um gene foi clonado em *Saccharomyces cerevisiae* expressando uma proteína com motivos CCHC envolvido em *cis*

splicing. Essa proteína SLU7 interage com os pequenos RNAs U5 nucleares durante a seleção do sítio 3' apropriado (Covey *et al.*, 1986; Sato & Sargent, 1991; Frank *et al.*, 1992).

Chang *et al.* (2003) identificaram e caracterizaram uma proteína nucleolar denominada pN040. A mesma foi identificada pela técnica de duplo híbrido em leveduras utilizando uma biblioteca de cDNA de fígado humano e uma proteína denominada pinin (pnn) como isca. A proteína pinin é o produto de um suposto gene supressor de tumor, sendo identificada inicialmente como uma proteína associada à desmossomos (Ouyang & Sugure, 1996) e subsequentemente encontrada em núcleo de células em intérfase (Ouyang, 1999). A proteína pN040 apresenta vários domínios distintos, incluindo uma região de ligação a RNA presente em proteínas ribossomais, um *zinc finger* do tipo CX₂CX₄HX₄C e vários grupos de aminoácidos básicos representando um possível sinal de localização nuclear. Experimentos de imunocitocalização com a proteína endógena ou pNO40 transfectada em diversas culturas de células demonstraram a concentração da mesma em muitos componentes granulares do nucléolo. Sugere-se que a pN040, devido as suas características de seqüência e localização subcelular, possa ser uma nova proteína nucleolar com função relacionada à maturação e biogênese de ribossomos (Chang *et al.*, 2003).

Em *T. brucei* foi identificada uma proteína denominada RRM1 que contém três motivos de reconhecimento a RNA (RRMs), duas cópias *in tandem* de motivos *zinc fingers* CX₂CX₄HX₄C e uma região rica em arginina-serina (SR). Essa proteína de 70 kDa, se localiza no núcleo e é expressa tanto na forma sanguínea quanto na procíclica. RRM1 apresenta homologia estrutural com famílias de proteínas SR (em particular 9G8, um fator de *cis splicing* que atua em mRNAs de mamíferos), o que sugere que o mesmo possa desempenhar um papel relacionado ao *trans-splicing* em *T. brucei* (Manger & Boothoyd, 1998).

A HEXBP (*hexamer binding protein*) contém nove motivos *zinc finger* CX₂CX₄HX₄C que se liga na região 5' não traduzida do gene da mais abundante glicoproteína de membrana em *Leishmania chagasi*. A HEXBP pode potencialmente estar envolvida na replicação, recombinação, reparo e estrutura do DNA, além de estar envolvida com a regulação transcricional. Esta proteína se

liga a fita simples de ácidos nucleicos e se liga a oligonucleotídeos de forma similar à proteína humana CNBP (*cellular nucleic acid binding protein*) que apresenta sete motivos *zinc finger* CX₂CX₄HX₄C de ligação a ácido nucleico de fita simples. Estas proteínas se ligam a oligonucleotídeos que correspondem a uma fita de uma região consenso conhecida como elemento regulatório de esteróis e parecem funcionar como um repressor de vários genes envolvidos no metabolismo de esteróis (Webb & McMaster, 1993).

Em *Crithidia fasciculata* foi descrita uma proteína denominada de UMSBP (*UMS binding protein*) que apresenta cinco motivos *zinc finger* CX₂CX₄HX₄C que se liga à fita simples de DNA de minicírculo de cinetoplasto. Essa proteína se liga em uma região conservada contendo 12 nucleotídeos (5'-GGGGTTGGTGTA-3') conhecida como UMS (*universal minicircle sequence*) que é a seqüência que marca o início da replicação da fita pesada de minicírculos. Outra seqüência o hexâmero (5'-ACGCCC-3'), presente na fita leve de minicírculos também é alvo de interação por UMSBP. O fato de UMSBP interagir com a região do minicírculos e do maxicírculo a torna uma forte candidata à participação do início do processo de replicação nestes kDNAs. Essa proteína também pode interagir com RNA simples fita simples (Tzfati *et al.*, 1992; Abu-Elneel *et al.*, 1999).

Em *Trypanosoma brucei* foi identificada uma região genômica contendo um gene que codifica para uma subunidade beta proteassomal e um gene que codifica para uma proteína com o motivo *zinc finger* CX₂CX₄HX₄C, porém, a função desta *zinc finger* ainda é desconhecida (Radwanska *et al.*, 2000).

Em *Trypanosoma cruzi* foi identificada uma proteína denominada PDZ5 contendo cinco motivos *zinc finger* CX₂CX₄HX₄C (homóloga a proteína UMSBP de *Crithidia fasciculata*) que se liga especificamente à fita simples de DNA conhecida como UMS (Coelho *et al.*, 2003).

Em *Trypanosoma cruzi* foi também identificada uma proteína denominada PZFP1 contendo sete motivos CX₂CX₄HX₄C. Experimentos de gel de retardamento demonstraram que tanto a proteína nativa PZFP1 quanto a recombinante, interagem especificamente com oligonucleotídeos de DNA ou RNA carregando seqüências reconhecidas por outras proteínas contendo o motivo CCHC.

Esta proteína é localizada tanto no núcleo quanto no citoplasma. Ensaio de duplo híbrido em leveduras demonstraram que PZFP1 também interage com uma proteína rica em resíduos de serinas e argininas de *T. cruzi* (TcSR), sugerindo o papel da mesma em processamento de mRNA (Espinosa *et al.*, 2003).

Outras proteínas contendo um motivo denominado *zinc finger* CXXCXGXG tem sido descritas em chaperoninas da família DnaJ de *Trypanosoma cruzi*, dentre as quais tcj2, tcj3 e tcj4. Este motivo parece estar envolvido na interação destas proteínas com outras proteínas pertencentes à família DnaJ (Tibbetts *et al.*, 1998).

II – Relevância e Justificativa do projeto

A doença de Chagas é uma das mais importantes endemias dentre as que ocorrem na América Latina, afetando cerca de 16 milhões de pessoas.

O recente seqüenciamento do genoma de tripanosomatídeos parasitas humanos e animais (*T. cruzi*, *T. brucei* e *L. major*) (El-Sayed *et al.*, 2005; Berriman *et al.*, 2005; Ivens *et al.*, 2005) confirmou diferenças e revelou outras novas entre estes e os demais eucariotos. Contudo várias moléculas sinalizadoras, em um grupo diverso de proteínas quinases e fosfatases, proteínas que se ligam a ácidos nucléicos, dentre outras já identificadas implicam previamente em interações ainda não conhecidas, bem como processos regulatórios também desconhecidos e que poderão vir a ser possíveis alvos para a intervenção na doença. Os mecanismos moleculares, dentre os quais, os de expressão gênica e a sua regulação nestes parasitas são pouco conhecidos. Assim, a caracterização do genoma do *Trypanosoma cruzi* e os demais tripanosomatídeos e a elucidação da função de suas proteínas e suas interações podem trazer respostas importantes para a compreensão da biologia molecular do mesmo, bem como possibilitar o desenvolvimento de novas terapias para doença de Chagas.

Recentemente, Zapata-Estrella *et al.* (2006) demonstrou que a imunoterapia com vacinas de DNA codificando o antígeno TSA-1 (*trypomastigote surface antigen 1*) de *T. cruzi* foi capaz de reduzir danos ao tecido cardíaco e melhorar a sobrevivência de camundongos quando administrados durante as fases aguda ou crônica da infecção por *Trypanosoma cruzi*. Ainda assim pode-se dizer que desde a sua descoberta em 1909 até o presente momento não foi disponibilizada uma vacina testada em humanos ou uma droga realmente eficaz para o tratamento da Doença de Chagas. Os quimioterápicos disponíveis possuem baixa toxicidade seletiva, sendo restritos ao tratamento da fase aguda da doença. Deste modo, a investigação estrutural e funcional de regiões do genoma de *Trypanosoma cruzi* e suas proteínas pode vir a solucionar várias questões relacionadas à biologia do mesmo, como por exemplo a compreensão das relações interação hospedeiro-parasita e moléculas envolvidas.

Em trabalhos anteriores do nosso grupo, Bartholomeu *et al.* (2001) isolaram um clone genômico correspondente a um fragmento de 17 kb na banda XX de *T. cruzi*, utilizando como sonda um cDNA que codifica a subunidade $\alpha 5$ de proteassoma. Após subclonagens e seqüenciamento da região a montante a este gene, foi verificado que havia a presença de uma região com fase de leitura de 5,2 kb predizendo uma proteína homóloga à cadeia pesada de clatrina (Assis *et al.*, 1998) e uma fase de leitura de 1,1 kb predizendo uma subunidade β proteassomal (Lima & Assis, 1998) (Figura 13).

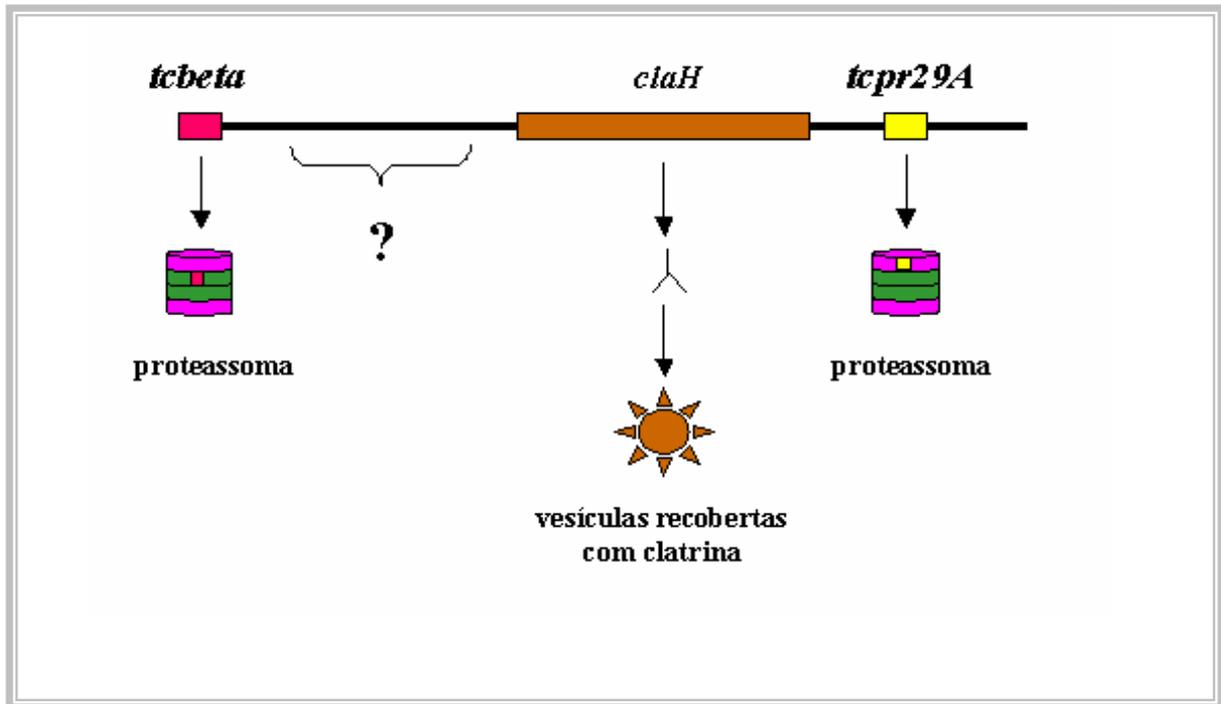


Figura 13 – Esquema da região genômica de 17 kb da banda XX de *T. cruzi*. Retângulos correspondem às regiões de fase de leitura. *TcBETA5* - gene para subunidade beta proteassomal; *TcPR29A* – gene para subunidade alfa proteassomal; *TcCLAH* – gene para uma proteína homóloga à cadeia pesada de clatrina. A região indicada com colchete e interrogação foi estudada nesta tese.

III – Objetivos

Esse trabalho teve como objetivo geral a caracterização da região genômica localizada no cromossoma XX de *Trypanosoma cruzi*, entre o gene que codifica para subunidade beta proteassomal e o gene que codifica para a proteína homóloga à cadeia pesada de clatrina.

Para tal objetivo, as seguintes metas foram propostas:

- Clonagem e seqüenciamento dos genes presentes na região genômica acima;
- Análise das fases de leitura dessa região e análise de homologias das proteínas preditas em bancos de dados;
- Comparação desta região genômica com a mesma região em *T. brucei* e *L. major*;
- Análise filogenética das proteínas identificadas e suas homólogas;
- Expressão heteróloga do gene *TcZFP8* em *E. coli*;
- Purificação da proteína recombinante TcZFP8;
- Produção de anticorpos policlonais contra a proteína TcZFP8;
- Análise da expressão do gene *TcZFP8* em experimentos de *Northern blot* e *Western blot* nas diferentes formas de *T. cruzi*: epimastigotas, amastigotas e tripomastigotas;
- Citolocalização da proteína TcZFP8 em células epimastigotas de *T. cruzi*;
- Análise da ligação da proteína TcZFP8 a ácidos nucleicos (DNA de fita simples, DNA de fita dupla ou RNA);
- Seleção da seqüência nucleotídica de ligação da proteína TcZFP8 em experimento de SELEX.

MATERIAL E MÉTODOS

I – LINHAGEM DE CÉLULAS

1 – *Trypanosoma cruzi*

Foram utilizados a cepa CL Brener (Brener & Chiari, 1963), Dm 28c (Contreras *et al.*, 1988) e cepa Y a partir de uma amostra venezuelana de *Didelphis marsupialis* gentilmente cedidos pelo Dr. Samuel Goldenberg (Instituto Oswaldo Cruz - RJ). As linhagens de *T. cruzi* foram mantidas em nitrogênio líquido ou mantidas em 2 ml de meio LIT a 28°C, sendo que os repiques foram realizados semanalmente com um inóculo de 10%.

2 - Células L6

Foi utilizada a linhagem L6 (células musculares cardíacas de camundongos) para os experimentos de infecção. As células foram gentilmente cedidas pelo Dr. Klaus Scherrer (Instituto Jacques Monod – França) e mantidas em nitrogênio líquido ou mantidas em meio 199 com repiques semanais, como descrito a seguir: o meio 199 suplementado com 10% de soro fetal bovino inativado foi utilizado para a multiplicação das células. Após o crescimento celular, o meio de cultivo era descartado e 3 ml de solução de verseno contendo 1% de tripsina eram adicionados às garrafas e as células eram mantidas nesta solução à temperatura ambiente por aproximadamente 30 seg. Após este período, a solução de verseno era descartada e as células incubadas a 37°C por 10 min para que se descolassem da garrafa. Posteriormente, 2 ml de meio 199 suplementado com soro fetal bovino eram adicionados em cada cultura para a homogeneização das células e então 200µ da suspensão de células eram inoculadas em garrafas de 25 cm³ (Falcon) contendo 10 ml de meio 199 suplementado com soro fetal bovino. As células eram incubadas a 37°C por uma semana, quando era realizado um novo repique.

3 - *Escherichia coli*

As linhagens de *E. coli* utilizadas neste trabalho, com os seus respectivos genótipos, estão apresentadas na Tabela 1. As linhagens bacterianas foram mantidas a -80°C em meio LB contendo 15% de glicerol.

Tabela 1: Linhagens de *E. coli* utilizadas.

Estirpe	Genótipo	Referência
DH5 α	endA1, recA1, hsdR17 (rK ⁺ , mK ⁺), supE44, thi-1, gyrA, relA1, F ⁻ , (Φ 801lacZ Δ M15), Δ (lacZYA-argF) U169	Sambrook <i>et al.</i> , 1989.
BL21(DE)3	F- ompT gal[dcM][lon] hsdS _B (r _B - m _B -, an <i>E. coli</i> B stain) with DE3, a λ prophage carrying the T7 RNA polymerase gene	Studier <i>et al.</i> , 1990
XL-1 blue	F ['] ::Tn10 proA+B+ lacI ^q Δ (lacZ)M15/recA1 endA1 gyr96 (NaI ^r) thi hsdR17 (r _k -m _k +) supE44 relA1 lac	Bullock <i>et al.</i> , 1987

II – PLASMÍDEOS

Foram utilizados os plasmídeos: pBluescript SK+ (Stratagene), pGEM-T (Promega), pUC18 (Biolabs), pGEX-4T-1 (Pharmacia), pET28 a+ (Novagen), pGFP-C1(Clontech) e p33 (gentilmente cedido pelo Dr. Schenkman).

III – MEIOS DE CULTURA

Os reagentes utilizados no preparo das soluções e meios de cultura foram os de maior grau de pureza disponível no laboratório e a água usada era bidestilada ou milli-Q.

As soluções e meios de cultura eram esterilizados a 120°C durante 20 minutos e/ou, quando necessário, eram filtradas em membrana do tipo Millipore com poro de $0,22\ \mu\text{m}$.

1 - Meio SOB

bacto-triptona 2,0%

extrato de levedura	0,5%
NaCl	0,06%
KCl	0,05%

pH 7,4

Após esterilização em autoclave, era adicionado 10 mM de MgCl₂.

2 - Meio SOC

Meio SOB adicionado de 20 mM de glicose.

3 - Meio 2YT

peptona de caseína	1,6%
extrato de levedura	1,0%
NaCl	0,5%

pH 7,2

4 - Meio LIT

NaCl	0,4%
KCl	0,04%
Na ₂ HPO ₄	0,8%
Glicose	0,2%
Triptose	0,5%
Infuso de Fígado	0,5%
Hemina	0,0025 %
Soro Fetal Bovino	10%

O extrato de fígado foi previamente dissolvido em água destilada, fervido em placa aquecedora sob agitação durante 1 h e filtrado em papel Whatman nº1. A hemina foi adicionada ao meio a partir de uma solução estoque de 0,5% em 5% de quadrol. O pH do meio foi ajustado para 7,2 com HCl concentrado e este foi esterilizado por filtração. Após a esterilização do meio, o soro fetal bovino já inativado e esterilizado (Gibco BRL) e a ampicilina (100 µg/ml) esterilizada por filtração através de filtro Millipore 0,22 µm eram adicionados.

5 - Meio 199 (10 X)

Meio 199 (Gibco BRL)	20 g
Água bidestilada qsp.	200 ml

O meio foi esterilizado por filtração através de filtro Millipore 0,22 µm, distribuído assepticamente em alíquotas de 50 ml e armazenado a -20°C.

6 - Meio 199 suplementado com soro fetal bovino

Meio 199 (10 X)	10%
Bicarbonato de Sódio (4,4%)	2%
Soro Fetal Bovino inativado (Gibco BRL)	5%
Tampão Hepes 1M pH 7,0	1%

Todos os componentes do meio foram previamente esterilizados por filtração através de filtro Millipore 0,22 µm. O meio foi alíquotado assepticamente em garrafas e armazenado a 4°C.

IV – REAGENTES E SOLUÇÕES

Os reagentes utilizados no preparo das soluções foram os de maior grau de pureza disponível no laboratório e a água usada era bidestilada ou milli-Q.

1 - Enzimas e tampões

As enzimas utilizadas foram adquiridas principalmente das empresas Promega, Pharmacia, New England Biolabs e Amersham. O procedimento usado foi o indicado pelo fabricante.

2 - Soluções

Solução I	Tampão Tris-HCl pH 8,0	25 mM
	EDTA	10 mM
	Glicose	50 mM
Solução II	NaOH	0,2 M
	SDS	1 %
Solução III	Acetato de Sódio	3 M
	Ácido Acético	2 M

pH 4,8 - 5,0

RNase A 10 mg da enzima por mL de água Milli-Q ou bidestilada e fervida em banho-maria durante 20 minutos.

Fenol Saturado Fenol 1 V
 β -hidroxiquinolina 0,05 %
 equilibrada com tampão TE

Tampão TE Tampão Tris-HCl 25 mM, pH 8,0
 EDTA 10 mM

Clorofil Clorofórmio 24 V
 Álcool Isoamílico 1 V
 equilibrada com tampão TE

Solução STET NaCl 100 mM
 Tris-HCl pH 8,0 10 mM
 EDTA 1 mM
 Triton X-100 5 %

Tampão de eluição em gel de poliacrilamida

Acetato de amônio 0,5 M
 EDTA 1,0 mM
 SDS 0,1 %
 pH 8,0

Tampão PBS 10X

NaCl 137 mM
 Na_2HPO_4 7 mM
 NaN_3 0,02 % (p/v)
 pH 7,2

Tampão PBSM 1X

Tampão PBS 1 X
 Leite desnatado em pó 2 % (p/v)

Solução reveladora para peroxidase

Cloranaftol 0,3% 2,5 ml
 PBS 1X 7,5 ml
 H_2O_2 10V 0,05 ml

Tampão de transferência

Tris-base 48 mM
 Glicina 39 mM
 SDS 0,037 %
 Metanol 20 %

Tampão de corrida TEB 10X

Trizma-base	0,89 M
Ácido Bórico	0,89 M
EDTA	0,08 M
pH 8,0	

Tampão de amostra TEB 10X

Glicerol	50 %
Azul de Bromofenol	0,01%
Xileno cianol	0,01%
TEB	20 X

Tampão de lise SDS-PAGE 2X

Tris-HCl	250 mM (pH 6,8)
SDS	8 %
2-Mercaptoetanol	20 %
Azul de Bromofenol	0,04 %

Tampão de amostra SDS-PAGE 2X

Tris-HCl	160 mM (pH 6,8)
Sacarose	24,0 %
SDS	4 %
2-Mercaptoetanol	4 %
Azul de Bromofenol	0,04 %

Tampão de Corrida SDS-PAGE 5X

Trizma-base	125 mM
Glicina	0,96 M
SDS	5,0 %

Solução Corante SDS-PAGE

Metanol	40%
Ácido acético	10 %
Coomassie blue R250	1 %

Solução Descorante SDS-PAGE

Metanol	30 %
Ácido acético	7 %

Tampão de equilíbrio

Fosfato monobásico de sódio	0,1 M
NaCl	50 mM
pH 7,3	

Tampão de lavagem

Fosfato Monobásico de sódio	50 mM
-----------------------------	-------

Uréia	8 M
Imidazol	0,08 M
pH 6,5	

Tampão de eluição (Níquel)

Fosfato Monobásico de sódio	50 mM
Uréia	8 M
Imidazol	0,3M
pH 6,5	

Tampão de eluição (GST)

Glutaciona reduzida	50 mM
Tris-HCl pH 6,8	50 mM

Tampão de Klenow 10X

Tampão Tris-HCl pH 7,6	0,5 M
MgCl ₂	0,1 M
DTT	1 mM

Tampões de alquilação:**Solução A (4X)**

Tris-HCl pH 8,0	20 mM
SDS	4 %
DTT	4 %

Solução B (4X)

Tris-HCl pH 8,0	20 mM
Iodoacetamida	500 mM

Tampão TENM2 10X:

Tris-HCl pH 7,0	100 mM
NaCl	100 mM
MnCl ₂	10 mM
MgCl ₂	10 mM
2-Mercaptoetanol	0,036%
H ₂ O	85,9%

Tampão de transfecção para *T. cruzi*

NaCl	137 mM
HEPES pH 7,0	21,0 mM
Na ₂ HPO ₄	5,5 mM
Glicose	0,77 mM

Tampão de ligação SELEX

HEPES pH 7,9	10 mM
KCl	70 mM
MgCl ₂	2,5 mM
EDTA	1,0 mM
ZnCl ₂	1,0 mM

ZnSO ₄	0,01 mM
Glicerol	4,0%
DTT	1 mM
Poli (dIdC)	100 µg/ml

Tampão de lavagem SELEX

HEPES pH 7,9	10 mM
KCl	70 mM
MgCl ₂	2,5 mM
EDTA	1,0 mM
ZnCl ₂	1,0 mM
ZnSO ₄	0,01 mM
Glicerol	4,0%
DTT	1 mM

Tampão de ligação EMSA (10X)

NaH ₂ PO ₄	10 mM
Na ₂ HPO ₄	10 mM
MgCl ₂	5,0 mM
EDTA	2,5 mM
ZnCl ₂	1,0 mM
ZnSO ₄	0,1 mM
Glicerol	50,0 %
Poli (dIdC)	100 µg/ml

pH 7,6

Solução Bicarbonato de Sódio a 4,4%

A solução foi preparada sob agitação com a adição de 2 a 4 gotas de vermelho de fenol a 0,5%. Foram adicionadas pedras de gelo seco até que a solução se tornasse róseo-amarelada (pH 7,0). Em seguida, a solução foi distribuída rapidamente em alíquotas de 20 ml em tubos com tampa de rosca bem vedada. As alíquotas foram esterilizadas em autoclave a 121°C por 30 min e armazenadas a 4°C por até 2 meses.

Solução Verseno

EDTA	2 mM, pH 7,0
PBS	1 X

A esterilização foi realizada em autoclave por 20 min a 120°C.

Solução de Tripsina

Tripsina	5%
PBS	1 X

O pH foi ajustado para 7,8 a 8,0. A solução foi armazenada a 4°C durante a noite para decantação. No dia seguinte, o pH foi checado novamente e ajustado se necessário. A solução foi esterilizada por filtração através de filtro Millipore 0,22 µm, distribuída em alíquotas de 0,5 ml e

armazenada a -20°C .

V – MÉTODOS UTILIZADOS

Todos os procedimentos de digestões com enzimas de restrição, tratamento com a enzima DNA polimerase I fragmento Klenow, desfosforilações, purificação de fragmentos de DNA em gel de agarose, ligações, análises em géis de agarose, SDS-PAGE, extração de DNA plasmidial foram realizados por procedimentos padrões (Sambrook *et al.*, 1989) ou conforme instruções dos fabricantes (Biolabs, Promega e Gibco BRL). Todas as transformações bacterianas foram realizadas seguindo a metodologia descrita por Inoue *et al.* (1990).

1 – Isolamento e análise da seqüência da região genômica de 17 kb de *Trypanosoma cruzi*

O banco genômico (EMBL3) de *Trypanosoma cruzi* (clone Dm28C) (Fragoso *et al.*, 1992) foi selecionado com o cDNA da subunidade proteassomal ALPHA6 (*TcPR29A*) (Bartholomeu *et al.*, 2001) e um clone positivo foi caracterizado, contendo um fragmento de 17.303 pb localizado no cromossoma XX de *T. cruzi*. O DNA do clone positivo foi isolado e os fragmentos obtidos através de digestões com enzimas de restrição foram sub-clonados no pBluescript SK⁺ ou pUC18. Todos os sub-clones foram submetidos a sequenciamento automático e a análise de seqüência foi feita usando o software GCG da Universidade de Wisconsin. As seqüências de nucleotídeos estão disponíveis no banco de dados GenBank com os números de acesso: AF104214 e AY728266. O alinhamento das seqüências foi realizado utilizando a versão 1.82 do programa CLUSTALW (Thompson *et al.*, 1994). Um grupo de 1000 seqüências foi gerado com o programa SeqBoot; o cálculo da matriz da distância e a construção da árvore filogenética foi feita utilizando os programas ProtDist e Fitch respectivamente. O SeqBoot, o ProtDist, o Fitch e o programa de consenso para árvores fazem parte de um pacote de program chamado Phylip, cuja versão aqui utilizada foi a 3.8 (Felsenstein *et al.*, 2002).

2 – Produção da proteína TcZFP8 em *E. coli*

Linhagens de *E. coli* contendo o plasmídeo foram inoculados em frascos Erlenmeyers de 250 ml contendo 20 ml de meio LB adicionado de ampicilina a 100 µg/ml e incubados a 37°C sob agitação de 200 rpm. As culturas foram incubadas até atingirem uma absorbância de 0,8 a 600 nm. Retirou-se uma alíquota de 1 ml de cada frasco antes de iniciar a indução. Foram acrescentados IPTG para uma concentração de 1,0 mM em cada um dos frascos e retiradas alíquotas de 1 ml em intervalos pré-estabelecidos. Essas alíquotas foram centrifugadas a 12.000 rpm por 3 min. O sobrenadante foi descartado e o sedimento de células foi ressuscitado com 50 µl de tampão de lise 1X. Os extratos protéicos total foram incubados em banho a 90°C por 5 min e analisados em gel desnaturante de poliacrilamida (SDS-PAGE) (Laemmli, 1970).

3 – Purificação da proteína TcZFP8 expressa em *E. coli*

3.1 - Cromatografia de afinidade para purificação da proteína TcZFP8 fusionada a proteína Glutathione-S-Transferase (*Shistosoma japonicum*)

Para cada sistema utilizou-se 12 µl da resina glutathione sepharose 4B (Pharmacia) para a ligação GST-TcZFP8. A resina foi previamente ressuscitada em PBS 1X adicionado de 1 mM de EDTA e centrifugada a 5.000 rpm por 5 min. Após a centrifugação, as resinas foram ressuscitadas em 200 µl de PBS 1X suplementado com 1 mM de EDTA e distribuídos em dois tubos Eppendorfs contendo cada um deles 200 µl de extrato protéico de *E. coli*. Após homogeneização, os tubos foram incubados a 4°C sob agitação de 200 rpm durante 1 hora. Em seguida foram centrifugados a 5.000 rpm durante 5 min e os sobrenadantes descartados. Os sedimentos foram lavados três vezes com 200 µl de PBS 1X e finalmente ressuscitados em 40 µl de tampão de eluição. Após a centrifugação, os sobrenadantes foram transferidos para outros tubos Eppendorf e estocados a – 80°C. Uma alíquota de 10 µl da proteína eluída era analisada em gel de poliacrilamida a 15% (SDS-PAGE).

3.2 - Cromatografia de afinidade Ni-NTA (*Nickel-nitrilotriacetic acid*) para purificação da proteína TcZFP8 fusionada a *His-tag*

Os extratos protéicos de *E. coli* foram previamente alquilados conforme descrito por Yan *et al.* (1999). Utilizou-se 1 ml de resina Ni-NTA ativada com 50 mM de sulfato níquel pH 6,8 para a ligação TcZFP8-*His-tag*. A resina foi previamente ressuspensa em tampão de ressuspensão (Novagen) centrifugada a 2.000 rpm por 2 min. Após centrifugação, a resina Ni-NTA foi ressuspensa novamente no mesmo tampão e distribuída em tubos Eppendorfs contendo cada um deles 200 µl de extrato protéico alquilado. Após homogeneização, os tubos foram incubados a 4°C sob agitação de 200 rpm durante 1 hora. Em seguida, os tubos foram centrifugados a 2.000 rpm durante 2 min. Os sobrenadantes foram descartados e os sedimentos foram lavados com o tampão de lavagem (Novagen), sendo eluídas em tampão de eluição (Novagen). As proteínas foram analisadas em gel de poliacrilamida a 15% (SDS-PAGE).

4 - Produção de anticorpo anti-TcZFP8 em coelho

A proteína TcZFP8-*His-tag* purificada em Ni-NTA foi quantificada segundo método descrito por Ramagli & Rodrigues (1985). Previamente, o sangue de um coelho de 2 meses de idade foi retirado e centrifugado a 4.000 rpm durante 15 min, para coleta do soro pré-imune presente no plasma. Aliquotas do soro pré-imune foram mantidas em freezer -80°C. Em seguida, foram aplicadas 3 injeções com intervalos de 15 dias, contendo cada uma delas 100 µg de proteína purificada. A primeira injeção continha o adjuvante de Freund completo (50%), a segunda, o adjuvante de Freund incompleto (50%) e a terceira solução, salina estéril. Após 10 dias da aplicação da terceira injeção, o sangue do coelho foi retirado e o soro contendo os anticorpos foi processado como descrito anteriormente. O soro foi testado utilizando lisados das bactérias produtoras da TcZPF8-*His-tag* pela técnica de *Western blot*.

5 - Obtenção de Extratos Protéicos de *Trypanosoma cruzi*

As formas epimastigotas foram crescidas em meio LIT suplementado com 10% de soro fetal bovino inativado até atingirem fase exponencial ($2,0 \times 10^7$ /ml). Em seguida foram centrifugadas a 4.000 rpm durante 15 min. O sedimento foi lavado com PBS 1X gelado e novamente centrifugado. Para obtenção do extrato protéico total, as células foram ressuspensas em tampão de amostra 1X. As frações mitocondrial/nuclear e citoplasmática foram obtidas segundo o método descrito no item V.10.

As formas tripomastigotas e amastigotas foram obtidas através da infecção de células musculares cardíacas de camundongos (L6) crescidas em meio M199 (Gibco-BRL) - suplementado com soro fetal bovino (item 5.9.2). O número de tripomastigotas e amastigotas foi padronizado ($2,0 \times 10^8$). Em seguida foram centrifugadas a 4.000 rpm durante 15 min. Os sedimentos das respectivas células foram lavados com PBS 1X gelado e novamente centrifugados. As células foram ressuspensas em tampão de amostra 1X.

Os extratos protéicos foram analisados em gel de poliacrilamida a 15% (SDS-PAGE).

6 - Extração de RNA das três formas de *T. cruzi* e análise por *Northern blot*

O RNA total das formas epimastigota, amastigota e tripomastigota foi obtido pelo método de extração de RNA por TRIZOL segundo protocolo indicado pelo fabricante (Invitrogen). Esses RNAs foram submetidos à eletroforese em gel desnaturante de agarose (Ausubel *et al.*, 1995) e transferidos para a membrana de nylon. Para hibridação foi utilizada uma sonda de DNA de 1,8 kb contendo parte do gene *TcZFP8* (nucleotídeos 61 a 571). Os experimentos de *Northern blot* foram realizados como descrito em Sambrook *et al.* (1989).

7 - Análise de *Western Blot*

Os extratos proteicos de *E. coli* ou *T. cruzi* foram submetidos à eletroforese em gel de poliacrilamida a 15% (SDS-PAGE) e em seguida transferidos para uma membrana de nitrocelulose

através de um sistema de transferência semi-seco (BioRad), durante 1 hora. Em seguida, tampão de saturação de proteínas (tampão PBSM) foi adicionado à membrana, a qual foi incubada durante 3 horas a 4°C, sob agitação branda. A membrana foi lavada 3 vezes com PBS 1X e o anticorpo primário (monoclonal anti-His-tag ou policlonal anti-TcZFP8) foi adicionado na diluição de 1:2.000 e incubado durante 3 horas a 4°C, sob agitação branda. Em seguida, foram realizadas 3 lavagens com PBS 1X e posteriormente, o anticorpo secundário acoplado a peroxidase foi adicionado na diluição de 1:2000 e incubado durante 3 horas a 4°C, sob agitação branda. Após o período de incubação, a membrana foi lavada 3 vezes com PBS 1X e revelada com solução reveladora para peroxidase.

8 - Transfecção em *Trypanosoma cruzi*

As formas epimastigotas foram crescidas em meio LIT suplementado com 10% de soro fetal bovino inativado até atingirem fase exponencial ($2,0 \times 10^7$ células/ml). Em seguida foram centrifugadas a 4.000 rpm durante 15 min. O sedimento foi lavado com PBS 1X gelado e novamente centrifugado e ressuscitado em tampão de transfecção 1X (Ramirez *et al.*, 2000). A suspensão contendo os parasitas foi transferida para cuvetas de 2 mm (400 µl em cada) sendo adicionado 50 µg de plasmídeo em cada sistema. O sistema foi submetido a eletroporação (Bio-Rad Gene Pulser II) com 2 pulsos consecutivos de 300V e 500 µF. Em seguida os parasitas eletroporados foram transferidos para o meio LIT suplementado com 10% de soro fetal bovino inativado e incubados a 28°C. Após 12 horas, foi adicionado 100 µg/ml de Geneticin (G418)(Gibco-BRL) para selecionar os parasitas transfectados. Aliquotas de 200 µl foram retiradas após 12, 36 e 48h de transfecção, centrifugadas a 4.000 rpm durante 15 min e o sedimento foi lavado com solução salina estéril. Em seguida as células foram ressuscitadas em salina estéril suplementada com 10 % de BSA e fixadas em lâmina com metanol absoluto durante 5 minutos. As lâminas foram observadas em microscópio de fluorescência sob o comprimento de onda de 365 a 480 nm (faixa de visualização da Green fluorescent protein)

9 – Obtenção das diferentes formas de *Trypanosoma cruzi*

9.1 – Obtenção da forma epimastigotas de *Trypanosoma cruzi*

Epimastigotas das linhagens CL Brener, cepa Y e Dm 28c foram crescidos em meio LIT de modo que o volume ocupado pelas culturas fosse de aproximadamente 25 % da capacidade total do frasco. Os inóculos foram iniciados com aproximadamente 2×10^6 células/ml e incubados durante uma semana a 28°C. Ao final desse tempo, as culturas alcançavam uma densidade celular de aproximadamente 3×10^7 células/ml, sendo então, centrifugadas a 5.000 rpm durante 15 minutos. O sedimento contendo as formas epimastigostas foi utilizado para a realização dos experimentos necessários.

9.2 – Obtenção das formas tripomastigotas e amastigotas de *Trypanosoma cruzi*

As células L6 foram cultivadas em garrafas de 25 cm³ contendo meio 199 suplementado com soro fetal bovino inativado. As células foram incubadas a 37°C em estufa com 5% de CO₂ por um período necessário para obtenção de crescimento confluyente (48 - 72 horas). Após este período, o meio foi retirado, descartado e as monocamadas de células foram lavadas 3 vezes com PBS 1X e em seguida tratadas com solução verseno/tripsina. Posteriormente, 4 a 5 gotas da suspensão de células foram inoculadas em garrafas de 80 cm³ contendo 20 ml de meio 199 suplementado com soro fetal bovino e incubadas a 37°C até atingirem confluência de 10 a 20%. O meio de cultura foi retirado e novamente 20 ml de meio 199 contendo soro fetal bovino foram adicionados à cultura. Em seguida, as formas tripomastigotas de *T. cruzi* foram adicionadas na proporção de 40 parasitas para cada célula hospedeira L6. As células L6 infectadas foram incubadas a 37°C durante 15 – 30 dias para obtenção das formas tripomastigotas e amastigotas.

A maior concentração de tripomastigotas estava na fase inicial da infecção e posteriormente, a forma amastigota. As formas foram separadas como descrito a seguir: o meio de cultura foi coletado, centrifugado a 2.500 rpm durante 10 minutos e incubado a 37°C durante 3 – 12 horas para migração das formas tripomastigotas para o sobrenadante. A parte superior do sobrenadante

(aproximadamente 75%) foi transferida para outro tubo e centrifugado novamente para obtenção de um sobrenadante contendo as formas tripomastigotas. As formas amastigotas foram separadas dos debrís celulares (L6) pela centrifugação da parte inferior previamente reservada durante 5 minutos a 500 rpm.

10 – Obtenção das frações mitocondrial/nuclear e citoplasmática da forma epimastigota de *Trypanosoma cruzi*

As formas epimastigotas de *T. cruzi* (cepas Dm28c e Y), obtidas conforme item V.9.1, foram lavadas três vezes com PBS 1X. Em seguida, foi adicionado tampão TENM2 1X contendo 1% de Triton X-100 e a solução foi incubada por 5 min no gelo. Após o período de incubação, a suspensão celular foi homogeneizada com aproximadamente 20 movimentos em um homogeneizador de vidro (*dounce*). A lise celular foi acompanhada ao microscópio óptico e a osmolaridade foi restaurada com a adição de sacarose a 2 M em TENM2 1 X para uma concentração final de 0,25 M. O lisado celular foi então centrifugado a 5.000 rpm (Rotor R20A2 46, Hitachi) por 10 min a 4°C. O sobrenadante (fração citoplasmática) e o sedimento (fração nuclear/mitocondrial) foram coletados estocados a -80°C.

11 – Imunofluorescência indireta

As formas epimastigotas de *T. cruzi* cepa Y obtidas conforme item V.9.1 foram lavadas três vezes com PBS 1X e ressuspensas em PBS 1X contendo 1% de BSA. Em seguida, os parasitas foram fixados em lâmina com metanol 100% durante 15 minutos à temperatura ambiente. Após secagem, os parasitas foram impermeabilizados com uma solução de PBS 1X suplementada com 0,2% de Triton X-100 durante 15 minutos. Após este período, as lâminas foram saturadas com PBS 1X contendo 5% de leite desnatado durante 4 horas a 4°C. Posteriormente, as lâminas foram lavadas três vezes com PBS 1X e incubadas com anticorpo policlonal anti-TcZFP8 na diluição 1:50 em PBS 1X contendo 1% de leite desnatado, durante 4 horas à temperatura ambiente. Em seguida, as lâminas foram lavadas três vezes com PBS 1X e incubadas com anticorpo secundário anti-coelho

conjugado com FITC, na diluição 1:50 em PBS 1X contendo 1% de leite desnatado, durante 4 horas à temperatura ambiente. Após este período, as lâminas foram lavadas com PBS 1X e incubadas com DAPI (1 µg/ml) em PBS 1X durante 20 minutos e novamente lavadas com PBS 1X. As lâminas foram observadas em microscópio Olympus BX60 usando os filtros apropriados.

12 – Evolução Sistemática de Ligantes por Enriquecimento Exponencial (SELEX) - adaptado de Tsai & Reed, 1998

O SELEX (*systematic evolution of ligands by exponential enrichments*) (Klug & Famulok, 1994) foi realizado do seguinte modo: inicialmente, três oligonucleotídeos foram sintetizados quimicamente: PR102 - 5'TAATACGACTCACTATAGGGAGGATCCGCGCAGC(N)₃₀CCAATGGTACCAGCAAC3'; PR103 - 5'TAATACGACTCACTATAGGGAGGATCCGCGCAGC 3' e PRQ2.7 - 5'GTTGCTGGTACCATTGG 3'. O oligonucleotídeo PR102 contém uma região randômica de 30 pb, um sítio de restrição *Bam*HI (5'GGATCC3') e uma seqüência que codifica o promotor T7 (5'TAATACGACTCACTATA3'). Os oligonucleotídeos PR103 e PRQ2.7 foram desenhados para amplificação do oligonucleotídeo 102, anelando na região 5' e 3' respectivamente.

O oligonucleotídeo PR102 (utilizado como fita molde) foi amplificado por PCR utilizando os oligonucleotídeos PR103 e PRQ2.7 como primers. Foram utilizadas as seguintes condições para a amplificação por PCR: um ciclo inicial de desnaturação a 94°C por 1 min, trinta ciclos de desnaturação a 94°C por 15 segundos, anelamento a 50°C por 30 segundos e extensão a 72°C por 30 segundos e um ciclo final de extensão a 72°C por 1 minuto. O produto da PCR foi analisado em gel de agarose a 2%. Em seguida, o produto da PCR, DNA fita dupla ou DNA fita simples (desnaturado a 95°C durante 5 minutos) foram adicionados nos seguintes sistemas: (1) proteína GST ligada à resina glutationa-sepharose 4b e (2) proteína de fusão GST-TcZFP8 ligada à resina glutationa-sepharose 4b. A cada sistema foi adicionado 25 µl de tampão de ligação SELEX. Os sistemas foram incubados durante 30 min à temperatura ambiente. Em seguida, os sistemas foram lavados 5 vezes com tampão de lavagem SELEX (5 volumes) e centrifugados a 3.000 rpm durante 2 min. O sobrenadante foi descartado e o sedimento foi ressuscitado em 20 µl de tampão para

PCR da GIBCO (1X), completando 1 ciclo de SELEX. Um novo ciclo de SELEX foi iniciado, utilizando 10 µl do produto obtido anteriormente como template para a PCR. Após 6 ciclos de SELEX, o produto da PCR foi ligado ao pGEMT-easy. O sistema de ligação foi utilizado para transformar células de *E. coli* DH5α e os transformantes foram analisados pelo perfil de restrição pela digestão com a enzima *Bam*H I. Os plasmídeos obtidos foram seqüenciados (MegaBace 1000 – Pharmacia) pelo método *Dye Terminator*.

13 – Ensaio de ligação proteína-DNA (*gel shift*)

Os plasmídeos 33, 36 e 38 oriundos do SELEX foram digeridos com a enzima de restrição *Eco*R I e os fragmentos foram separados em gel de poliacrilamida a 10%. Os fragmentos de DNA de 80 pb foram eluídos e marcados com γ -³²P-ATP e a incorporação quantificada segundo método descrito em Sambrook *et al.* (1989). Após a marcação, os fragmentos foram utilizados como sondas para o ensaio de ligação com a proteína de fusão TcZFP8-GST e/ou GST. Em cada sistema foi adicionado de 4.000 a 6.000 cpm da sonda obtida. Competidores específicos em maior concentração (sondas obtidas sem marcação - 100 vezes) e/ou competidor inespecífico (poli dIdC.dIdC - 100 vezes) foram também utilizados como controles.

A ligação DNA-proteína foi realizada do seguinte modo: a um tubo Eppendorf mantido no gelo foram adicionados nesta ordem: (1) tampão de ligação *gel shift*, (2) 200 a 400 ng da proteína, (3) polidIdC e (4) DNA marcado desnaturado a 90°C durante 5 minutos. Após a adição, os tubos foram homogeneizados e incubados durante 20 minutos à temperatura ambiente.

Em seguida as amostras foram submetidas à eletroforese em gel poliacrilamida não desnaturante a 5% em câmara fria durante 4 horas a 150 V. Após a corrida, o gel foi transferido para um papel de filtro, secado a 80°C e exposto em filmes de raios X. As autoradiografias foram reveladas conforme indicado pelo fabricante.

RESULTADOS

I - Identificação de um cluster contendo três genes *in tandem* codificando para proteínas zinc finger CCHC em *T. cruzi*

Em trabalho anterior, Bartolomeu *et al.* (2001) isolou e caracterizou o gene para a subunidade ALPHA6 proteassomal, designado como *TcPR29A* e mostrou que esse gene é localizado na banda XX de *T. cruzi*. Um fragmento genômico de 17.303 pb foi isolado e o gene *TcPR29A* localizado no final dessa região (Figura 14A). Após subclonagens e análise de seqüenciamento dessa região, foi evidenciada a presença de outras cinco fases abertas de leitura (ORF), com a mesma direção de transcrição. A primeira, uma região codante de 788 pb (*TcBETA5*), foi encontrada na outra extremidade deste fragmento e a análise em bancos de dados mostrou ser um gene incompleto codificando a subunidade BETA5 proteassomal por homologia (Figura 14A). A montante do gene *TcPR29A*, foi encontrada uma fase aberta de leitura de 5.103 pb (*TcCLAH*) codificando para a cadeia pesada de clatrina (Assis *et al.*, 1998). Em uma região de 8,6 kb, entre os genes *TcBETA5* e *TcCLAH* foram identificadas outras três fases abertas de leituras para proteínas contendo cinco, sete e oito domínios *zinc finger* do tipo CX₂CX₄HX₄C, respectivamente.

A pesquisa no banco de dados da proteína predita com cinco domínios *zinc finger* mostrou que essa proteína de 14,32 kDa corresponde à proteína PDZ5 de *T. cruzi* (Coelho *et al.*, 2003). A proteína PDZ5 foi descrita de um clone de cDNA da cepa CL Brener e o gene correspondente foi localizado na banda XX usando eletroforese em campo pulsante e *Southern Blot* (Coelho *et al.*, 2003). O gene correspondente na cepa Dm28c de *T. cruzi* foi aqui identificado, com a localização neste fragmento de 17 kb da banda XX a jusante do gene *TcBETA5*. Em *T. brucei* a mesma posição do gene ZFP em relação ao gene *BETA5* proteassomal foi observado na sua banda X (Figura 14A). Desta forma, PDZ5 e ZFP são fortes candidatas a serem as proteínas de *T. cruzi* e *T. brucei* homólogas à proteína UMSBP de *C. fasciculata*, uma proteína que se liga especificamente a fita simples na região de 12 nucleotídeos chamada da seqüência universal de minicirculos (UMS)

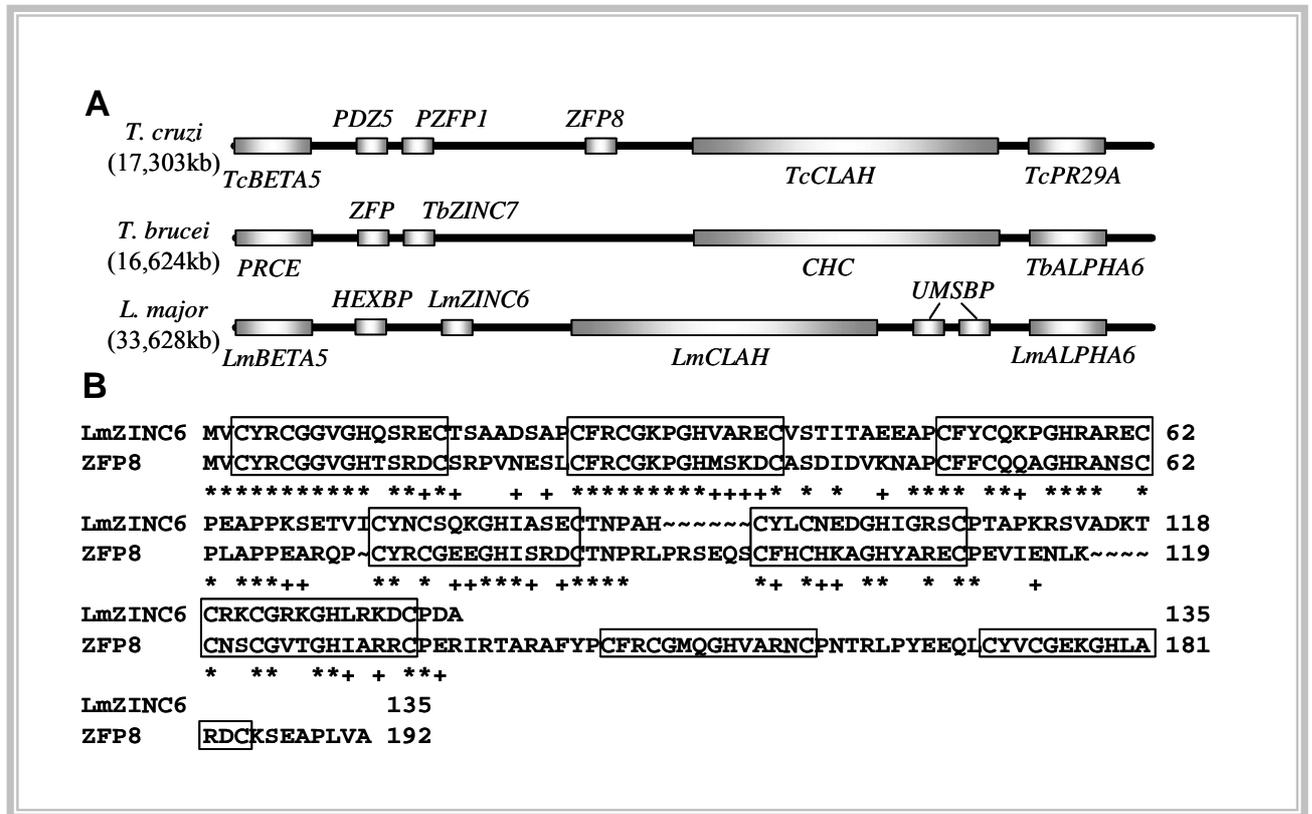


Figura 14 – Comparação das regiões genômicas homólogas de *Trypanosoma cruzi*, *Trypanosoma brucei* e *Leishmania major* e caracterização de uma família de genes codificando proteínas zinc finger CX₂CX₄HX₄C. (A) Esquema mostrando a região de 17 kb presente na banda XX de *Trypanosoma cruzi*, onde os retângulos representam genes na ordem seguinte: *TcBETA5* - gene da subunidade beta 5 proteassomal; *PDZ5*, *PZFP1* e *ZFP8* - genes codificando cinco, sete e oito domínios zinc finger, respectivamente; *TcCLAH* – gene da cadeia pesada de clatrina; *TcPR29A* – gene para a subunidade alfa proteassomal. A mesma região entre os genes para a subunidade BETA5 proteassomal e ALPHA6 são mostrados em *T. brucei* (banda 10, banco de dados TIGR) e *L. major* (contig 36.1 do banco de dados Sanger). (B) Alinhamento das proteínas ZFP8 e LmZINC6. Asteriscos indicam os aminoácidos idênticos e os sinais de “+” indicam os aminoácidos similares. Os domínios *zinc finger* estão representados em caixas fechadas.

localizada em minicírculos no cinetoplasto (Tzfati *et al.*, 1992). Em *L. major*, foram encontradas duas cópias idênticas *in tandem* do gene homólogo UMSBP de *C. fasciculata*, a montante do gene *LmCLAH* (Figura 14A), que foram denominados de genes *LmUMSBP*. As proteínas preditas desses genes possuem 80% de identidade e 90% de similaridade com a UMSBP de *C. fasciculata* (números de acesso do GenBank: AAC3814 e A54598).

A análise em banco de dados da proteína predita com sete domínios *zinc finger* (193 aa e 20,42 kDa) mostrou que essa proteína corresponde à PZFP1 de *T. cruzi* (Espinosa *et al.*, 2003). Esta proteína PZFP1 foi localizada principalmente no citoplasma e se liga a fitas simples de DNA ou RNA carreando seqüências de reconhecimento conhecidas de outras proteínas $CX_2CX_4HX_4C$, tais como o UMS de *C. fasciculata* e a seqüência hexâmera da GP63 de *L. major* (Espinosa *et al.*, 2003). Em *T. brucei* foi identificado o gene homólogo a *PZFP1* denominado *TbZINC7* na mesma posição em relação ao gene *ZFP* (Figura 14A). Análise do alinhamento entre as proteínas preditas PZFP1 e *TcZINC7* mostrou 67% de identidade e 70% de similaridade. Essas proteínas são também similares a uma proteína de ligação a ácido nucléico de *Trypanosoma equiperdum* (número de acesso no GenBank: AAB47542), apresentando 49% de identidade e 53% de similaridade com PZFP1 e 74% de identidade e 75% de similaridade com *TbZINC7*. Essas proteínas ainda não têm funções conhecidas.

O alinhamento de *PDZ5* e *PZFP1* mostrou uma identidade de 95% em 375 nucleotídeos e 94% em 125 aa entre os 5 domínios de *PDZ5* e os últimos 5 domínios de *PZFP1*. Isso sugere que esses genes podem ser derivados da duplicação de um mesmo gene. O mesmo foi observado em *T. brucei* quando *ZFP* e *TbZINC7* foram comparados.

A proteína predita com oito domínios *zinc finger* revelou ser um novo gene denominado *TcZFP8* contendo uma fase aberta de leitura codificando uma proteína de 192 aa. A análise em banco de dados mostrou a presença desse gene também em cepa CL Brener de *T. cruzi*, onde três genes com menor diferença foram observados. Duas seqüências de genes obtidas do banco de dados TIGR, números de acesso: 7413 e 5774, e um do GenBank: AF204398. O alinhamento da proteína

TcZFP8 com outras proteínas conhecidas CCHC de tripanossomatídeos (TcPDZ5, TcZFP1, TbZFP, TbZINC7 e LmUMSBP) revelou que a similaridade entre elas é limitada às cisteínas e histidinas constituintes do domínio *zinc finger* CX₂CX₄HX₄C. O alinhamento dos dados mostra que a identidade é cerca de 27% para 34% e a similaridade é cerca de 40% a 49%, confirmando que TcZFP8 é uma nova proteína de *T. cruzi*.

Nenhum gene similar a *TcZFP8* foi encontrado na mesma região (Figura 14A) ou no genoma total de *T. brucei* (TIGR e no banco de dados Sanger). Em *Leishmania major*, a presença do gene *HEXBP* (Webb *et al.*, 1993) a montante do gene BETA5 proteassomal foi observado (Figura 14A). O gene candidato a ser o homólogo de *TcZFP8* foi identificado nesta região e denominado *LmZINC6* (Figura 14A), mas a proteína predita contém somente seis domínios zinc finger ao invés de oito. A Figura 14B mostra o alinhamento de TcZFP8 com LmZINC6. Observa-se que essas proteínas têm um alto grau de identidade em seus três primeiros domínios.

Com o objetivo de se avaliar a relação evolucionária entre essas proteínas *zinc finger*, realizamos uma análise filogenética. Como mostrado na Figura 15, três grupos distintos monofiléticos puderam ser observados. A proteína TcZFP8 é mais próxima a LmZINC6 do que as outras proteínas, o que sugere que ambas as proteínas possam ter mesma função celular em *L. major* e *T. cruzi*.

Nenhuma descrição sobre o gene/proteína TcZFP8 foi encontrada na literatura, além dos genes homólogos para a cepa CL Brener encontrados nos bancos de dados. Desse modo, nós prosseguimos na caracterização dessa proteína de *Trypanosoma cruzi*.

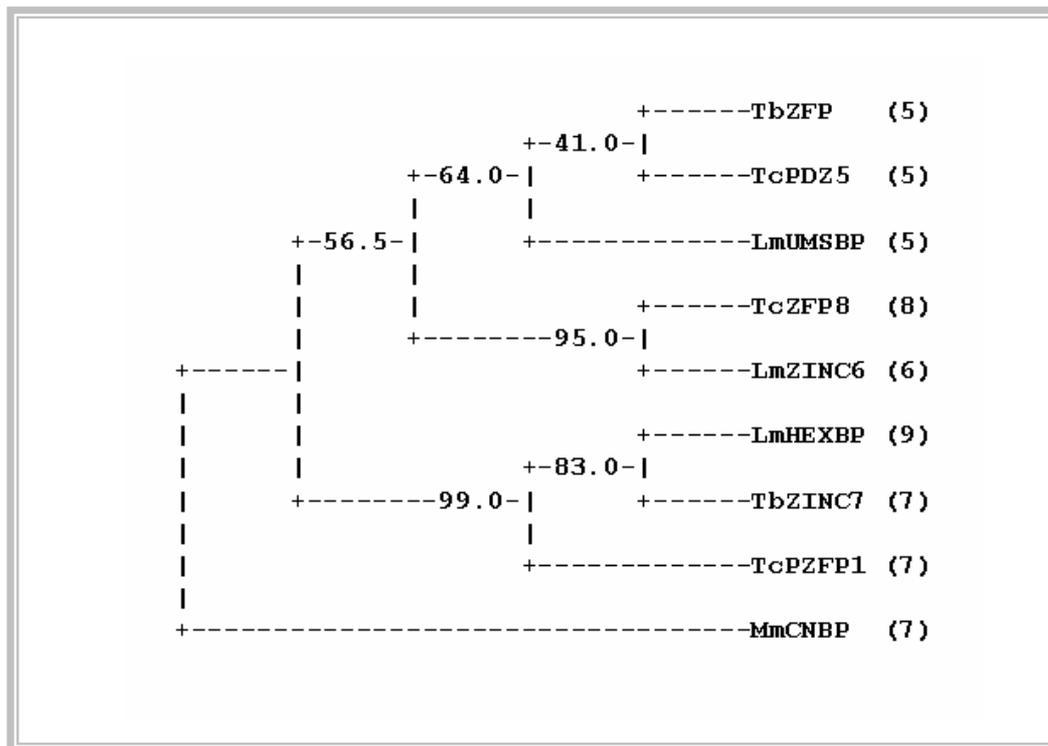


Figura 15 - Árvore filogenética descrevendo o relacionamento evolucionário entre as proteínas *zinc finger*. Os números nos parênteses representam o número de motivos *zinc finger* nas proteínas. MmCNBP (número de acesso no GenBank: P53996): proteína zinc finger de *Mus musculus*.

II – Análise do gene *TcZFP8* e da sua proteína predita

Uma análise individual da fase de leitura do gene *TcZFP8* está mostrada na Figura 16, onde podemos observar que este gene apresenta uma fase de leitura de 579 pb. Após a tradução desta fase de leitura obtivemos, então, a proteína predita de 193 aminoácidos, com massa molecular estimada de 21.382 Da. Verificou-se uma grande porcentagem do aminoácido cisteína: 11,5% da massa desta proteína. O alinhamento desta proteína internamente nos seus resíduos de aminoácidos permitiu verificar a presença dos oito motivos *zinc finger* do tipo $CX_2CX_4HX_4C$ (Figura 17).

Com o objetivo de prosseguirmos com os estudos desse gene identificado *TcZFP8*, está mostrado na Figura 18 o esquema geral contendo as etapas do trabalho a ser desenvolvido visando principalmente à expressão heteróloga da fase de leitura do gene *TcZFP8* na bactéria *Escherichia coli*, a caracterização e a localização da proteína *TcZFP8* em células de *T. cruzi*. Deste modo, a região já mencionada (na banda XX) foi subclonada e um fragmento de 4,5 kb correspondente a esta parte desta região foi subclonado no plasmídeo pUC18 a partir do DNA do fago recombinante contendo a região de 17 kb, obtendo-se assim o plasmídeo recombinante pU5.4A. A partir deste plasmídeo, um fragmento de 1,8 kb contendo a fase de leitura do gene *TcZFP8* foi utilizado nas subclonagens posteriores para a expressão deste gene em *E. coli*.

TcZFP8

```

          1/1                                31/11
ATG GTG TGT TAT CGT TGT GGC GGC GTG GGT CAT ACA AGC CGC GAC TGC TCA AGG CCT GTC
M V C Y R C G G V G H T S R D C S R P V
61/21
AAT GAA AGC CTT TGT TTC CGC TGC GGC AAA CCT GGC CAC ATG AGT AAG GAC TGC GCC ACT
N E S L C F R C G K P G H M S K D C A S
121/41
GAC ATT GAC GTA AAA AAT GCG CCC TGC TTT TTT TGC CAG CAG GCC GGC CAC AGG GCC AAC
D I D V K N A P C F F C Q Q A G H R A N
181/61
AGC TGC CCG CTC GCA CCG CCG GAA GCA CCG CAG CCG TGC TAC AGG TGT GGA GAG GAG GGC
S C P L A P P E A R Q P C Y R C G E E G
241/81
CAC ATC TCC CGC GAC TGC ACG AAT CCC CGA CTG CCA CGT TCA GAG CAG TCA TGC TTT CAC
H I S R D C T N P R L P R S E Q S C F H
301/101
TGT CAC AAG GCA GGC CAC TAC GCC CGC GAG TGC CCG GAA GTG ATT GAA AAC CTG AAG TGC
C H K A G H Y A R E C P E V I E N L K C
361/121
AAC AGC TGC GGC GTG ACT GGC CAC ATT CCA AGC CGG TGC CCG GAA CCG ATA AGG ACG GCG
N S C G V T G H I A R R C P E R I R T A
421/141
AGA GCG TTT TAT CCG TGC TTT CGC TGC GGC ATG CAG GGC CAC GTT GCT CGC AAC TGC CCG
R A F Y P C F R C G M Q G H V A R N C P
481/161
AAC ACG CGT TTG CCA TAT GAG GAG CAG CTT TGT TAT GTC TGC GGC GAG AAG GGC CAC CTG
N T R L P Y E E Q L C Y V C G E K G H L
541/181
GCA CCG GAC TGC AAG TCC GAG GCC CCG CTA GTC GCA TAA
A R D C K S E A P L V A *
```

Figura 16 - Sequência nucleotídica do gene *TcZFP8*, com a proteína predita, identificada na região da banda XX de *T. cruzi*.

<u>TcZFP8</u>	
Motivos:	
1	CYRCGGVGHTSRDC
2	CFRCGKPGHMSKDC
3	CFFCQQAGHRANSC
4	CYRCGEEGHISRDC
5	CFHCHKAGHYAREC
6	CNSCGVTGHIARRC
7	CFRCGMQGHVARNC
8	CYVCGEKGLARDC
consenso:	CxxCxxxGHxxxxC

Figura 17: Alinhamento interno dos motivos CCHC da proteína predita TcZFP8.

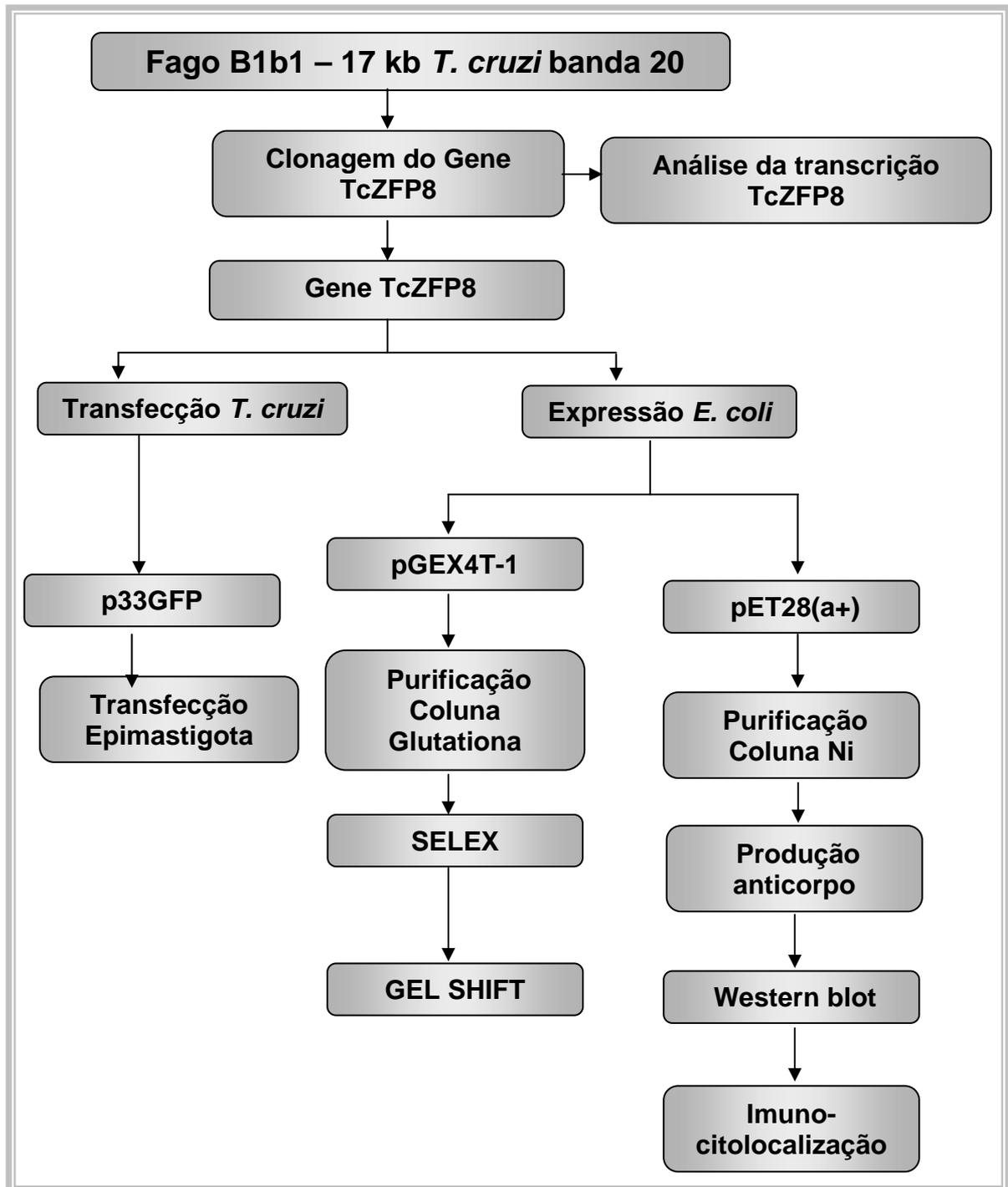


Figura 18: Esquema geral utilizado para estudar o gene *TcZFP8* de *Trypanosoma cruzi*, localizado na região genômica entre os genes que codificam a subunidade beta-proteassomal e a cadeia pesada de clatrina.

III - Análise da expressão do gene *TcZFP8* em células de *T. cruzi* por Northern blot

A partir de células das três formas de *T. cruzi*, epimastigotas, tripomastigotas e amastigotas, RNA total foi purificado e experimento de *Northern blot* foi realizado, utilizando como sonda molecular o fragmento de 1,8 kb contendo o gene *TcZFP8*. A Figura 19 mostra os resultados desta hibridização, onde podemos verificar uma marcação nas três amostras correspondente ao sinal de maior massa molecular, confirmando que este gene apresenta uma região 3' não traduzida de pelo menos 2,5 kb, sugerido na análise do fragmento genômico. Os sinais de menores massas moleculares podem ser explicados por uma hibridização inespecífica com os outros genes de proteínas zinc finger ou serem produtos de degradação do mRNA do *TcZFP8*. Além disto, apesar da menor concentração do RNA na amostra da forma tripomastigota, podemos verificar que o mRNA de *TcZFP8* está presente nas três formas do parasita, sugerindo que esta proteína possa ser expressa nas três formas de *T. cruzi* (amastigota, epimastigota e tripomastigota).

IV – Expressão do gene *TcZFP8* como proteína de fusão com a glutathione S transferase de *Shistosoma japonicum* e a purificação da proteína recombinante

Com o objetivo de verificarmos se a proteína *TcZFP8* apresenta a capacidade de interação com ácidos nucleicos, foi realizada a clonagem do gene *TcZFP8* no vetor de expressão pGEX-4T1, visando a sua expressão como proteína de fusão com a glutathione S transferase. Para tal, o fragmento de 1,8 kb contendo o gene *TcZFP8* foi subclonado no sítio para a enzima *EcoRI* do plasmídeo pGEX-4T1 entrando em fase com o gene para a GST. Deste modo, o plasmídeo recombinante pEXZ foi obtido (Figura 20). Os plasmídeos recombinantes obtidos estão exemplificados na Figura 21, onde estão mostrados o perfil de restrição resultante da digestão com a enzima de restrição *EcoRI*, liberando uma banda de 1,8 kb relativo ao fragmento clonado contendo o gene *TcZFP8*, e também das digestões com *BamHI* e *PstI*, mostrando a inserção na orientação correta do fragmento no vetor de expressão (uma vez que *PstI* está presente no início do gene *TcZFP8*).

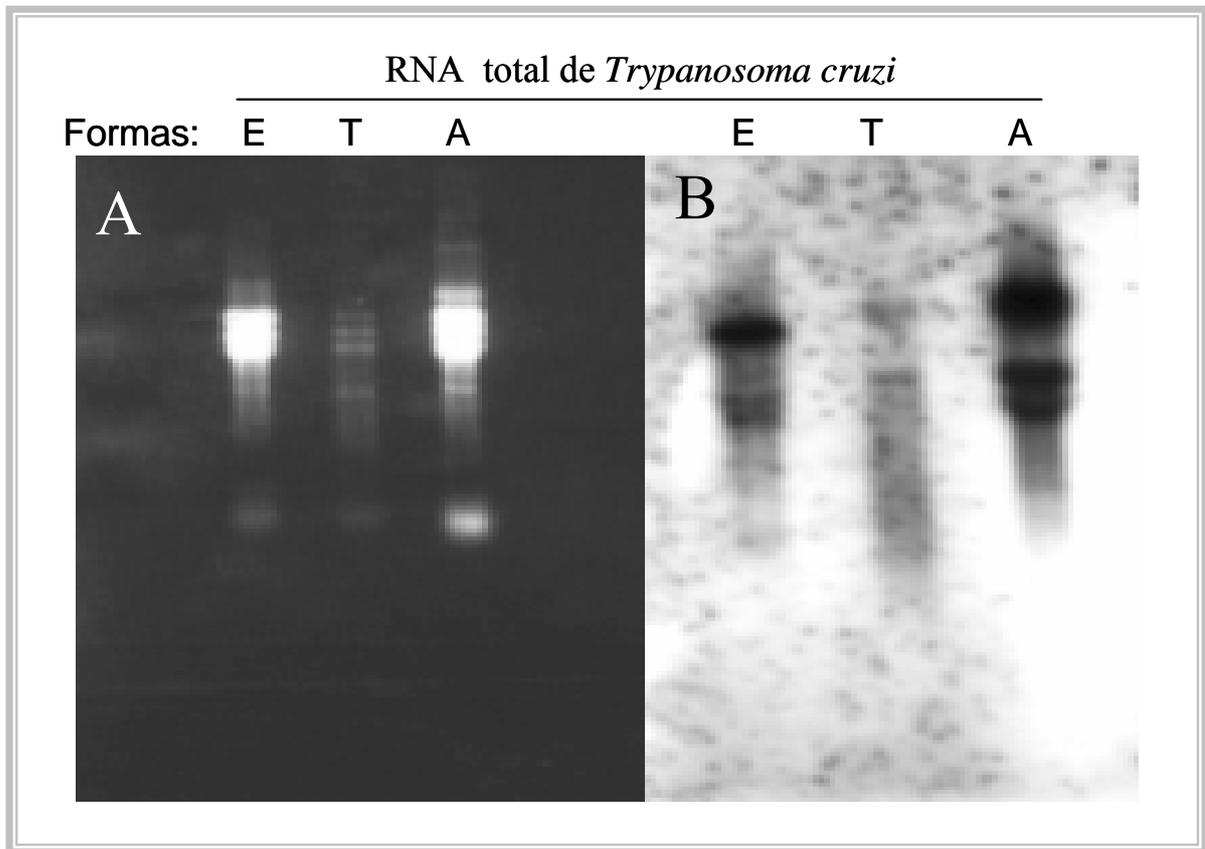


Figura 19 - Análise por *Northern-Blot* de RNA total de diferentes formas de *Trypanosoma cruzi*. Painel A - gel desnaturante de agarose com os RNAs de *T. cruzi*. Painel B – *Northern blot* – sonda: fragmento genômico do pBC1 (1,8 Kb), contendo parte do gene *TcZFP8* (7 motivos CCHC) e sua região 3' UTR. E = epimastigotas, T = tripomastigotas e A = amastigostas.

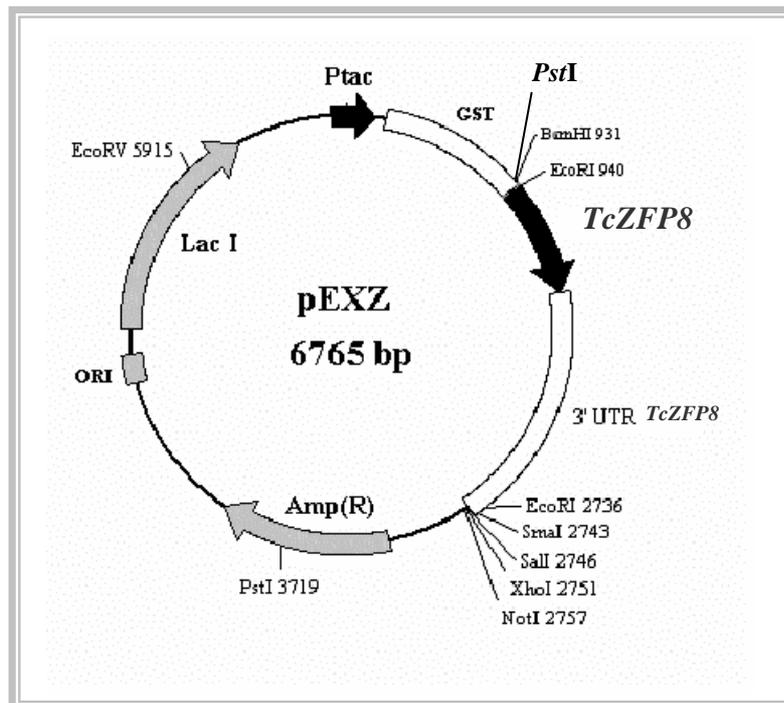


Figura 20: Mapa genético do plasmídeo pEXZ, utilizado para a expressão da proteína TcZFP8 fusionada à proteína GST em *E. coli*.

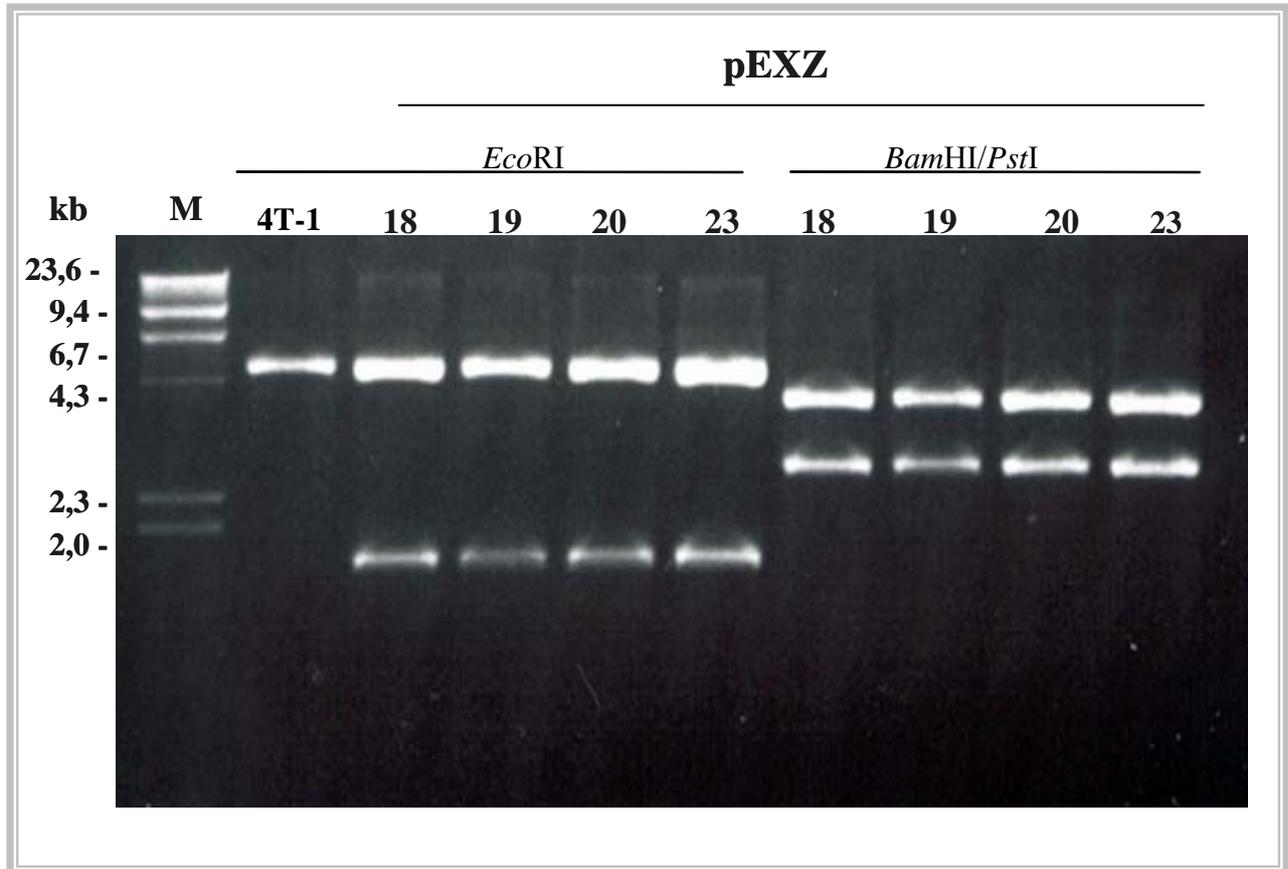


Figura 21: Análise em gel de agarose a 1% das digestões dos plasmídeos recombinantes pEXZ18, 19, 20 e 23 (plasmídeo pGEX4T-1 + gene tcZFP8) com as enzimas *EcoRI* e *BamHI/Pst I*. M – marcador de massa molecular; 4T-1 – plasmídeo pGEX4T-1 digerido com *EcoRI*.

A partir de culturas de *E. coli* contendo esses plasmídeos recombinantes pEXZ, foram realizados experimentos de indução da expressão do gene de fusão *GST-TcZFP8* em diferentes tempos e o perfil protéico total das culturas foi analisado em gel de poliacrilamida desnaturante. A Figura 22 mostra esses resultados onde podemos verificar a expressão da proteína de fusão GST-TcZFP8 apresentando aproximadamente 46 kDa nas amostras induzidas. Como controle, foi utilizado cultura contendo o plasmídeo pGEX-4T1 sem o gene *TcZFP8* clonado, onde observamos a indução apenas da GST (28 kDa) sem fusão com TcZFP8.

As proteínas GST-TcZFP8 e GST oriundas da expressão dos clones denominados pEXZ18, pEXZ23 e pGEX4T-1, respectivamente, foram purificadas por cromatografia de afinidade em glutathiona sepharose e analisadas em gel de poliacrilamida desnaturante (Figura 23), para a sua posterior utilização em ensaios de ligação com ácidos nucleicos.

V - Análise da interação da TcZFP8 com ácidos nucleicos por SELEX e *gel shift*

O experimento de SELEX foi realizado para determinarmos a região de ligação da proteína TcZFP8 nos ácidos nucleicos. O desenho dos oligonucleotídeos utilizados para tal estão esquematizados na Figura 24. Devido à possibilidade da proteína TcZFP8 poder utilizar um RNA como alvo de sua ligação, o oligonucleotídeo PR102 foi sintetizado contendo a seqüência do promotor T7, para uma provável necessidade de uma etapa de transcrição *in vitro*. Além desse promotor, este oligonucleotídeo contém 30 bases degeneradas para a seleção da seqüência de ligação no ácido nucleico da TcZFP8.

As etapas do experimento de SELEX estão mostradas na Figura 25. Esse experimento constitui de ciclos de ligação da proteína com o ácido nucleico, purificação da proteína e amplificação por PCR do fragmento ligado à proteína. Desde modo, ocorre o enriquecimento da seqüência específica de ligação do ácido nucleico com a proteína. Assim, o experimento de SELEX foi realizado utilizando as proteínas purificadas GST-TcZFP8 e GST, esta última como controle negativo, e o conjunto de oligos da Figura 24. O experimento foi realizado com fita simples de

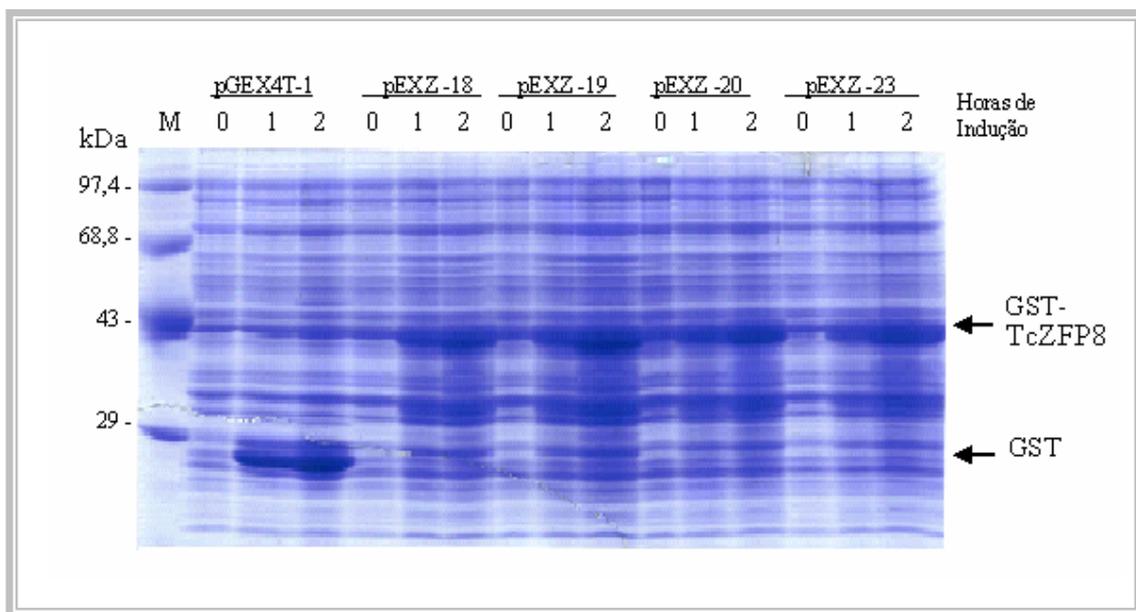


Figura 22: Análise em gel desnaturante de poliacrilamida da expressão de TcZFP8 Fusionada a GST, antes e depois da indução. As setas indicam as proteínas GST-TcZPP8 e GST induzidas.

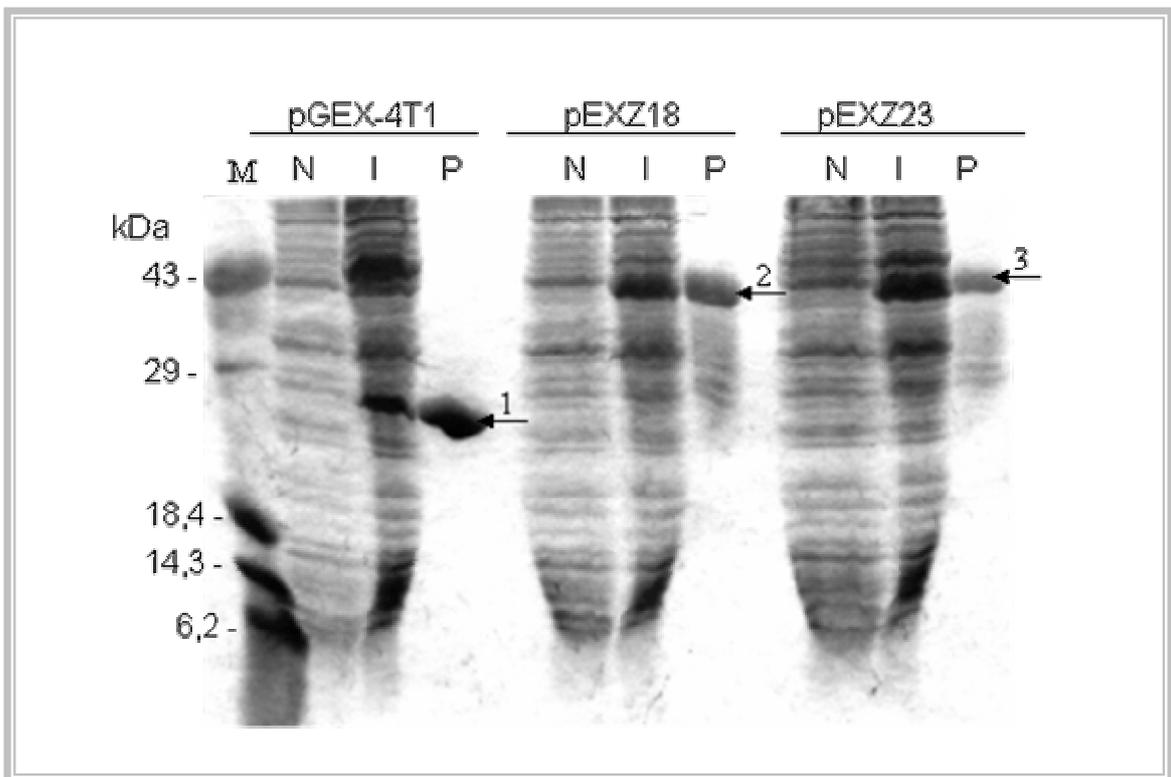


Figura 23: Purificação de GST e GST-TcZFP8 por cromatografia de afinidade. A seta 1 indica a purificação da proteína GST(pGEX-4T1). A seta 2 indica a purificação da proteína GST-ZFP8 (pEXZ18) e a seta 3 indica a purificação da proteína GST-TcZFP8 (pEXZ23). N- Extrato protéico não induzido, I- Extrato protéico induzido e P- Proteína purificada.

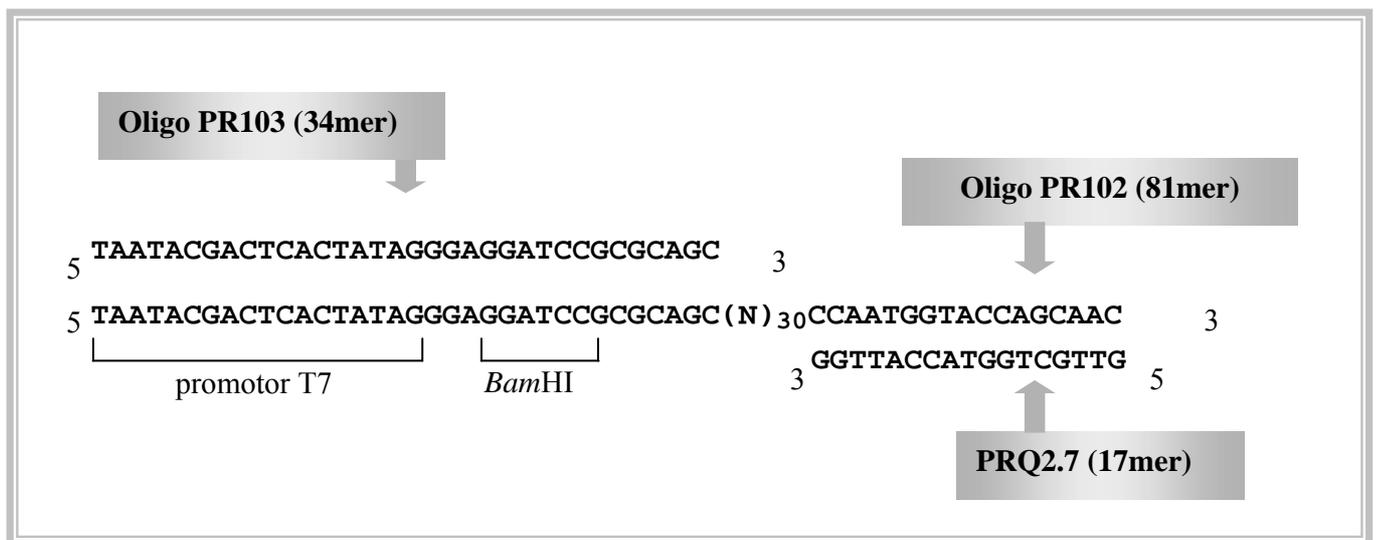


Figura 24: Seqüência dos oligonucleotídeos utilizados nos experimentos de SELEX.

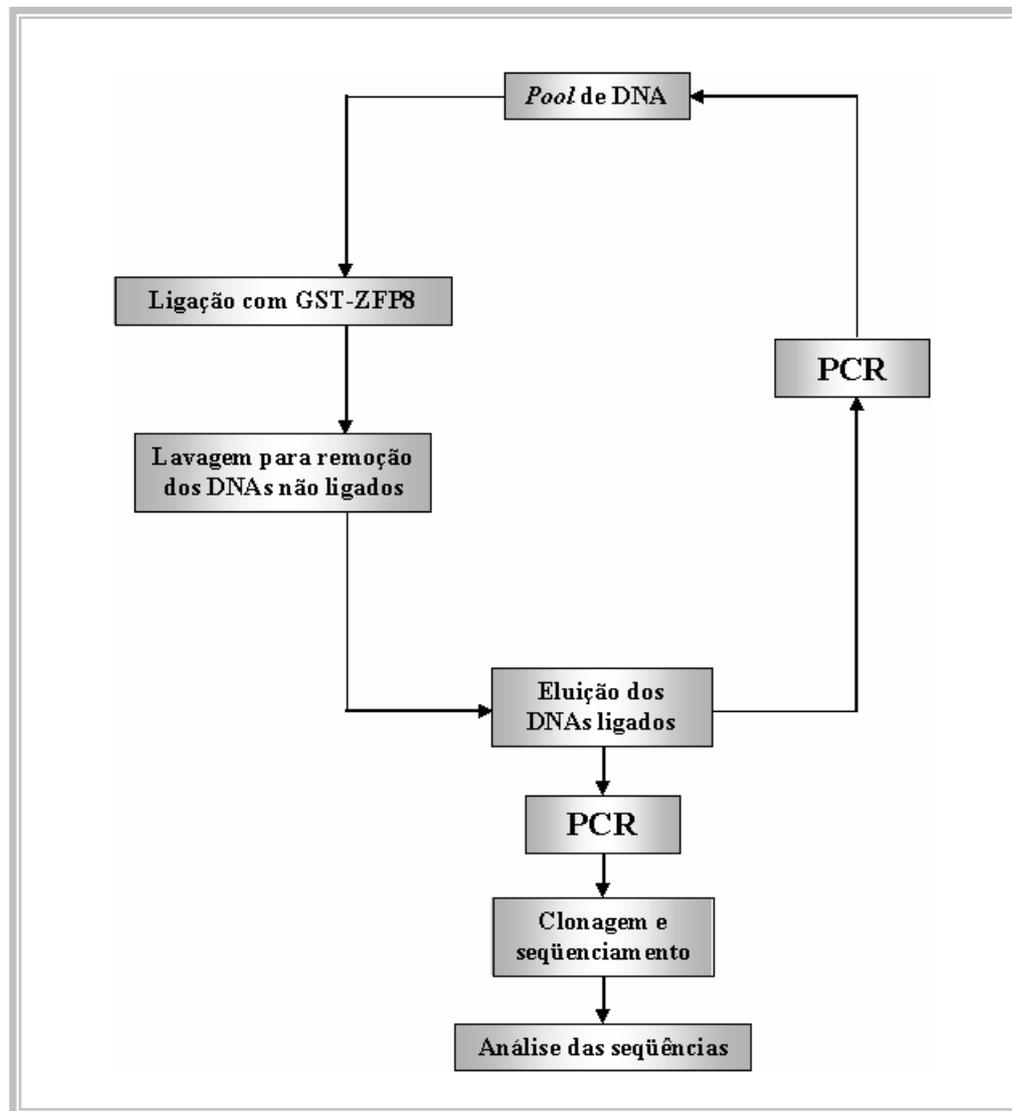


Figura 25: Estratégia do experimento SELEX para a seleção da seqüência nucleotídica de interação da proteína TcZFP8 (adaptado de Famulok & Klug, 1984).

DNA, sendo que o cada ciclo, após o PCR, uma etapa de desnaturação era realizada antes da etapa de ligação às proteínas. Após cinco e seis ciclos, o conjunto de fragmentos amplicados foram clonados no vetor para fragmentos de PCR pGEM-T *easy* para a individualização das seqüências e seu seqüenciamento.

Os resultados do seqüenciamento e do alinhamento das seqüências, conforme mostrado na Figura 26, confirmaram que TcZFP8 se liga a DNA fita simples e mostraram quatro populações distintas de seqüências consensuais (Figura 26A), sugerindo que o sítio de ligação não é restrito a uma seqüência específica, porém todas as populações mostraram regiões ricas em C e/ou G.

Com o objetivo de confirmar a especificidade de ligação da proteína TcZFP8 às seqüências nucleotídicas obtidas no experimento de SELEX, foi realizado experimento de *gel shift*. A seqüência nucleotídica do clone 33 selecionada do experimento SELEX foi utilizada como sonda para o ensaio de interação com a proteína recombinante GST-TcZFP8 e também com a proteína GST (controle negativo). A Figura 26B mostra o resultado do ensaio de ligação onde podemos observar a formação de complexo apenas no sistema da proteína GST ZFP8. Esse resultado confirma que a proteína TcZFP8 foi capaz de se ligar à seqüência do clone 33.

VI - Citolocalização da proteína TcZFP em células de *T. cruzi* por intermédio da fusão com a proteína verde fluorescente

Para determinar a citolocalização da proteína TcZFP8 em células de *T. cruzi* foram construídos vetores para transfecção onde a proteína é expressa como fusão com a proteína fluorescente verde (GFP) a partir do vetor p33 (Ramirez *et al.*, 2000). A Figura 27 mostra o mapa genético do vetor construído denominado p33GZ (p33 contendo o gene GFP fusionado ao fragmento de 1,8 kb contendo o gene *TcZFP8*). O perfil de restrição dos plasmídeos recombinantes p33GZ liberando os fragmentos de 1,8 kb contendo o gene *TcZFP8* e 0,7 kb relativo ao gene GFP na digestão *EcoRI/SalI* são mostrados na Figura 28, assim como as digestões dos plasmídeos p33 e

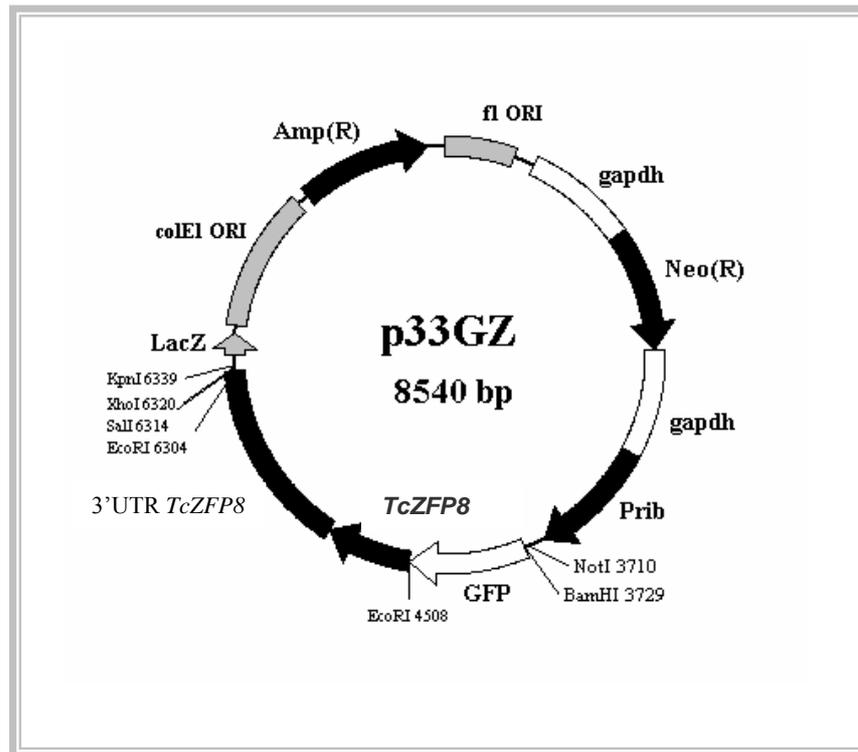


Figura 27: Mapa genético do plasmídeo p33GZ. Plasmídeo utilizado para expressão da proteína TcZFP8 fusionada a proteína GFP em *T. cruzi*.

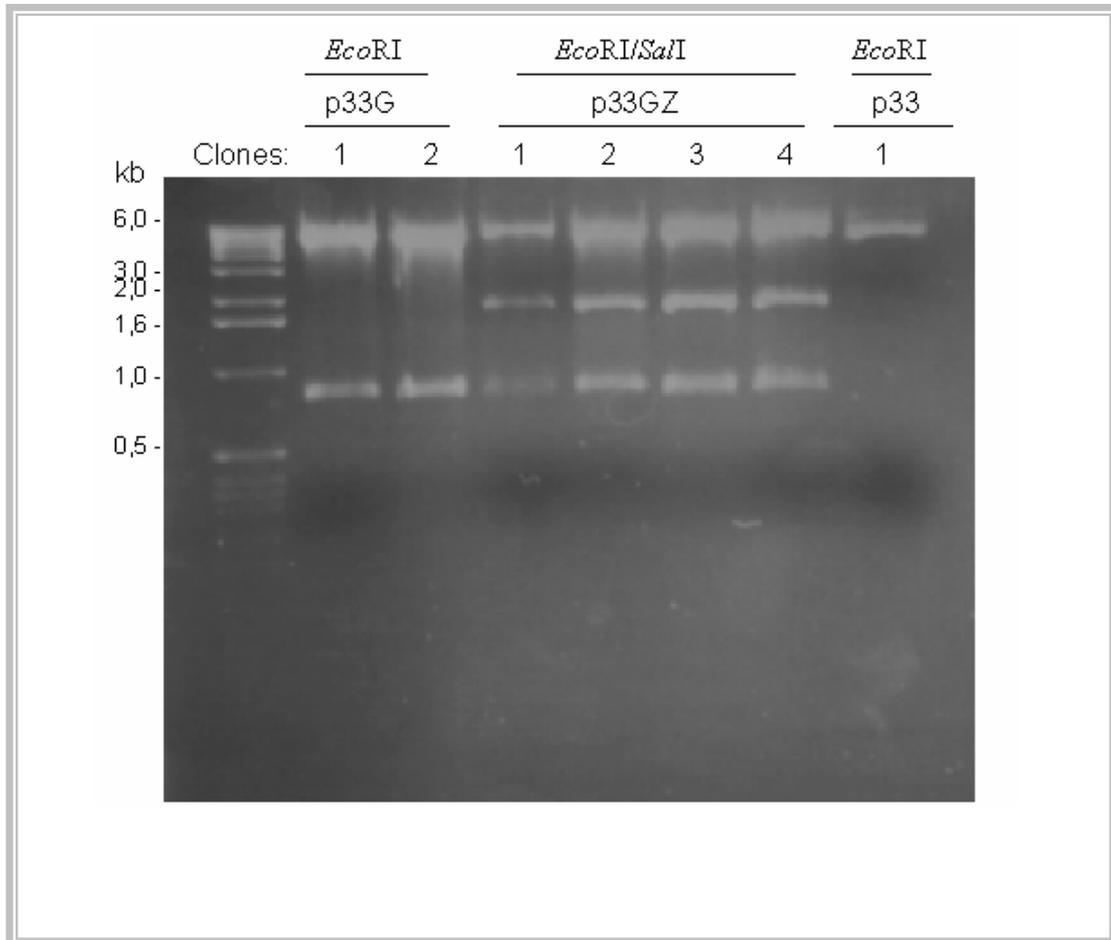


Figura 28: Análise em gel de agarose 1% do perfil de restrição das construções p33G e p33GZ.

p33G com *EcoRI*. Esses plasmídeos recombinantes foram utilizados em experimentos de transfecção do gene *TcZFP8* em formas epimastigotas de *T. cruzi*, porém os resultados não foram conclusivos, já que no controle apenas com a expressão da GFP em fusão não detectamos a proteína na célula. Alternativamente então foi realizado um experimento de imunocitocalização da TcZFP8 em células de epimastigotas de *T. cruzi*.

VII - Expressão em *E. coli* da TcZFP8 com fusão a His-tag, purificação da proteína recombinante e produção de anticorpo policlonal contra esta proteína

Com o objetivo então de realizar a imunocitocalização da TcZFP8 em células epimastigotas de *T. cruzi*, em uma primeira etapa, foi obtida a expressão da TcZFP8 com fusão a His-tag para a purificação posterior da proteína em coluna de níquel. Para tal, o fragmento de 1,8 kb contendo o gene TcZFP8 foi subclonado, a partir do plasmídeo pGEX utilizando as enzimas de restrição *BamHI* e *SalI*, no plasmídeo pET28a+, para a clonagem direcionada e em fase do gene *TcZFP8*, obtendo-se assim o plasmídeo recombinante pETZ2.2 (Figura 29). Células de *E. coli* BL21 contendo esse plasmídeo foram utilizadas em experimentos de indução da expressão do gene *TcZFP8* em diferentes tempos e o perfil protéico total foi analisado em gel de poliacrilamida desnaturante. A Figura 30 mostra os resultados desta análise onde foi verificada a expressão da proteína de fusão TcZFP8-His tag de aproximadamente 21,5 kDa. Como controle negativo foi utilizado células de *E. coli* BL21 contendo o plasmídeo pET28a+ sem o gene clonado, onde não visualizamos proteína induzida.

Deste modo, a proteína TcZFP8-His tag oriunda da expressão de cultura contendo o plasmídeo pETZ2.2 foi produzida e purificada por cromatografia de afinidade ao níquel. O resultado da purificação da proteína TcZFP8-His tag em Ni-NTA revelou um aumento de sua massa molecular de 21,4 kDa para 25,9 kDa, fato explicado devido ao processo de alquilação (Figura 30).

Após a purificação da proteína TcZFP8-His tag procedemos à produção de anticorpos policlonais em coelho. Após 45 dias, coletamos o soro e testamos em extratos protéicos de *E. coli*

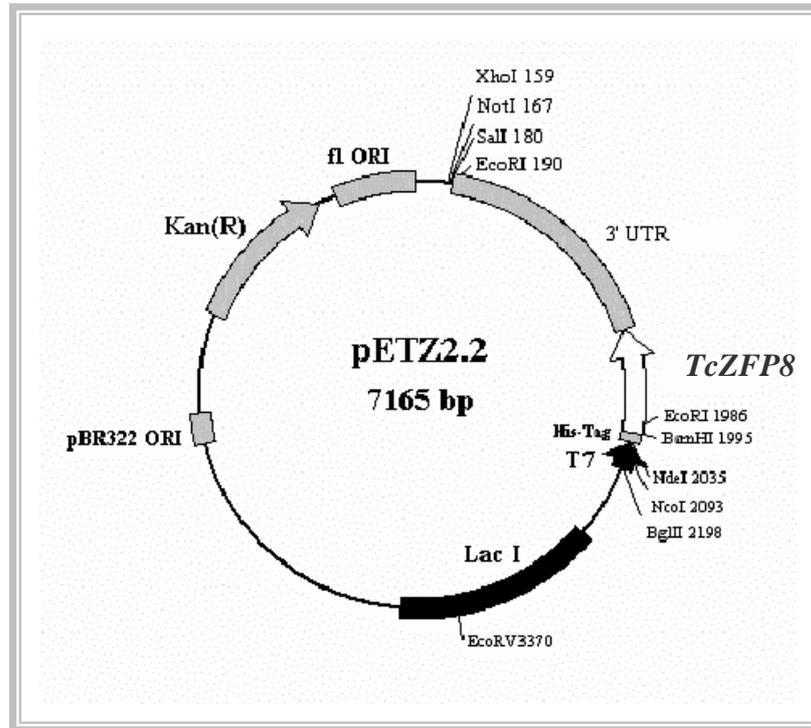


Figura 29: Mapa genético do plasmídeo pETZ2.2, utilizado para expressão da proteína TcZFP8 fusionada ao His-tag em *E. coli*.

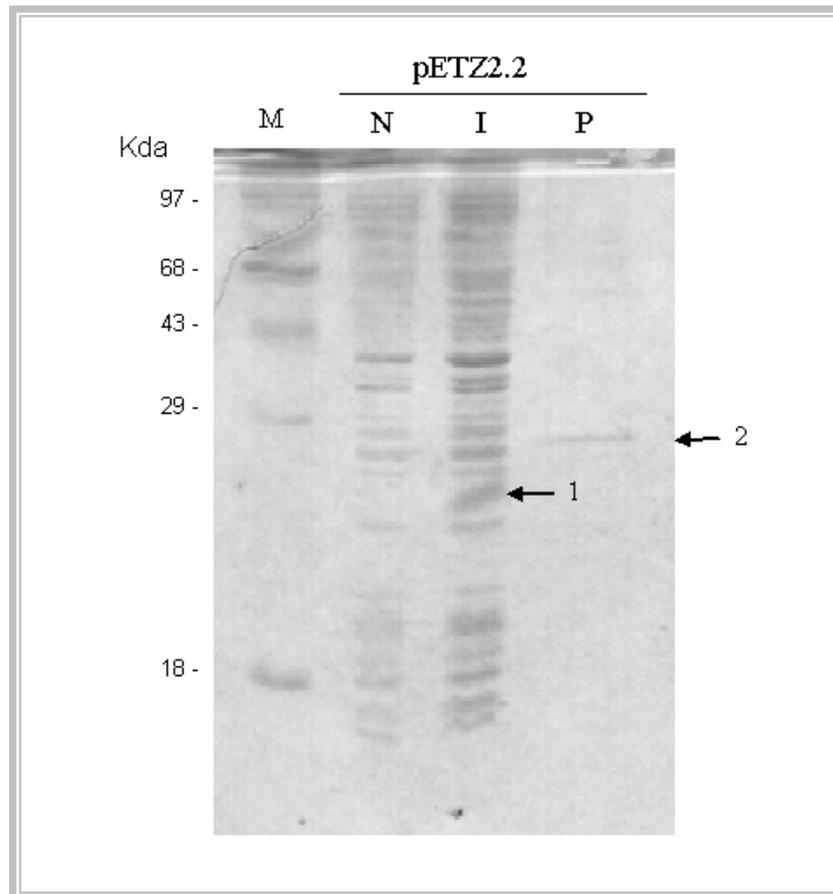


Figura 30: Expressão e purificação da TcZFP8-His-tag por cromatografia de afinidade em Ni-NTA. A seta 1 indica a presença da proteína recombinante TcZFP8-His tag em extratos protéicos de *E. coli*. A seta 2 indica a proteína purificada. N - Extrato protéico não induzido, I- Extrato protéico induzido e P- Proteína purificada. M - Marcador de massa molecular.

contendo as construções pEXZ (pGEX-4T-1/inserto) e pETZ2.2 (pET28a+/inserto). O anticorpo policlonal produzido foi capaz de reconhecer as proteínas induzidas dos extratos protéicos de *E. coli* tanto contendo pEXZ quanto pETZ2.2.

VIII - Análise da expressão e localização da proteína TcZFP8 em lisados de células de *T. cruzi* por *Western blot*

Para confirmarmos a expressão da TcZFP8 nas três formas do parasita sugerida pelo experimento de *Northern blot*, foi realizada a análise por *Western blot* utilizando extratos protéicos totais das três formas de *T. cruzi* e o anticorpo policlonal produzido. A Figura 31A mostra a presença da proteína TcZFP8 nas três formas de *T. cruzi* (amastigota, epimastigota e tripomastigota), confirmando assim que TcZFP8 provavelmente não é uma proteína estágio-específica.

Foi realizada também a análise por *Western-blot* com extratos protéicos de frações mitocondrial/nuclear e citoplasmática de formas epimastigotas de *T. cruzi* (cepa Dm28c e cepa Y). A presença da proteína TcZFP8 foi detectada apenas nas frações mitocondrial/nuclear de ambas as cepas, conforme mostrado na Figura 31B.

IX - Imunocitocalização da proteína TcZFP8 em formas epimastigotas de *T. cruzi*:

Com o objetivo de confirmar a citocalização nuclear e/ou mitocondrial da proteína TcZFP8, foram realizados experimentos de imunofluorescência indireta em células da forma epimastigota de *T. cruzi*. Como observado na Figura 32, os resultados desses experimentos mostraram a localização da TcZFP8 como sendo nuclear e não mitocondrial, co-localizando-a com o DNA genômico marcado com DAPI (Figura 32).

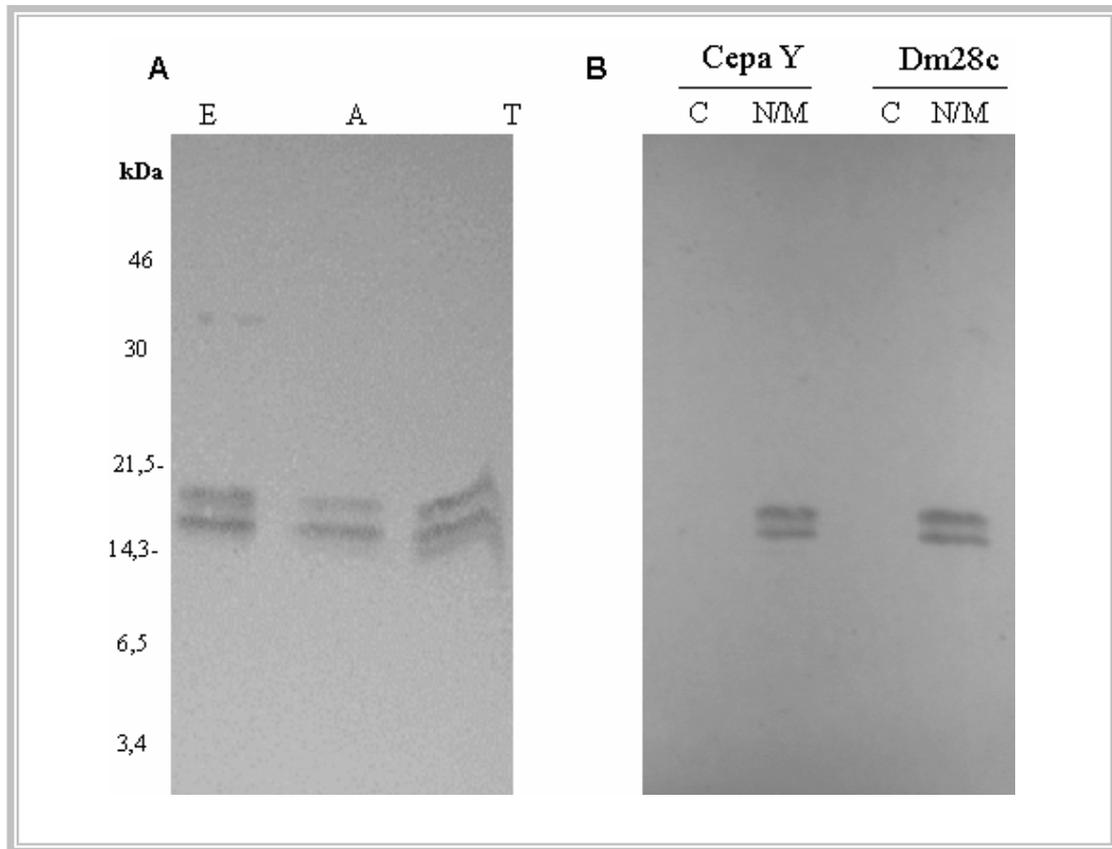


Figura 31: Análise por *Western blot* de lisados protéicos de *T. cruzi* utilizando o anticorpo policlonal anti-TcZFP8 produzido em coelhos. O painel A contém 40 μ g de extrato protéico total de células de formas epimastigotas, amastigotas e tripomastigotas: E - epimastigotas, A - amastigotas e T - tripomastigotas. O Painel B contém 40 μ g de frações protéicas nuclear/mitocondrial e fração citoplasmática de duas linhagens celulares de formas epimastigotas de *T. cruzi* (Dm28c e cepa Y): C-fração citoplasmática e M/N - fração nuclear/mitocondrial. M - Marcador de massa molecular em kDa.

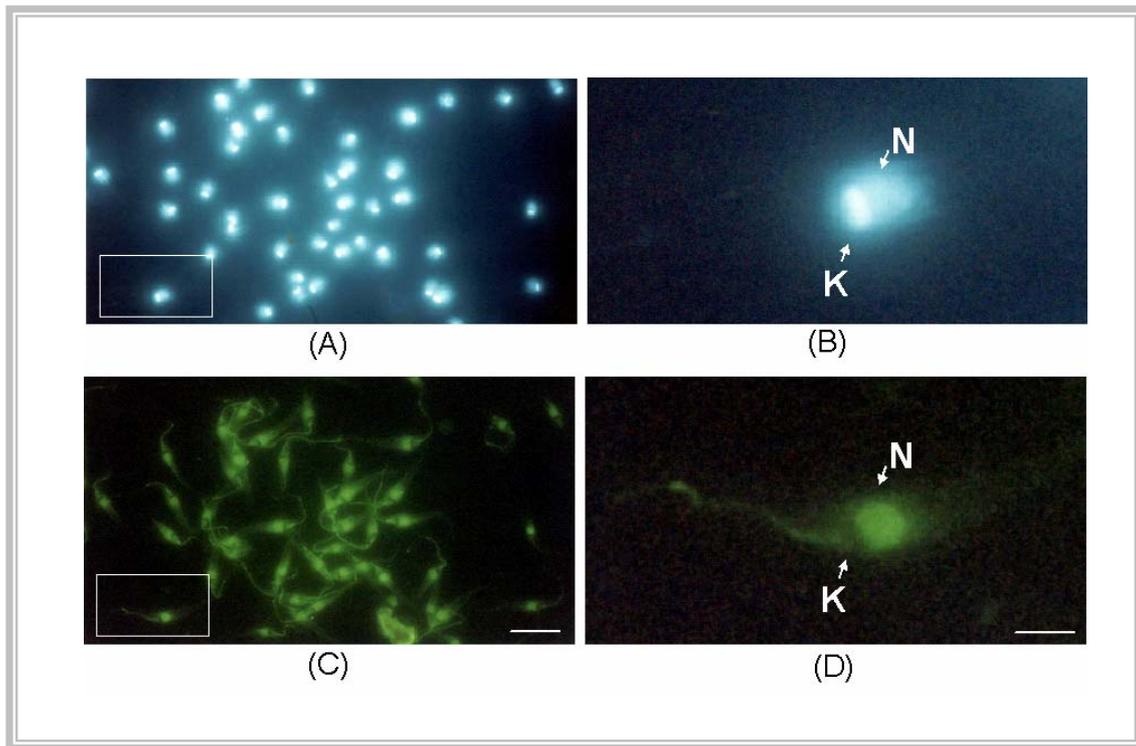


Figura 32: Análise por imunofluorescência indireta em células epimastigotas de *T. cruzi* demonstram a localização nuclear da proteína TcZFP8. Painéis A e B (escala bar = 2 μm); células epimastigotas tratadas com DAPI. Painéis C e D: células epimastigotas tratadas com o anticorpo secundário *goat anti-rabbit* conjugado com FITC (escala em barra = 10 μm). N = núcleo; K = Cinetoplasto.

DISCUSSÃO

Várias proteínas contendo o motivo *zinc finger* têm sido encontradas em diversos organismos e microrganismos e estão envolvidas em vários eventos celulares, dentre os quais a regulação da expressão de genes eucarióticos. Algumas proteínas contendo esse motivo foram relatadas em tripanossomatídeos.

Em *Trypanosoma cruzi* já foram identificados alguns genes codificando proteínas com motivo *zinc finger*: Um elemento retrotransposon non-LTR (*long terminal repeat*) presente em alto número de cópias contendo várias fases abertas de leitura, dentre as quais, uma denominada ORF 3 contendo o motivo *zinc finger* CCHH (Martin *et al.*, 1995); a família de proteínas chaperoninas denominadas DnaJ (tcj1, tcj2 e tcj3) contendo uma região rica em aminoácidos glicina e fenilalanina e o motivo *zinc finger* CXXCXGXG (Tibbets *et al.*, 1998); uma proteína denominada poly-*zinc finger* (PZFP1) contendo sete motivos CCHC (Espinosa *et al.*, 2003); uma proteína denominada PDZ5 contendo cinco motivos *zinc finger* CCHC (Coelho *et al.*, 2003) e a proteína denominada tcZFP1 contendo cada uma um motivo *zinc finger* CCCH (Morking *et al.*, 2004) dentre outras descritas na introdução deste trabalho.

Em nosso trabalho nós identificamos três genes codificando para proteínas *zinc finger* pertencentes a um grupo de genes contidos na banda XX de *T. cruzi* (Figura 33). As proteínas preditas denominadas de TcPDZ5, TcPZFP1 e TcZPF8 compartilham o motivo *zinc finger* CCHC encontrado em um grupo de diversas proteínas que se ligam a ácidos nucléicos. Essas proteínas podem estar, portanto, envolvidas em vários aspectos do controle da homeostase e diferenciação celular. Especificamente, as proteínas aqui identificadas possuem o motivo CX₂CX₄HX₄C que é presente na proteína de nucleocapsídeo viral de HIV-1. É descrito na literatura que esse tipo de estrutura se liga a fitas simples de ácidos nucléicos.

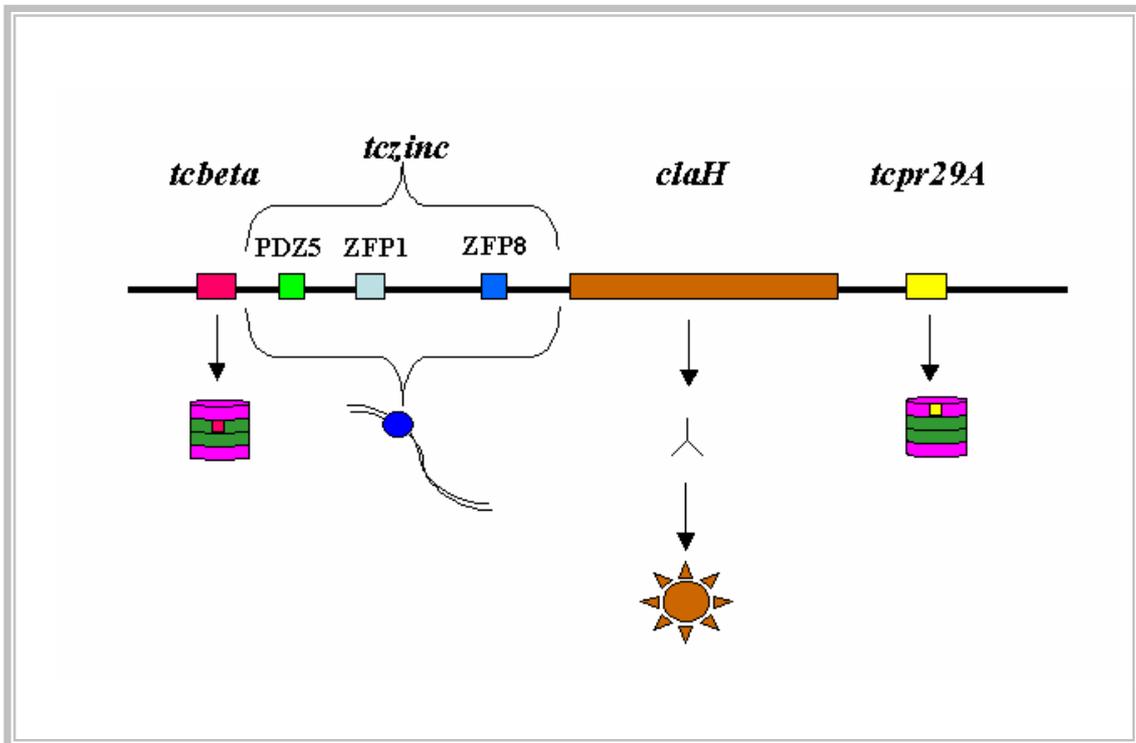


Figura 33 – Esquema da região genômica de 17 kb da banda XX de *T. cruzi*. Retângulos correspondem às regiões de fase de leitura. *TcBETA5* - gene para subunidade beta proteossomal; *TcPR29A* - gene para subunidade alfa proteossomal; *TcCLAH* - gene para a cadeia pesada de clatrina. As regiões indicadas com colchete contêm três fases de leitura para *zinc fingers* CCHC: PDZ5, ZFP1 e *TcZFP8*.

A proteína TcPDZ5 é homóloga a UMSBP de *C. fasciculata* (Tzfati *et al.*, 1995) e a UMSBP de *T. brucei*, uma proteína que se liga DNA fita simples, especificamente a uma região conservada de 12 nucleotídes denominada seqüência universal de minicírculos (UMS) localizada na fita pesada de uma origem de replicação de DNA minicírculo de cinetoplasto (Tzfati *et al.*, 1995). A proteína TcPDZ5 também mostrou se ligar a seqüência duodecâmera UMS, o que fortemente sugere que esta proteína é uma proteína de ligação a UMS do *T. cruzi* (Coelho *et al.*, 2003).

A proteína TcPFZP1 é uma proteína que se localiza principalmente no citoplasma e se liga a oligonucleotídeos de DNA/RNA fita simples carreando seqüências de reconhecimento de outras proteínas *zinc finger* do tipo CX₂CX₄HX₄C, como a UMS de *C. fasciculata* e a seqüência hexâmera da GP63 em *L. major* (Espinosa *et al.*, 2003)

O terceiro gene aqui identificado é um novo gene de *T. cruzi* (TcZFP8), codificando uma proteína *zinc finger* com oito domínios CX₂CX₄HX₄C e apresentando alto grau de identidade e similaridade com *ZINC6* de *L. major*. Nós encontramos três seqüências gênicas correspondendo para os seus alelos na cepa CL Brener, mas uma descrição sobre a caracterização das mesmas é ausente na literatura. Além disto, não encontramos nenhum gene homólogo no genoma de *T. brucei*. De posse desses resultados, prosseguimos então com a caracterização do gene *TcZFP8* e a sua proteína em *T. cruzi*.

Para esclarecermos o padrão de expressão desse gene em células de *T. cruzi*, experimentos de *Northern blot* foram realizados inicialmente. Esses experimentos mostraram a presença de transcritos estáveis desse gene nas três formas de *T. cruzi*. Com o objetivo de caracterizar, citolocalizar e determinar a provável função da proteína TcZFP8 de *T. cruzi*, outros experimentos foram realizados.

Após a expressão heteróloga desse gene em *E. coli*, purificação da proteína recombinante e obtenção de anticorpos policlonais contra a TcZFP8 por imunização de coelhos, experimentos de *Western blot* com lisados de *T. cruzi* mostraram a presença da TcZFP8 das três

formas do parasita e em extratos nucleares/mitocondrial das formas epimastigotas, sendo completamente ausente nos extratos citoplasmáticos, sugerindo uma localização nuclear ou mitocondrial dessa proteína.

Esses experimentos de *Western blot* detectaram duas bandas de proteínas ao invés de uma banda como esperado, tanto nos extratos totais das três formas do *T. cruzi* quanto na fração mitocondrial/nuclear. Da mesma forma, Hendriks *et al.* (2003) descreveu o mesmo fato quando produziu anticorpos policlonais anti-TbCPSF30, uma proteína *zinc finger* do tipo CCCH, que deveria detectar somente uma banda de 30 kDa correspondente à proteína em extratos protéicos de *T. brucei*, mas duas bandas com diferença pequena de massa molecular como para a TcZFP8 foram detectadas. Acredita-se que as duas bandas sejam de fato a mesma proteína com formas diferentes, resultado de modificações pós-traducionais.

O fato de TcZFP8 ter sido encontrada nas três formas do parasita: amastigota, tripomastigota e epimastigota em concentrações equimolares, nos revela que TcZFP8 não é estágio específica de uma das formas de *T. cruzi* e que provavelmente ela está envolvida em algum processo comum às três formas do parasita. Uma vez que as formas tripomastigotas são células não replicativas e TcZFP8 está presente na mesma concentração em que nas células replicativas amastigotas e epimastigotas, supõe-se que esta proteína não está envolvida no controle da replicação, como sugerido para a proteína UMSBP e TcPDZ5 (Tzfati *et al.*, 1995; Coelho *et al.*, 2003).

Com o objetivo de citolocalizar a proteína TcZFP8 e esclarecer a sua localização nuclear ou mitocondrial, experimentos de transfecção para a expressão ectópica do gene *TcZFP8* fusionado ao gene que codifica a proteína fluorescente verde (GFP) em células de *T. cruzi* foram realizados, porém os resultados não foram satisfatórios já que não foi detectada a expressão da GFP. Alternativamente, foram realizados experimentos de imunofluorescência indireta e foi verificada a citolocalização nuclear da proteína TcZFP8 em formas epimastigotas de *T. cruzi*, colocalizando com a marcação do DNA nuclear e não com o mitocondrial.

Os resultados da imunofluorescência, juntamente com as análises por *Western blot* de lisados nucleares/mitocondriais e citoplasmáticos, confirmaram o fato de que TcZFP8 é uma proteína nuclear de *T. cruzi*. Resultados diferentes foram encontrados para a proteína PZFP1 que foi localizada principalmente no citoplasma (Espinosa *et al.*, 2003) e TcPDZ5 supostamente mitocondrial (Coelho *et al.*, 2003).

O fato da proteína TcZFP8 de *T. cruzi* apresentar oito motivos *zinc finger* do tipo CCHC e estar localizada no núcleo a torna uma forte candidata à ligação a ácidos nucléicos. Esse motivo tem sido descrito em várias proteínas que se ligam a fitas simples de ácidos nucléicos e, desta maneira, proteínas que o possuem podem participar da regulação da expressão gênica. Os oito motivos CCHC identificados na proteína TcZFP8 são idênticos aos 2 motivos CCHC também identificados em proteínas de nucleocapsídeo do vírus HIV tipo 1. Estudos de interação dessas proteínas de HIV revelam que os seus dois motivos CCHC se ligam a 4-5 pares de bases no RNA, em uma região denominada Psi (Rosbach *et al.*, 1997). Através desta informação foram desenhados oligonucleotídeos para o experimento SELEX, dentre os quais o oligonucleotídeo denominado PR102, utilizado como fita molde, contendo uma região randômica de 30 pares de base (considerando-se que cada 2 motivos se ligariam a 5 pares de base no DNA ou RNA, os 8 motivos de TcZFP8 se ligariam em 20 pares de base no DNA ou RNA). Ainda foram adicionadas 10 pares de base extras no oligonucleotídeo randômico para melhor acomodação dos motivos de TcZFP8.

Resultados preliminares obtidos com os experimentos de SELEX e *gel shift*/EMSA mostraram que essa proteína pode se ligar a DNA fita simples. Foram obtidos os possíveis DNAs alvos de interação *in vitro* com a proteína TcZFP8. Esses DNAs foram divididos por “consenso” em 4 populações. Essas populações não apresentaram uma seqüência específica comum a todas, mas todas demonstraram regiões ricas em C e/ou G. Dentre essas populações foi escolhida primeiramente a seqüência contendo o consenso N₂₂**GAGGN**₄ (clone 33/população 4). Resultados preliminares do experimento de *gel shift* indicaram a possível ligação da proteína TcZFP8 à sonda N₂₂**GAGGN**₄.

Interessantemente foi demonstrada a interação de uma proteína CX₂CX₄HX₄C de nucleocapsídeo viral (NC) com somente 4 nucleotídeos GGAG presentes no tetraloop de RNA (elemento de reconhecimento SL3 psi-RNA) (Guzmam *et al.*, 1998; Zeffman *et al.*, 2000). Em termos gerais as sequências: GGNG (SL2 e SL3) e GNRA (SL4) (N=G, C, A ou T; R= purinas) são escolhidas pelas proteínas de nucleocapsídeo viral de acordo com as diferenças de afinidade de ligação intrínsecas ao nucleocapsídeo do HIV (Amarasinghe *et al.*, 2001). Algumas proteínas *zinc finger* do tipo CCHC em tripanosomatídeos têm também demonstrado a capacidade de se ligarem às sequências conhecidas de ácidos nucléicos, tais como: o elemento de resposta ao esterol (5'-GTGGGGTGTGGT-3'), a seqüência universal de replicação de minicírculos em tripanosomatídeos (5'-GGGGTTGGTGTA-3') e a seqüência hexâmera da GP63 (5'-GGCGAGGCGAG-3') (Webb *et al.*, 1993; Tzfati *et al.*, 1995; Espinosa *et al.*, 2003).

A distribuição nuclear da proteína TcZFP8 e sua possível atividade de ligação ao DNA podem ser significantes *in vivo*. Porém, a obtenção dos possíveis alvos de DNAs ligantes a TcZFP8 por SELEX, bem como o resultado preliminar da ligação *in vitro* à sonda N₂₂GAGGN₄ não garante que a proteína nativa TcZFP8 se ligue a estas seqüências *in vivo*. Espinosa *et al.* (2003) demonstraram que a proteína PZFP1 de *T. cruzi* foi capaz de interagir *in vitro* especificamente com fitas simples de DNA ou RNA contendo uma variedade de consensos de elementos regulatórios, dentre os quais a seqüência UMS (*universal minicircle sequence*). Experimento de imunocitocalização demonstrou a presença da proteína PZFP1 no núcleo e citoplasma de *T. cruzi*. Apesar da proteína PZFP1 se ligar especificamente ao elemento UMS *in vitro*, a ausência desta proteína no cinetoplasto exclui o papel da mesma em replicação de minicírculos como sugerido para UMSBP de *Crithidia fasciculata* (Tzfati *et al.*, 1992). Essas observações demonstraram que a dedução da função de uma proteína baseando-se somente em estudos de ligação a ácidos nucléicos *in vitro* pode ser errônea.

Ainda assim, ensaios de EMSA utilizando clones oriundos das outras três populações e também competidores específicos e inespecíficos poderão ser úteis para determinação da afinidade e especificidade da proteína TcZFP8 para com as diferentes populações aqui identificadas.

Hendriks *et al.* (2003), através da técnica de interferência de RNA, silenciaram o gene TbCPSF30 de *T. brucei* causando a morte dos parasitas, devido ao acúmulo de moléculas de mRNA não processadas por reações de *trans*-splicing. Através deste e outros resultados, o autor pôde inferir a possível função da proteína *poly zinc finger* CPSF30 como sendo um dos componentes da maquinaria requerida para processamento policistrônico de RNA neste parasita.

Ramboarina *et al.* (2004) discute a forte relação estabelecida entre a organização tri-dimensional de NCp7 do HIV-tipo 1 (contendo os dois motivos CX₂CX₂HX₄C idênticos aos 8 de TcZFP8) e sua função biológica. Nesta publicação, os autores revelam que mutações no motivo CCHC da zinc finger N-terminal, tornado-o CCCC ou CCHH, leva a uma completa baixa de infectividade. No caso do CCCC, o comportamento foi o surgimento de uma baixa proximidade espacial entre as duas zinc fingers, impedindo interações das mesmas com ácidos nucleicos, o reconhecimento pela transcriptase reversa e diminuição de atividade chaperonina.

A possível função da proteína TcZFP8 poderia também ser elucidada através da técnica de interferência de RNA, mas até o presente momento não foi detectada a presença da maquinaria celular deste tipo de regulação em células de *T. cruzi*, impossibilitando, portanto a utilização desta técnica. Outros experimentos como: modelagem molecular, mutações internas aos motivos CCHC, *knockout*, análises de *microarray*, dentre outros, poderão elucidar a provável função de TcZFP8 no *T. cruzi*.

Como mostrado nesse trabalho, a família *zinc finger* em tripanosomatídeos é muito complexa e serão necessários esforços para estudar e determinar as funções exatas destes genes e proteínas. Os nossos resultados nos levam a crer que o fato de TcZFP8 apresentar o domínio conservado *zinc finger* presente em proteínas que se ligam a fitas simples de DNA/RNA, localizar-se no núcleo e preliminarmente se ligar a DNA fita simples, a torna uma forte candidata à proteína nuclear

envolvida em atividade gênica. Assim, a identificação e a caracterização de fatores *trans-acting* que se ligam a moléculas de DNA ou RNA muito nos ajudará a compreender os mecanismos de expressão gênica e sua regulação em tripanosomatídeos. A continuidade dos estudos moleculares poderá dissecar o papel dessas moléculas que podem também estar envolvidas no processo de infecção pelo *Trypanosoma cruzi* e avançar na descoberta de novos alvos de drogas e vacinas eficazes ao tratamento e prevenção da doença de Chagas.

CONCLUSÃO

Em uma região genômica de 17 kb da banda XX de *Trypanosoma cruzi*, foram identificadas três fases de leituras entre o gene que codifica para a subunidade beta proteassomal e o gene que codifica a cadeia pesada de clatrina e os genes foram denominados de *TcPDZ5*, *PZFP1* e *TcZFP8*. Esses três genes codificam proteínas com cinco, sete e oito motivos *zinc finger* do tipo $CX_2CX_4HX_4C$, respectivamente.

O gene *TcZFP8* foi estudado mais profundamente neste trabalho e nossos resultados mostram que o mRNA de *TcZFP8* é transcrito nas três formas do *T. cruzi*. A proteína também é expressa nas três formas do *T. cruzi* e se localiza no núcleo onde possivelmente interage com fitas simples de ácidos nucleicos.

Para elucidar o papel desta proteína em *T. cruzi* propomos os seguintes experimentos como perspectivas deste trabalho:

1. Citolocalização da proteína TcZFP8 nas formas amastigota e tripomastigota de *T. cruzi* para confirmação da localização nuclear também nessas formas do parasita;
2. Análise da ligação da proteína TcZFP8 às diferentes populações de DNAs obtidos com o experimento SELEX, realizando experimentos de *gel shift* com competidores específicos e inespecíficos;
3. Silenciamento do gene *TcZFP8* e análise dos mutantes;
4. Mutação nos motivos internos de CCHC da proteína TcZFP8 e análise da interação ao ácido nucleico;
5. Modelagem molecular da proteína TcZFP8.

BIBLIOGRAFIA

- Abu-Elneel, K.; Kapeller, I. & Shlomai, J. 1999. Universal minicircle sequence-binding protein, a sequence-specific DNA-binding protein that recognizes the two-replication origins of the kinetoplast DNA minicircle. *J. Biol. Chem.*, 274: 13419-26.
- Alberts, B.; Bray, D.; Lewis, J.; Raff, M.; Roberts, K. & Watson, D. 1997. *Biologia Molecular da Célula.*, Ed. Artes Médicas, São Paulo. 3ª Ed.
- Amarasinghe, G.K.; Miskimon, M.; Chancellor K.J.; McDonald J.A.; Miller R.R.; Rouse M.D. & Summers, M.F. 2001. Steam-loop SL4 of the HIV-1 psi RNA packaging signal exhibits weak affinity for the nucleocapsid protein: Structural studies and implications for genome recognition. *J. Mol. Biol.*, 314: 961-70.
- Aranda, A. & Pascual, A. 2001. Nuclear Hormone Receptors and Gene Expression. *Physiol. Rev.*, 81: 1269-304.
- Assis, S.M., Simões, D.G. & Lima, B.D. 1998. Molecular cloning of a *Trypanosoma cruzi* clathrin heavy chain gene. *Mem. Inst. Oswaldo Cruz*, 93 (suppl. II): 155.
- Ausubel, F. M.; Brent, R.; Kingston, R.E.; Moore, D.D.; Seidman, J.G.; Smith, J.A. & Struhl, K. 1995. *Current protocols in molecular biology*. John Wiley & Sons, Inc., New York, N.Y.
- Bartholomeu, D.C.; Batista, J.A.N.; Vainstein, M.H.; Lima, B.D. & Martins de Sá, C. 2001. Molecular cloning and characterization of a gene encoding the 29-kDa proteasome subunit from *Trypanosoma cruzi*. *Mol. Genet. Genomics*, 265: 986-92.
- Berriman, M.; Ghedin, E.; Hertz-Fowler, C.; Blandin, G.; *et al.* 2005. The genome of the African trypanosome *Trypanosoma brucei*. *Science* 15; 309 (5733): 416-22.
- Borden, K.L. 2000. RING domains: Master builders of molecular scaffolds. *J. Mol. Biol.*, 295: 1103-12.
- Brecht, M.; Niemann, M.; Schluter, E.; Muller, U. F.; Stuart, K.; Goring, H. U. 2005. TbMP42, a protein of the RNA Editing Complex in African Trypanosomes, has Endo-Exoribonuclease activity. *Mol. Cell.*, 17: 621-30.
- Brener, Z. & Chiari, E. 1963. Variações morfológicas observadas em diferentes amostras de *Trypanosoma cruzi*. *Rev Inst Med. Trop. São Paulo*, 5: 220-4.

- Brener, Z. 1965. Comparative studies of different strains of *Trypanosoma cruzi*. *Annu. Trop. Med. Parasitology.*, 59: 19-26.
- Bullock, W.O., Fernandez, J.M. & Short, J.M. 1987. XL1-Blue: A highly efficiency plasmid transforming recA *Escherichia coli* strain with beta-galactosidase selection. *BioTechniques*, 5: 376-8.
- Burri, M.; Schlimme, W.; Betschart, B; Kampfler, U. & Hecker, H. 1993. Biochemical and functional characterization of histone H1-like proteins in procyclic *Trypanosoma brucei brucei*. *Parasitol. Research.*, 79: 649-59.
- Caro, F.; Bercovich, N.; Atorrasagasti, C.; Levin, M.J. & Vazquez, M.P. 2005. Protein interactions within the TcZFP zinc finger family members of *Trypanosoma cruzi* implications for their functions. *Biochem. Biophys. Res. Commun.*, 333: 1017-25.
- Chagas, C. 1909. Nova Tripanomiase humana. Estudos sobre morfologia e o ciclo evolutivo do *Schizotrypanum cruzi*, n. gen., n. sp., agente etiolojico da nova entidade morbida do homem. *Mem. Inst. Oswaldo Cruz* 1: 159-218.
- Chang, W.; Lee, D.; Leu, S.; Huang, Y. Lu, M. & Ouyang, P. 2003. Molecular characterization of a novel nucleolar protein, pN040. *Biochem. Biophys. Res. Comm.*, 307: 569-77.
- Coelho, R.E.; Urményi, T.P.; Silveira, J.F.; Rondinelli, E. & Silva, R. 2003. Identification of PDZ5, a candidate universal minicircle sequence binding protein of *Trypanosoma cruzi*. *Int. J. Parasitol.*, 33: 853-8.
- Contreras, V.T.; Araújo-Jorge, T.C.; Bonaldo, M.C.; Thomas, N., Barbosa, H.S.; Meirelles, M.N.L. & Goldenberg, S. 1988. Biological Aspects of the Dm28c Clone of *Trypanosoma cruzi* after metacyclogenesis in chemically defined media. *Mem. Inst. Oswaldo Cruz*, 83: 123-33.
- Covey, S.N. 1986. Amino acid sequence homology in gag region of reverse transcribing elements and the coat protein gene of cauliflower mosaic virus. *Nucl. Acids Res.*, 14: 623-33.
- Cranz, S.; Beger, C.; Baici, A.; Jelesarov, I. & Bosshard. 2004. Monomeric and Dimeric bZIP Transcription Factor GCN4 Bind at the Same Rate to Their Target DNA Site. *Biochem.*, 43: 718-27.
- De Sousa, M.A. 1999. Morphological Characterization of *Trypanosoma cruzi* Chagas (1909) and its Distinction from other Trypanosomes. *Mem. Inst. Oswaldo Cruz*, 94(suppl. I): 205-210.
- De Souza, W. 1984. Cell Biology of *Trypanosoma cruzi*. *Int. Rev. Cytol.*, 86: 197-283.

- De Souza, W. 1999. A Short Review on the Morphology of *Trypanosoma cruzi*: from 1909 to 1999. *Mem. Inst. Oswaldo Cruz*, 94(suppl. I): 17-36.
- Degrave, W.J.; da Silveira, J.F. & Levin, M.J. 1997. Parasite Genome Projects and the *Trypanosoma cruzi* Genome Initiative. *Mem. Inst. Oswaldo Cruz*, 92: 859-862.
- Drozd, M.; Palazzo, S.S.; Salavati, R.; O'Rear, J.; Clayton, C. & Stuart, K. 2002. TbMP81 is required for RNA editing in *Trypanosoma brucei*. *EMBO J.*, 21: 1791-9.
- El-Sayed, N.M.; Myler, P.J.; Bartholomeu, D.C.; Nilsson, D.; *et al.* 2005a. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* 15; 309: 409-15.
- El-Sayed, N.M.; Myler, P.J.; Blandin, G.; Berriman, M. *et al.* 2005b. Comparative genomics of trypanosomatid parasitic protozoa. *Science* 15; 309: 404-9.
- Espinosa, J.M.; Portal, D.; Lobo, G.S.; Pereira, C.A.; Alonso, G.D.; Gómez, E.B.; Lan, G.H.; Pomar, V.R.; Flawiá, M.M. & Torres, H.N. 2003. *Trypanosoma cruzi* poly-zinc finger protein: a novel DNA/RNA-binding CCHC-zinc finger protein. *Mol. Biochem. Parasitol.*, 131: 35-44.
- Famulok, M & Klug, S.J. 1994. All you wanted to know about SELEX. *Mol. Biol. Reports*, 20: 97-107.
- Felsenstein, J. 2002. PHYLIP (Phylogeny Inference Package) version 3.6a3. Distributed by the author. Department of Genome Sciences. University of Washington. Seattle.
- Fragoso, S.P. & Goldenberg, S. 1992. Cloning and characterization of the gene encoding *Trypanosoma cruzi* DNA topoisomerase II. *Mol Biochem Parasitol* 55: 127-34.
- Frank, D. & Guthrie, C. 1992. An essential splicing factor, SLU7, mediates 3' splice site choice in yeast. *Genes Dev.*, 6: 2112-24.
- García-Estrada, C.; Reguera, R.M.; Villa, H.; Requena, M.J.; Muller, S.; Pérez-Pertejo, Y.; Balaña-Fouce, R. & Ordóñez, D. 2003. Identification of a gene in *Leishmania infantum* encoding a protein that contains a SP-RING/MIZ zinc finger domain. *Biochem. Biophys. Acta*, 1629: 44-52.
- Gehring, W.J. & Affolter M. & Burglin, T. 1994. Homeodomain Protein. *Annu. Rev. Biochem.*, 63: 487-526.
- Graham, S.V.; Jefferiers, D. & Barry, J.D. 1996. A promoter directing α -amanitin-sensitive transcription of GARP, the major surface antigen of insect stage *Trypanosoma congolense*. *Nucl. Acids Res.*, 24: 272-81.

- Guzman, R.N.; Wu, Z.R.; Stalling, C.C.; Pappalardo, L.; Borer, P.N.; Summers, M.F. 1998. Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science*, 279: 384-388.
- Hajduk, S.L. & Sabatini, R.S. 1996. *RNA editing: pos-transcriptional restructuring of genetic information*. In: Molecular Biology of Parasitic Protozoa. D.F. Smith & M. Parsons, eds., IRL Press, New York; págs 134-58.
- Harrison, S.C. 1991. A structural taxonomy of DNA-binding domains. *Nature*, 353: 715-9.
- Heath, J.M.; Derebail, S.S.; Gorelick, R.J. & DeStefano, J.J. 2003. Differing roles of the N-terminal and C-terminal zinc fingers in human immunodeficiency virus nucleocapsid protein-enhanced nucleic acid annealing. *J. Biol. Chem.*, 278: 30755-63.
- Hecker, H. & Gander, E.S. 1985. The compaction pattern of the chromatin of trypanosomes. *Biol. of the Cell*, 53: (199-208).
- Hecker, H. & Gander, E.S. 1989. Instability of the nuclear chromatin of procyclic *Trypanosoma brucei brucei*. *Mol. Biochem. Parasitol.*, 37: 225-234.
- Hendriks, E.F. & Matthews, K.R. 2005. Disruption of the developmental programme of *Trypanosoma brucei* by genetic ablation of TbZFP1, a differentiation-enriched CCCH protein. *Mol. Microbiol.*, 57: 706-16.
- Hendriks, E.F.; Abdul-Razak, A. & Matthews, K.R. 2003. tbCPSF30 depletion by RNA interference disrupts polycistronic RNA processing in *Trypanosoma brucei*. *J. Biol. Chem.*, 78: 26870-8.
- Hendriks, E.F.; Robinson, D.R.; Hinkins, M. & Matthews, K.R. 2001. A novel CCCH protein which modulates differentiation of *Trypanosoma brucei* to its procyclic form. *EMBO J.*, 20: 6700-11.
- Henriksson, J.; Aslund, L. & Pettersson, U. 1996. Karyotype variability in *Trypanosoma cruzi*. *Parasitol. Today*, 12: 108-14.
- Horie, K.; Urano, T.; Ikeda, K. & Inoue, S. 2003. Estrogen-responsive RING Finger Protein Controls Breast Cancer Growth. *J. Steroid Biochem. & Mol. Biol.*, 85: 101-4.
- Huang, C.E.; O'Hearn, S.F. & Sollner-Webb, B. 2002. Assembly and function of the RNA editing complex in *Trypanosoma brucei* requires band III protein. *Mol. Cell. Biol.*, 22: 3194-203.
- Ikeda, K.; Orimo, A.; Higashi, Y.; Muramatsu, M. & Inoue, S. 2000. Epf as a Primary Estrogen-Responsive gene in human breast cancer. *Febs Lett.*, 472: 9-13.
- Inoue, H.; Nojima, H. & Okayama, H. 1990. High efficiency transformation of *Escherichia coli* with plasmids. *Gene*, 96: 23-28.

- Ivens, A.C.; Peacock, C.S.; Worthey, E.A.; Murphy, L.; *et al.* 2005. The genome of the kinetoplastid parasite *Leishmania major*. *Science* 15: 309: 236-42.
- Jensen, K.; Shiels, P.S. & Freemont, P.S. 2001. PML Protein Isoforms and the RBCC/TRIM Motif. *Oncogene*, 20: 7223-33.
- Kang, X; Falick, A. M.; Nelson, R.E.; Gao, G.; Rogers, K., Aphasizhev, R. & Simpson, L. 2004. Disruption of the Zinc Finger Protein in the *Leishmania tarentolae* LC-4 (=TbMP63) L-complex Editing Protein Affects the Stability of the L-complex. *J. Biol. Chem.*, 279: 3893-9.
- Klug, S. & Famulok, M. 1994. All you wanted to know about SELEX. *Mol. Biol. Reports*, 20: 97-107.
- Kooter, J.M & Borst, P. 1984. Alfa-amanitin-insensitive of variant surface glycoprotein genes provides further evidence for discontinuous transcription in trypanosomes. *Nucl. Acids. Res.*, 12: 9457-72.
- Laemmli, U. K. 1970. Cleavage of structure proteins during the assembly of the head of bacteriophage T4. *Nature*, 227: 680-85.
- Lazar, M.A. 2002. *Mechanism of action of hormones that act on nuclear receptors*. In: Larsen, P.R.; Kronenberg, H.M.; Melmed, S. & Polonsky, K.S., eds. Williams Textbook of Endocrinology, Saunders Publishing, Philadelphia, 10a ed., cap. 4, pp. 35-44.
- Lima, B.D. & Assis, S.M. 1998. Cloning and analysis of a gene encoding a beta type subunit of *Trypanosoma cruzi* proteasome. *Mem. Inst. Oswaldo Cruz*, 93 (suppl. II): 166.
- Lorry, J.A & Atchey, R.W. 2000. Molecular evolution of the GATA family of transcription factor: conservation within the DNA-binding domain. *J. Mol. Evol.*, 50: 103-15.
- Mabrouk, T. & Lemay, G. 1994. Mutations in a CCHC zinc-binding motif of the reovirus δ -3 protein decrease its intracellular stability. *J. Virol.*, 68: 5287-90.
- Mair, G.; Shi, H.; Li, H.; Djikeng, A.; Aviles, H.O.; Bishop, J.R.; Falcone, F.H.; Gavrilescu, C.; Montgomery, J.L.; Santori, M.I.; Stern, L.S.; Wang, Z.; Ullu, E. & Tschudi, C. 2000. A new twist in trypanosome RNA metabolism: cis-splicing of pre-mRNA. *RNA*, 6: 163-9.
- Manger, I.D. & Boothroyd, J.C. 1998. Identification of a nuclear protein in *Trypanosoma brucei* with homology to RNA-binding proteins from cis-splicing systems. *Mol. Biochem. Parasitol.*, 97: 1-11.

- Martin, F.; Maranon, C.; Olivares, M.; Alonso, C. & Lopes, M.C. 1995. Characterization of a non-long terminal repeat retrotransposon cDNA (L1Tc) from *Trypanosoma cruzi*: homology of the first ORF with the ape family of DNA repairs enzymes. *J. Mol. Biol.* 274: 49-59.
- Marzluf, G.A. 1997. Genetic Regulation of nitrogen metabolism in the Fungi. *Microb. Mol. Biol. Rev.*, 61: 17-32.
- Miller, J.; Fairall, L. & Rhodes, D. 1989. A novel method for the purification of the *Xenopus* transcription factor IIIA. *Nucl. Acids. Res.*, 17: 9185-92.
- Morking, P.; Dallagiovanna, B.M.; Foti, L.; Garat, B.; Picchi, G.F.A.; Umaki, A.C.S.; Probst, C.M.; Krieger, M.A.; Goldenberg, S. & Fragoso, S. 2004. TcZFP1: a CCCH zinc finger protein of *Trypanosoma cruzi* that binds poly-C oligoribonucleotides *in vitro*. *Biochem. Biophys. Res. Comm.*, 319: 169-77.
- Niikura, T.; Hashimoto, Y.; Tajima, H.; Ishizaka, M.; Yamagishi, Y.; Kawasumi, M.; Nawa, M.; Tersashita, K.; Aiso, S. & Nishimoto, I. 2003. A tripartite motif protein TRIM11 binds and destabilizes Humanin, a neuroprotective peptide against Alzheimer's disease-relevant insults. *Eur. J. Neurosci.*, 17: 1150-8.
- Nilsen, T.W. 1994. Unusual strategies of gene expression and control in parasites. *Science*, 264: 1868-9.
- Ouyang, P. & Sugure, S.P. 1996. Characterization of Pinin, a Novel Protein Associated with the Desmosome-intermediate filament complex. *J. Cell. Biol.*, 135:1027-42.
- Ouyang, P. 1999. Antibodies differentiate desmosome-form and nucleus-form Pinin: evidence that Pinin is a moonlighting protein with dual localization at the desmosome and within the nucleus. *Biochem. Biophys. Res. Commun.*, 263: 192-200.
- Pabo, C.O. & Sauer, R.T. 1992. Transcription Factors: Structural Families and Principles of DNA Recognition. *Annu. Rev. Biochem.*, 61: 1053-95.
- Panigrahi, A.K.; Schnauffer, A.; Carmean, N.; Igo, R.P. Jr.; Gygi, S.P.; Ernst, N.L.; Palazzo, S.S.; Weston, D.S.; Aebersold, R.; Salavati, R. & Stuart, K. 2001. Four related proteins of the *Trypanosoma brucei* RNA editing complex. *Mol. Cell. Biol.*, 21: 6833-40.
- Pavletich, N.P. & Pabo, C.O. 1991. Zinc Finger-DNA Recognition: Crystal Structure of Zif268-DNA Complex at 2.1 Å. *Science*, 252: 809-17.
- Pays, E. 1993. Genome organization and control of gene expression in trypanosomatids. *Symp. Soc. Gen. Microbiol.*, 50: 127-160.

- Peisach, E. & Pabo, C.O. 2003. Constraints for zinc finger linker design as inferred from X-ray crystal structure of tandem Zif268-DNA complexes. *J. Mol. Biol.*, 330: 1-7.
- Radwanska, M.; Couvreur, B.; Dumont, N.; Pays, A.; Vanhamme, L.; Pays, E. 2000. A transcript encoding a proteasome beta-subunit and zinc finger protein in *Trypanosoma brucei brucei*. *Gene* 255: 43-50.
- Ramagli, L.S. & Rodriguez, L.V. 1985. Quantitation of microgram amount of protein in two dimensional polyacrylamide gel electrophoresis sample buffer. *Electrophoresis*, 6: 559-63.
- Ramboarina, S.; Druillenec, S.; Morellet, N.; Bouaziz, S. & Roques, B.P. 2004. Target specificity of Human Immunodeficiency Virus Type 1 NCp7 requires an intact conformation of its CCHC N-Terminal Zinc Finger. *J. Virol.*, 78: 6682-7.
- Ramboarina, S.; Moreller, N.; Fournie-Zaluski, M.C.; Roques, B.P. & Moreller, N. 1999. Structural investigation on the requirement of CCHC Zinc Finger Type in Nucleocapsid Protein of Human immunodeficiency Virus 1. *Biochemistry*, 38: 9600-7.
- Ramirez, M.I.; Yamauchi, L.M.; de Freitas, L.H.G. Jr.; Uemura, H. & Schenkman, S. 2000. The use of the Green Fluorescent Protein to monitor and improve transfection in *Trypanosoma cruzi*. *Mol. Biochem. Parasitol.*, 111: 235 - 240.
- Robinson-Rechavi, M.; Garcia, H.E. & Laudet, V. 2003. The Nuclear Receptor Superfamily. *J. Cell. Sci.*; 116: 585-586.
- Romaniuk, P.J. 1985. Characterization of the RNA binding properties of transcription factor IIIA of *Xenopus laevis* oocytes. *Nucl. Acids. Res.*, 13: 5369-87.
- Rosbash, M.; Chapentier, B. & Berlund, J.A. 1997. A high affinity binding site for the HIV-1 nucleocapsid protein. *Nucl. Acids Res.*, 25: 1042-49.
- Roth, M. & Leon, O. 2000. Zinc fingers: DNA binding and protein-protein interactions. *Biol. Res.*, 33:1-17.
- Ruffner, H.; Joazeiro, A.C.; Hemmati, D. & Verma, I.M. 2001. Cancer-predisposing mutations within the RING domain of BRCA 1: loss of ubiquitin protein ligase activity and protection from radiation hypersensitivity. *Proc. Natl. Acad. Sci. USA*, 98: 5134-9.
- Sambrook, J.; Fritsch, E.F. & Maniatis, T. 1989. *Molecular Cloning - a laboratory manual*. USA, Cold Spring Harbor Laboratory Press.

- Sato, S.M. & Sargent, T.D. 1991. Localized and inducible expression of *Xenopus-posterior* (Xpo) gene active in early frog embryos encoding a protein with a 'CCHC' finger domain. *Development*, 112: 747-53.
- Shapiro, T. & Englund, P.T. 1995. The structure and replication of kinetoplast DNA. *Annu. Rev. Microbiol.*, 49: 117-43.
- Simpson, L; Neckelmann, N.; de la Cruz V.F.; Simpson A.M.; Feagin, J.E.; Jasmer, D.P. & Stuart, J.E. 1987. Comparison of the maxicircle (mitochondrial) genomes of *Leishmania tarentolae* and *Trypanosoma brucei* at the level of nucleotide. *J. Biol. Chem.*, 262: 6182-96.
- Studier, F.W.; Rosenberg, A.H.; Dunn, J.J. & Dubendorff, J.W. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.*, 185: 60-89.
- Thompson, J.D.; Higgins, D.G. & Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucl. Acids Res.*, 22: 4673-4680.
- Tibbetts, R.S.; Lensen, J.L.; Olson, C.L.; Wang, F.D. & Engman, D.M. 1998. The DnaJ family of protein chaperones in *Trypanosoma cruzi*. *Mol. Biochem. Parasitol.*, 91: 319-26.
- Tsai, R.Y.L. & Reed, R.R. 1998. Identification of DNA recognition sequences and protein interaction domains of the multiple-Zn-finger protein Roaz. *Mol. Cell. Biol.*, 18: 6447-56.
- Tzfati, Y.; Abeliowich, H.; Kapeller, I. & Shlomai, J. 1992. A single-stranded DNA-binding protein from *Crithidia fasciculata* recognizes the nucleotide sequence at the origin of replication of kinetoplast DNA minicircles. *Proc. Natl. Acad. Sci. USA*, 89: 6891-5.
- Urrutia, R. 1997. Exploring the role of homeobox and zinc finger proteins in pancreatic cell proliferation, differentiation, and apoptose. *Int. J. Pancreatol.*, 22: 1-14.
- Vanhamme, L. & Pays, E. 1995. Control of Gene Expression in Trypanosomes. *Microbiological Reviews*, 59: 223-40.
- Vickerman, K. & Preston, T.M. 1970. Spindle microtubules in the dividing nuclei of trypanosomes. *J. Cell Sci.*, 6: 365-83.
- Wagner, W. & So, M. 1990. Genomic Variation of *Trypanosoma cruzi*: Involvement of Multicopy Genes. *Infec. Immun.*, 58: 3217-24.
- Webb, J.R & McMaster. 1993. Molecular Cloning and Expression of a *Leishmania major* Genes Encoding a Single-stranded DNA-binding Protein Containing Nine CCHC Zinc Finger Motifs. *J. Biol. Chem.*, 268: 13994-4002.

- Worthey, E.A.; Schnauffer, A.; Mian, S.; Stuart, K. & Salavati, R. 2003. Comparative Analysis of Editosome Proteins in trypanosomatids. *Nucl. Acids. Res.*, 31: 6392-478.
- Yan, J.X.; Sanchez, J.C.; Rouge, V.; Williams, KL & Hochstrasser, D.F. 1999. Modified immobilized pH gradient gel strip equilibration procedure in SWISS-2DPAGE protocols. *Electrophoresis* 20: 723-6.
- Zapata-Estrella, H.; Hummel-Newell, C.; Sanchez-Burgos, G.; Escobedo-Ortegon, J.; Ramirez-Sierra, M.J.; Arjona-Torres, A. & Dumonteil, E. 2006. Control of *Trypanosoma cruzi* infection and changes in T-cell populations induced by a therapeutic DNA vaccine in mice. *Imunol.*, 103: 186-191.
- Zeffman, A.; Hassard, S.; Varani, G. & Lever, A. 2000. The major HIV-1 packaging signal is an extended bulged steam loop whose structure is altered on interaction with the Gag polyprotein. *J. Mol. Biol.* 297: 877-893.

ANEXO

TcZFP8, a Novel Member of the *Trypanosoma cruzi* CCHC Zinc Finger Protein Family with Nuclear Localization.

Running Title: Novel *T. cruzi* nuclear CCHC zinc finger protein

Alessandra O. Ericsson, Liliam O. Faria, Waldenor B. da Cruz, Cezar Martins de Sá and Beatriz D. Lima

Laboratório de Biologia do Gene - Departamento de Biologia Celular – Universidade de Brasília, Brasília, Distrito Federal, CEP 70910-900, Brazil

Corresponding author: B. D. Lima

Mail address: Laboratório de Biologia do Gene - Departamento de Biologia Celular – Universidade de Brasília, Brasília, Distrito Federal, CEP 70910-900, Brazil

Telephone number: 55-61-33072598

FAX number: 55-61-32734608

e-mail: bdlima@unb.br

Abstract:

In a 17-kb genomic fragment of *Trypanosoma cruzi* chromosome XX, we identified three tandemly linked genes coding for CX₂CX₄HX₄C zinc finger proteins. We also showed that similar genes are present in *Trypanosoma brucei* and *Leishmania major*, sharing three monophyletic group among these trypanosomatids. In *T. cruzi*, *TcZFP8* corresponds to a novel gene coding for a protein containing eight zinc finger motifs. Molecular cloning of this gene and heterologous expression as a fusion with a His-tag were performed in *Escherichia coli*. The purified recombinant protein was used to produce antibody in rabbits. Using western blot analysis, we observed the presence of this protein in all three forms of the parasite: amastigote, trypomastigote and epimastigote. An analysis of cytoplasmic and nuclear cell extracts showed that this protein is present in nuclear extracts, and indirect immunofluorescence microscopy confirmed the nuclear localization of *TcZFP8*. Homologues of *TcZFP8* in *T. brucei* are apparently absent, while one candidate in *L. major* was identified.

Key words: CCHC-type zinc finger, nuclear protein, *TcZFP8*, *Trypanosoma cruzi*, trypanosomatid.

Introduction

Trans-acting factors such as DNA- or RNA-binding proteins have conserved structural domains that mediate and stabilize their binding to nucleic acid molecules. The zinc finger domain is a very ubiquitous structural element, whose hallmark is the coordination of a zinc atom by several amino acid residues (cysteines and histidines, and occasionally aspartate and glutamate). These structural elements are associated with protein-nucleic acid recognition as well as protein-protein interactions (Leon and Roth 2000), where they are found in proteins involved in differentiation and growth signals, in proto-oncogenes, in general transcription factors, and in the development of eukaryotic organisms (Pabo and Sauer, 1992). One subtype of this zinc finger domain is the CX₂CX₄HX₄C structure that occurs twice in the NCp7 protein of the nucleocapsid of HIV-1. This domain contributes to multiple steps of the viral life cycle, including the proper encapsidation of the HIV RNA through interactions with single-strand nucleic acids and viral proteins (Ramboarina et al., 1999). Another protein that contains seven CX₂CX₄HX₄C zinc finger domains, eukaryotic CNBP, a single-strand DNA-binding protein, plays a fundamental biological role across different species as a transcription factor. For instance, mouse CNBP is located in the nucleus of cells and stimulates cell proliferation and increases *c-myc* promoter activity during embryogenesis (Shimizu et al., 2003).

Several zinc finger proteins sharing the motif CX₂CX₄HX₄C have been identified in trypanosomatids and have been shown to be involved in different cellular functions. In *Leishmania major*, the protein HEXBP, containing nine zinc finger CX₂CX₄HX₄C motifs, binds to the hexanucleotide repeat sequence found in the intervening region of the GP63 gene cluster, the most abundant surface glycoprotein of this protozoan, and it is likely to be involved in DNA replication, structure and repair (Webb and McMaster, 1993). In *Crithidia fasciculata*, a protein containing five CX₂CX₄HX₄C zinc finger domains designated UMSBP binds to the conserved universal sequence of kinetoplast DNA minicircles (Tzfati et al., 1992; Tzfati et al., 1995). Probable homologues of this protein are PDZ5 in *Trypanosoma cruzi* (Coelho et al., 2003) and TbZFP protein in *Trypanosoma brucei* (Radwanska et al., 2000). In *T. brucei*, the RRM1 protein containing three consensus RNA recognition motifs (RRM), two tandem copies of CX₂CX₄HX₄C zinc finger domain and an arginine-serine rich region was localized in the trypanosome nucleus, and it was suggested that it plays a role in *trans*-splicing (Manger and Boothroyd, 1998). Recently, in *T. cruzi*, a poly-zinc finger protein with seven CX₂CX₄HX₄C motifs designated as PZFP1 was identified and shown to bind specifically to the single-strand DNA or RNA oligonucleotides possessing binding sequences of other CX₂CX₄HX₄C proteins, such as the *C. fasciculata* conserved universal minicircle sequence and *L. major* hexanucleotide of GP63 (Espinosa et al., 2003).

We report here the identification of a tandemly linked family of CCHC-zinc finger protein genes in trypanosomatids. All predicted proteins share the zinc finger motif CX₂CX₄HX₄C that is present in the NCp7 protein of the nucleocapsid of HIV-1. This type of structure seems to bind to single-strand nucleic acids (Ramboarina et al., 2004). We also report the characterization of TcZFP8 as a new member of this protein family with nuclear localization.

Material and methods

Isolation and sequence analysis of a *T. cruzi* 17-kb genomic region. A *T. cruzi* (Dm28c clone) genomic λ EMBL3 library (Fragoso and Goldberg, 1992) was previously screened with the cDNA of the ALPHA6 proteasomal subunit (*TcPR29A* gene) (Bartholomeu et al., 2001), and

one of the positive clones was further characterized and found to contain a 17.303-kb fragment of chromosome XX (Figure 1A). DNA was isolated from this selected positive clone, and restriction fragments were sub-cloned into the pBluescript SK+ (Stratagene) or pUC18 (Biolabs). All sub-clones were submitted to automatic sequencing, and computer-aided sequence analysis was performed using University of Wisconsin GCG software package. Nucleotide sequence data reported in this paper are available in GenBank database under accession nos. AF104214 and AY728266. The sequence alignment was carried out using version 1.82 of CLUSTALW (Thompson et al., 1994). A set of 1,000 sequences was generated with the SeqBoot program; the calculation of the distance matrix and the tree construction was carried out with the aid of ProtDist and Fitch programs, respectively. Finally, the consensus tree was obtained with the Consensus tree program. SeqBoot, ProtDist, Fitch and Consensus tree programs are part of the Phylip package, version 3.8 (Felsenstein, 2002).

Heterologous expression of TcZFP8, its purification and antibody production. The sub-clone pU5.4A was used to isolate the 1.8-kb fragment that contains the *TcZFP8* gene by digestion with *Bam*HI and *Sal*I, which was then ligated to the linearized plasmid pET28a+ (Novagen) (*Bam*HI and *Sal*I digestion) to obtain the proper frame with His tag coding sequence. The correct cloning was checked by automatic sequencing. The new plasmid called pETZ2.2 was used to transform *E. coli* strain BL21DE3 to express *TcZFP8* gene.

The BL21DE3 bacteria harboring plasmid pETZ2.2 were grown in 2YT medium (1.6% tryptone, 1.0% yeast extract, 0.5% sodium chloride, pH 7.0) supplemented with 30 µg/mL kanamycin, to an absorbance of 0.6 at 600 nm. The *TcZFP8* production was then induced with 1.0 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) for 2 h at 37°C under aeration. Cells were harvested by centrifugation at 6,000 g for 15 min. The pellet from each 250 mL of culture was resuspended in 6 mL of PBS with protease inhibitors (5 mM phenyl methyl sulfonyl fluoride (PMSF), 2 µg/mL pepstatin, 0.5 µg/mL leupeptin). The cells were disrupted by sonication (5 cycles/30 s) in ice, and Triton X-100 was then added to a final concentration of 1%. The suspension was centrifuged for 20 min at 15,000 g. The supernatant was incubated for 15 min at 4 °C with 2 mL of solution I (20 mM Tris-HCl pH 8.0, 4% SDS, 4% dithiothreitol) followed by the addition of 2.5 mL of solution II (20 mM Tris-HCl pH 8.0, 0.5 M iodoacetamide) and incubation for 15 min at 28 °C in a dark place; this is a modified method of alkylation described by Yan et al. (1999). After the treatment, the lysate was incubated for 1 h at 4 °C with 1 mL of Ni-NTA (Pharmacia) previously treated with equilibrium buffer (0.1 M sodium phosphate, 50 mM sodium chloride, pH 7.3). The protein was eluted with elution buffer (0.1 M sodium phosphate, 50 mM sodium chloride, 7.5 M urea, 0.3 M imidazole, pH 6.3), and the protein concentration was determined using the method described by Ramagli and Rodriguez (1985).

The polyclonal anti-TcZFP8 immune serum was produced by immunization of rabbits with three intraperitoneal injections of 100 µg of recombinant purified protein TcZFP8 at 15-day intervals. Freund's complete adjuvant was used for the first injection and incomplete adjuvant in the subsequent injection. After two weeks from the last injection, blood was collected and the serum was stored at -20 °C.

Cultivation of *T. cruzi*. Epimastigote forms of *T. cruzi*, Dm 28c and Y strains, were maintained in logarithmic growth phase at 28 °C in liver infusion tryptose (LIT) medium supplemented with 10% fetal bovine serum as described elsewhere (Camargo, 1964). Trypomastigote and amastigote cells were obtained by infecting monolayers of HeLa cells, and purified by centrifugation as previously described (Teixeira et al., 1994).

Western blot. To obtain the total cell lysates, *T. cruzi* epimastigote cell culture at late log phase and amastigote and trypomastigote cells were centrifuged at 6,000 g for 15 min, washed twice in PBS and resuspended in SDS-PAGE sample buffer. Nuclear and cytoplasmic fractions of *T. cruzi* epimastigote cells were prepared as described by Batista et al. (1994). The cytoplasmic extract and nuclear enriched fraction were analyzed using 15% SDS-PAGE. Western blot analysis was performed following a standard protocol using polyclonal anti-TcZFP8 antiserum as the first antibody.

Indirect immunofluorescence assay. Exponentially growing *T. cruzi* epimastigote cells were centrifuged at 6,000 g for 3 min, washed twice in PBS and resuspended in PBS supplemented with 1% BSA. Next, the cells were fixed in 100% methanol for 15 min at room temperature and placed on microscopic slides. After air-drying, parasites were permeabilized in 0.2% Triton X-100 in PBS for 15 min followed by saturation with 5% nonfat milk in PBS for 4 h at 4 °C and probing for 4 h at room temperature with 1:50 polyclonal anti-TcZFP8 antiserum in 1% nonfat milk in PBS. Slides were then washed twice with PBS for 15 min and probed 4 h at room temperature with 1:50 FITC-conjugated goat anti-rabbit secondary antibody in 1% nonfat milk in PBS, followed by washing twice in PBS. After that, the slides were stained with 4'-6-diamidino-2-phenylindole (DAPI) for 30 min and washed twice in PBS. After air-drying, the parasites were observed with an Axiovert Carl Zeiss microscope with magnification of 1000X.

Results

In previous work (Bartholomeu et al., 2001), we isolated and characterized the proteasomal ALPHA6 subunit gene, designated as *TcPR29A*, and showed that this gene is located on chromosome XX of *T. cruzi*. A genomic fragment of 17.303 kb was isolated and the gene *TcPR29A* was localized at the end of this region (Figure 1A). Further sub-cloning and sequencing analysis of this fragment revealed the presence of five other open-reading frames with the same transcription direction (Figure 1A). The first one, a 788-bp coding region (*TcBETA5*), was found at the other end of this fragment, and an analysis of databases showed it to be an incomplete gene coding for a proteasomal BETA5 subunit by homology with the BETA5 cDNA of *T. brucei* (GeneBank accession no. AJ132959) and other eukaryotic proteasomal BETA5 proteins. Upstream of the *TcPR29A* gene, a 5.103 kb ORF (*TcCLAH*) was found coding for a clathrin heavy chain protein with 1,701 amino acids in length. Between the *TcBETA5* and *TcCLAH* genes in a region with 8.6 kb of length, we identified three other ORFs encoding proteins containing five, seven and eight CX₂CX₄HX₄C zinc finger domains, respectively.

Blast search of the predicted protein with five zinc finger domains showed that this 14.32 kDa protein corresponds to the PDZ5 of *T. cruzi* (Coelho et al., 2003), homologous to UMSBP of *C. fasciculata* (Tzfati et al., 1992) and CCHC zinc finger protein (ZFP) of *T. brucei* (Radwanska et al., 2000). PDZ5 was described from a cDNA clone of CL Brener strain and the corresponding gene was located on chromosome XX using pulse field electrophoresis and Southern blot (Coelho et al., 2003). The corresponding gene was identified in *T. cruzi* Dm28c strain with its localization in this 17-kb fragment of chromosome XX downstream of the *TcBETA5* gene. In *T. brucei* the same position of the ZFP gene in relation to the proteasomal BETA5 gene was observed in chromosome X (Figure 1A). Radwanska et al. (2000) performed a cDNA library screening and identified a 2.3-kb fragment encoding a proteasomal BETA5 subunit and a putative zinc finger protein, which they showed to be a mRNA precursor or a processing by-product. Thus, PDZ5 and ZFP are strong candidates to be the *T. cruzi* and *T. brucei* homologues of the *C. fasciculata* UMSBP, a single-stranded DNA-binding protein that binds specifically to the 12 conserved nucleotides termed universal minicircle sequence (UMS) located at the heavy strand of the replication origin of the kinetoplast DNA minicircles (Tzfati et

al., 1992). In *L. major* we found two identical copies in tandem of the homologous *C. fasciculata* UMSBP genes, downstream of the *LmCLAH* gene (Figure 1A), which we designated as *LmUMSBP* genes. The predicted protein from these genes shows 80% identity and 92% similarity with the *C. fasciculata* UMSBP (Genebank accession no. AAC32814 and A54598).

Blast search of the predicted protein with seven zinc finger domains (193 amino acids, 20.42 kDa) showed that this protein corresponds to PZFP1 of *T. cruzi* (Espinosa et al., 2003). It was shown that PZFP1 is located mainly in the cytoplasm and binds to single-stranded DNA or RNA oligonucleotides carrying recognition sequences of other CX₂CX₄HX₄C proteins, such as *C. fasciculata* UMS and *L. major* GP63 hexamer sequences (Espinosa et al., 2003). In *T. brucei*, we found a *PZFP1* homologous gene designated as *TbZINC7* in the same position in relation to the *ZFP* gene (Figure 1A). Alignment analysis between the predicted proteins PZFP1 and TcZINC7 showed 67% identity and 70% similarity. These two proteins are also similar to the *Trypanosoma equiperdum* nucleic acid-binding protein (Genebank accession no. AAB47542), showing 49% identity and 53% similarity with PZFP1 and 74% identity and 75% similarity with TbZINC7. These proteins still have unknown functions.

Alignment of PDZ5 and PZFP1 showed an identity of 95% in 375 nucleotides and 94% in 125 amino acids among the five domains of PDZ5 and the last five domains of PZFP1. This suggests that these two genes could be derived from a gene duplication. The same was observed in *T. brucei* when ZFP and TbZINC7 were compared.

The predicted protein with eight zinc finger domains revealed a novel gene designated as *TcZFP8* that contains an ORF encoding a 192-amino acid protein. An analysis of databases showed the presence of this gene in another strain of *T. cruzi*, CL Brener, where three genes with minor differences were observed. Two gene sequences were obtained from TIGR database, accession nos. 7413 and 5774, and one from GeneBank accession no. AF204398. A few changes in the predicted proteins showed us three patterns, suggesting three alleles of this gene in CL Brener (data not shown). Alignment of the TcZFP8 protein with other known CCHC zinc finger proteins of trypanosomatids (TcPDZ5, TcPZFP1, TbZFP, TbZINC7 and LmUMSBP) revealed that the similarity among them is limited to the cysteines and histidines at the CX₂CX₄HX₄C zinc finger domains. Quantitative alignment data showed that the identity is around 27 to 34% and the similarity is 40 to 49%, confirming that TcZFP8 is a novel gene/protein in *T. cruzi*.

Interestingly, no similar *TcZFP8* genes were found either in the same region (Figure 1A) or in the whole genome of *Trypanosoma brucei* (TIGR and Sanger database). In *Leishmania major*, the presence of the *HEXBP* gene (Webb and McMaster, 1993) downstream of the proteasomal BETA5 gene was observed (Figure 1A). A candidate for the *ZFP8* homologous gene was identified in this region designated as *LmZINC6* (Figure 1A), but the predicted protein contains only six zinc finger domains. Figure 1B shows the alignment of TcZFP8 with LmZINC6. Note that these proteins have a higher degree of identity in their three first domains.

In order to evaluate the evolutionary relationship between these zinc finger proteins, we performed a phylogenetic analysis, as shown in Figure 2, where three distinct monophyletic groups may be observed. TcZFP8 is more closely related to LmZINC6 than to the other proteins, which suggests that both proteins could have the same cellular functions in *L. major* and *T. cruzi*.

Although there are other homologous genes of CL Brener strain in the databases, no descriptions about the TcZFP8 gene/protein were found in the literature. Therefore, we pursued the characterization of this protein in *T. cruzi*.

To further characterize TcZFP8 in *T. cruzi* cells, the heterologous expression of the gene *TcZFP8* in *E. coli* was examined. Figure 3A shows the expression of TcZFP8 as a fusion protein with six histidines (His-tag), as expected with 21 kDa (lane I and arrow 1), which is not present at lane N (uninduced sample). The recombinant TcZFP8 fusion protein was purified

from bacterial lysates using Ni²⁺ affinity columns (Figure 3A, lane P). Note that the purified protein shows a higher molecular mass (26 kDa) than we expected (20 kDa); the alkylation process, to which the protein was submitted during the purification, explains that. This purified recombinant protein was used to produce a polyclonal antibody by immunization in rabbits. Western blot analysis using this antibody confirmed that the antiserum recognized the recombinant protein in *E. coli* total extracts (data not shown).

In order to analyze the TcZFP8 protein in *T. cruzi*, total cell extracts from amastigote, trypomastigote and epimastigote forms of the parasite were used and western blot analysis revealed the presence of this protein in all three forms (Figure 3B). As observed, a double band of proteins with slightly lower molecular weight was revealed by the anti-TcZFP8 antiserum, which probably represents processed or modified forms of the protein. The same was observed by Hendriks et al. (2001) working with a CCCH zinc finger protein of *T. brucei*.

The analysis by western blot using cytoplasmic and nuclear cell extracts of *T. cruzi* epimastigote cells from two different strains (Dm28c and Y) and the anti-TcZFP8 antiserum showed that the TcZFP8 protein is present in the nuclear extracts (Figure 3C). In order to confirm this result and discard a possible contamination of the nuclear extract with kinetoplast proteins, immunofluorescence microscopy analysis was performed on *T. cruzi* epimastigote cells using the anti-TcZFP8 antiserum. A strong nuclear signal was obtained, which co-localized with DAPI labeling (Figure 4A and 4B). An enlarged image of one parasite (Figure 4C and 4D) illustrates this general pattern: a DAPI labeling of the kinetoplast and nuclear DNA and the co-localization of the antiserum labeling in the nucleus, confirming that TcZFP8 is a nuclear zinc finger protein.

In order to determine the binding sequence of TcZFP8, the SELEX (systematic evolution of ligands by exponential enrichment) experiment (Ellington and Szostak, 1990) was performed using purified recombinant GST-TcZFP8 and single-strand DNA oligonucleotides with 30 degenerated bases. After six cycles of binding and PCR, the final amplified products were cloned and automatically sequenced. The resulting sequences did not show any consensus sequences (data not shown), suggesting that the binding site is not restricted to one specific sequence.

Discussion

In this work, we identified three zinc finger protein genes in a cluster of chromosome XX of *T. cruzi*. The predicted proteins, named TcPDZ5, TcPZFP1 and TcZFP8, share the zinc finger motif CCHC found in a diverse range of DNA/RNA-binding proteins involved in various aspects of the control of cell homeostasis and differentiation. Specifically, TcPDZ5, TcPZFP1 and TcZFP8 have the CX₂CX₄HX₄C zinc finger structure that is present in the NCp7 protein of the nucleocapsid of HIV-1. This kind of structure is shown to bind to single-strand nucleic acids.

TcPDZ5 protein is homologous to the *C. fasciculata* UMSBP (Tzfati et al., 1992) and *T. brucei* CCHC (Radwanska et al., 2000). UMSBP, a sequence-specific single-strand DNA-binding protein, binds specifically to the 12 conserved nucleotides termed universal minicircle sequence (UMS) located at the heavy strand of the origin of replication of kinetoplast DNA minicircles (Tzfati et al., 1992). TcPDZ5 protein was also shown to bind to the UMS sequence dodecamer, which strongly suggests that this protein is the *T. cruzi* UMS-binding protein (Coelho et al., 2003).

The TcPZFP1 protein is located mainly in the cytoplasm and binds to single-strand DNA or RNA oligonucleotides possessing recognition sequences of other CX₂CX₄HX₄C proteins, such as *C. fasciculata* UMS and *L. major* GP63 hexamer sequences (Espinosa et al., 2003).

The third gene identified here is the novel *T. cruzi* ZFP8 gene, coding for a zinc finger protein with eight domains CX₂CX₄HX₄C showing a high degree of identity and similarity with the *L. major* ZINC6. We found three gene sequences corresponding to the alleles in the CL

Brener strain, but the description about their characterization was missing. We opted to study and characterize this protein in *T. cruzi* cells. In this way, we obtained the heterologous expression of TcZFP8 and the production of a polyclonal antibody against this recombinant protein. Using this antiserum, we could detect TcZFP8 in equal amounts in all three forms of the parasite, amastigote, trypomastigote and epimastigote, suggesting that TcZFP8 is not developmentally regulated and is probably involved in a process occurring in all three forms. Since trypomastigotes are non-replicative cells and since this protein is present in the same amount as in the replicative cells, epimastigote and amastigote, it is hardly probable that this protein is involved in replication control as suggested for TcPDZ5.

Analysis of nuclear and cytoplasmic extracts by western blot showed that TcZFP8 is a nuclear-specific protein, which was confirmed by immunofluorescence microscopy. Different from PZFP1 which was localized mainly in the cytoplasm, TcZFP8 was present in the nucleus of the cell, and since its zinc finger domains have a high degree of identity and similarity with the same domains in single-strand DNA/RNA-binding proteins, we hypothesize here that this protein may be a nuclear protein involved in gene expression.

As shown in this work, the zinc finger family in trypanosomatids is very complex, and it will be necessary to study and to determine the exact functions of these genes/proteins. The identification and characterization of *trans*-acting factors that bind to DNA and RNA molecules will help us to understand the mechanisms of gene expression and regulation in trypanosomatids.

Acknowledgments

We thank to Dr. Ana D. de Lima, for critical reading of the previous version of this manuscript. We also thank the Wellcome Trust Sanger Institute and The Institute for Genomic Research for the parasite genome sequencing projects. This work was supported by Fundação Universidade de Brasília (FUB) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). A.O.E. was a fellow from Capes (Conselho de Aperfeiçoamento de Pessoal de Nível Superior).

References

- Bartholomeu DC, Batista JAN, Vainstein MH, Lima, BD and Martins de Sa C (2001). Molecular cloning and characterization of a gene encoding the 29-kDa proteasome subunit from *Trypanosoma cruzi*. *Mol Genet Genomics* 265: 986-92.
- Batista JAN, Teixeira SMR, Donelson JE, Kirchoff LV and Martins de Sá C (1994). Characterization of a *Trypanosoma cruzi* poly(A)-binding protein and its genes. *Mol Biochem Parasitol* 67: 301-12.
- Camargo EP (1964). Growth and differentiation in *Trypanosoma cruzi*: origin of metacyclic trypanosomes in liquid medium. *Rev Inst Med Trop Sao Paulo* 6: 93-100.
- Coelho ER, Ürményi TP, Silveira JF, Rondinelli E and Silva R (2003). Identification of PDZ5, a candidate universal minicircle sequence binding protein of *Trypanosoma cruzi*. *Int J Parasitol* 33: 853-58.
- Ellington AD and Szostak JW (1990). In vitro selection of RNA molecules that bind specific ligands. *Nature* 346: 818-22.
- Espinosa JM, Portal D, Lobo GS, Pereira CA et al. (2003). *Trypanosoma cruzi* poly-zinc finger protein: a novel DNA/RNA-binding CCHC-zinc finger protein. *Mol Biochem Parasitol* 131: 35-44.
- Felsenstein J (2002). PHYLIP (Phylogeny Inference Package) version 3.6a3. Distributed by the author. Department of Genome Sciences. University of Washington. Seattle.
- Fragoso SP and Goldenberg S (1992). Cloning and characterization of the gene encoding *Trypanosoma cruzi* DNA topoisomerase II. *Mol Biochem Parasitol* 55: 127-34.

- Hendriks EF, Robinson DR, Hinkins M and Matthews KR (2001). A novel CCCH protein which modulates differentiation of *Trypanosoma brucei* to its procyclic form. *EMBO J* 20: 6700-11.
- Leon O and Roth M (2000). Zinc fingers: DNA binding and protein-protein interactions. *Biol Res* 33: 21-30.
- Manger ID and Boothroyd JC (1998). Identification of a nuclear protein in *Trypanosoma brucei* with homology to RNA-binding proteins from *cis*-splicing systems. *Mol Biochem Parasitol* 97: 1-11.
- Pabo CO and Sauer RT (1992). Transcriptional factors: structural families and principles of DNA recognition. *Ann Rev Biochem* 61: 1053-95.
- Radwanska M, Couvreur B, Dumont N, Pays A, Vanhamme L and Pays E (2000). A transcript encoding a proteasome beta-subunit and a zinc finger protein in *Trypanosoma brucei brucei*. *Gene* 255: 43-50.
- Ramagli LS and Rodriguez LV (1985). Quantitation of microgram amount of protein in two dimensional polyacrylamide gel electrophoresis sample buffer. *Electrophoresis* 6: 559-63.
- Ramboarina S, Druillennec S, Morellet N, Bouaziz S and Roques BP (2004). Target specificity of human immunodeficiency virus type 1 NCp7 requires an intact conformation of its CCHC N-terminal zinc finger. *J Virol* 78: 6682-7.
- Ramboarina S, Moreller N, Fournié-Zaluski M-C and Roques BP (1999). Structural investigation on the requirement of CCHH zinc finger type in nucleocapsid protein of human immunodeficiency virus 1. *Biochem* 38: 9600-7.
- Shimizu K, Chen W, Ashique AM, Moroi R and Li Y-P (2003). Molecular cloning, developmental expression, promoter analysis and functional characterization of the mouse CNBP gene. *Gene* 307: 51-62.
- Teixeira SMR, Russel DG, Kirchhoff LV and Donelson JE (1994). A differentially expressed gene family encoding "Amastin," a surface protein of *Trypanosoma cruzi* amastigotes. *J Biol Chem* 269: 20509-16.
- Thompson JD, Higgins DG and Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucl Acids Res* 22: 4673-80.
- Tzfati Y, Abeliovich H, Avrahami D and Shlomai J (1995). Universal minicircle sequence binding protein, a CCHC-type zinc finger protein that binds the universal minicircle sequence of trypanosomatids. *J Biol Chem* 270: 21339-45.
- Tzfati Y, Abeliowich H, Kapeller I and Shlomai J (1992). A single-stranded DNA-binding protein from *Crithidia fasciculata* recognizes the nucleotide sequence at the origin of replication of kinetoplast DNA minicircles. *Proc Natl Acad Sci USA* 89: 6891-5.
- Webb JR and McMaster WR (1993). Molecular cloning and expression of a *Leishmania major* gene encoding a single-stranded DNA-binding protein containing nine "CCHC" zinc finger motifs. *J Biol Chem* 268: 13994-14002.
- Yan JX, Sanchez JC, Rouge V, Williams KL and Hochstrasser DF 1999. Modified immobilized pH gradient gel strip equilibration procedure in SWISS-2DPAGE protocols. *Electrophoresis* 20: 723-6.

Figure legends

Figure 1. Comparison of *Trypanosoma cruzi*, *Trypanosoma brucei* and *Leishmania major* homologous genomic regions and characterization of a gene family encoding CX₂CX₄HX₄C zinc finger proteins. **A.** Scheme showing the *T. cruzi* chromosome XX 17-kb region, where rectangles represent genes in the following order: *TcBETA5* - proteasomal BETA5 subunit gene; *PDZ5*, *PZFP1* and *ZFP8* – genes coding for proteins containing five, seven and eight zinc finger domains, respectively; *TcCLAH* - clathrin heavy chain gene; and *TcPR29A* - proteasomal ALPHA6 subunit gene. The same region between the proteasomal BETA5 and ALPHA6 subunit genes is shown in *T. brucei* (chr 10, TIGR database) and *L. major* (contig 36.1, Sanger database). **B.** Alignment of the ZFP8 and LmZINC6 proteins. Asterisks indicate conserved amino acids and plus signs indicate similar amino acids. Zinc finger domains are boxed.

Figure 2. Phylogenetic tree depicting the evolutionary relationships between the zinc-finger proteins. Numbers in parentheses represent the number of zinc finger motifs in the proteins. MmCNBP – accession no. P53996, the others in legend of Figure 1 and Material and methods.

Figure 3. TcZFP8 expression and nuclear localization. **A.** Heterologous expression in *Escherichia coli* of the TcZFP8 with His tag fusion and purification of the recombinant protein analyzed in SDS-PAGE. The N and I lanes represent *E. coli* total cell lysates of non-induced and 2 h IPTG-induced, respectively (arrow 1, induced TcZFP8). The P lane shows the purified recombinant TcZFP8 (arrow 2) purified by Ni²⁺ affinity chromatography, after alkylation treatment. M, molecular size markers in kDa. **B.** Western blot of *Trypanosoma cruzi* lysates from three forms of the parasite, epimastigote (lane E), amastigote (lane A), and trypomastigote (lane T), with the anti-TcZFP8 antiserum. M - molecular size markers in kDa. **C.** Western blot of cytoplasmic (lane C) and nuclear (lane N) lysates from *T. cruzi* epimastigotes Y and Dm28c strains, with the anti-TcZFP8 antiserum. M, molecular size markers in kDa.

Figure 4. Nuclear localization of TcZFP8 in *Trypanosoma cruzi* epimastigotes. **A.** Cells stained with DAPI. Scale bar = 10 μm. **B.** Cells stained with DAPI. Magnified area of the rectangle in Figure 4A. Scale bar = 2 μm. **C.** FITC-indirect immunofluorescence analysis with anti-ZFP8 antiserum. Scale bar = 10 μm. **D.** FITC-indirect immunofluorescence analysis with anti-ZFP8 antiserum. Magnified area of the rectangle in Figure 4C. Scale bar = 2 μm. K, kinetoplast, N, nucleus.

Figure 1

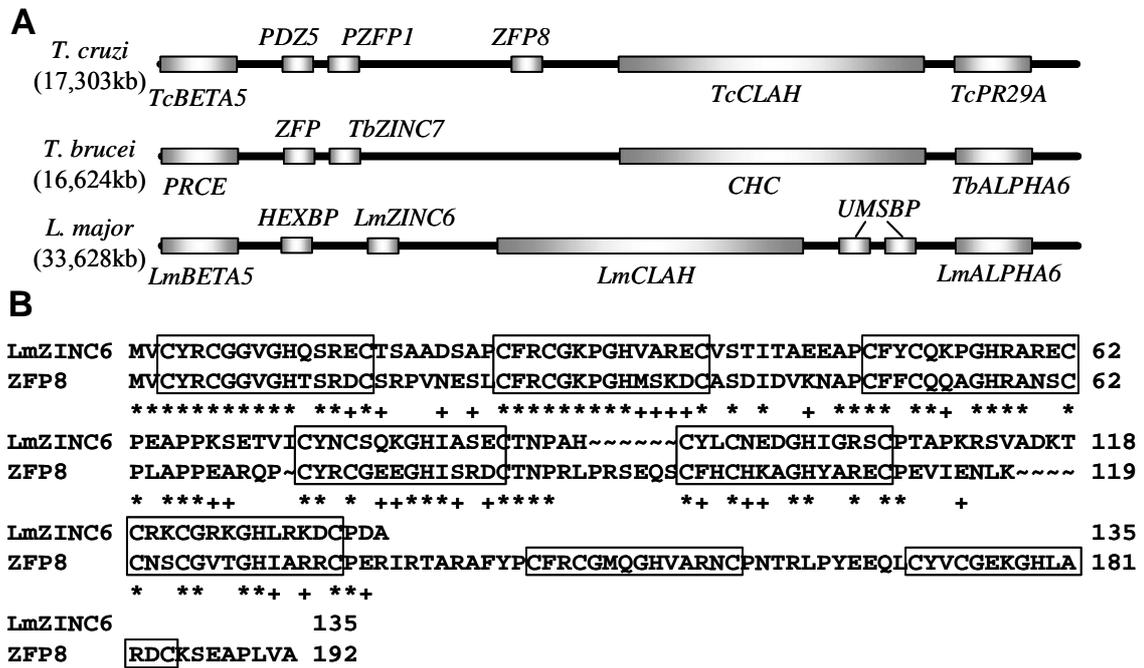


Figure 2

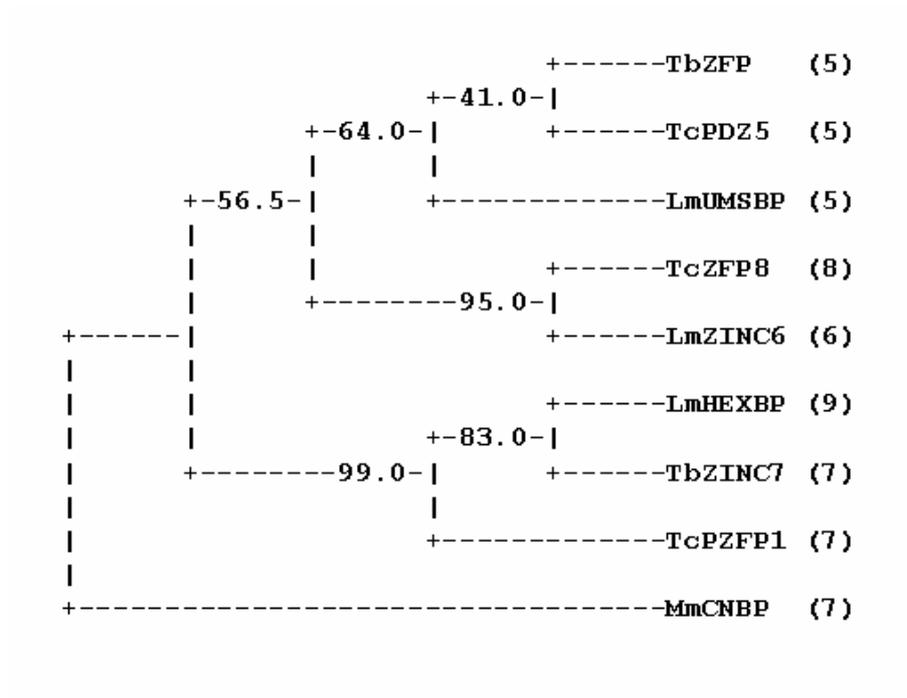


Figure 3

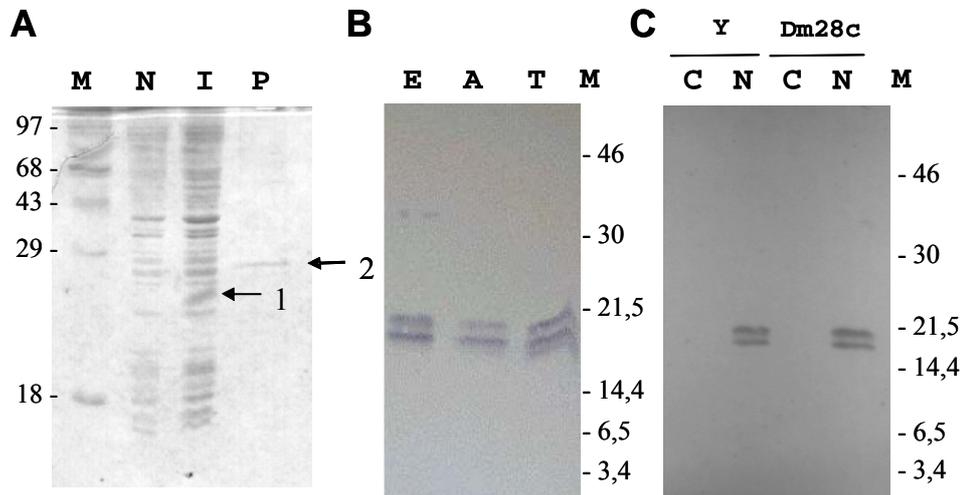


Figure 4

