

Universidade de São Paulo  
Instituto de Física de São Carlos

Ana Laura de Lima

Estudo estrutural e funcional das proteínas spliceossomais U5-15K, U5-102K e Prp43 de *Trypanosoma brucei*

São Carlos

2020



Ana Laura de Lima

Estudo estrutural e funcional das proteínas spliceossomais U5-15K, U5-102K e Prp43 de *Trypanosoma brucei*

Tese apresentada ao Programa de Pós-Graduação em Física do Instituto de Física de São Carlos da Universidade de São Paulo, para obtenção do título de Doutor(a) em Ciências.

Área de concentração: Física Aplicada

Opção: Física Biomolecular

Orientador: Prof. Dr. Otavio Henrique Thiemann

Versão Corrigida

(versão original na unidade que aloja o programa)

São Carlos

2020

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

Lima, Ana Laura

Estudo estrutural e funcional das proteínas spliceossomais U5-15K, U5-102K e Prp43 de Trypanosoma brucei / Ana Laura Lima; orientador Otavio Henrique Thiemann - versão corrigida -- São Carlos, 2020.  
123 p.

Tese (Doutorado - Programa de Pós-Graduação em Física Aplicada Biomolecular) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2020.

1. Trans-splicing. 2. snRNP. 3. Dbl1. 4. Prp6. 5. Prp43. I. Thiemann, Otavio Henrique, orient. II. Título.

## AGRADECIMENTOS

Ao Prof. Dr. Otavio H. Thiemann pela orientação, discussões e ideias que me ajudaram a desenvolver esse trabalho.

Ao Prof. Dr. Marco Tulio Alves da Silva, ao Dr Ivan Rosa e Silva e a M<sup>a</sup> Camila Boralli, ideias e discussões sobre *splicing*, que me ajudaram a desenvolver melhor esse trabalho.

À Perlina pela ajuda em parte dos experimentos desse trabalho.

Aos colegas de laboratório pelas discussões nas reuniões de grupo.

Aos técnicos e professores do Grupo de Biologia Molecular e Estrutural e Cristalografia de Proteína do Instituto de Física de São Carlos pela ajuda com os experimentos.

À Dra Júlia Pinheiro Chagas da Cunha por possibilitar a realização de parte dos experimentos.

Aos meus pais, Maria Lúcia e Marcos César, a minha irmã, Maria Cecília, ao Vinícius e todos os amigos de fora do laboratório pelo apoio e por não me deixarem desistir.

O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) código de financiamento 001.



## RESUMO

LIMA, A. L. **Estudo estrutural e funcional das proteínas spliceossomais U5-15K, U5-102K, Prp43 de *Trypanosoma brucei***. 2020. 123 p. Tese (Doutorado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2020.

A Tripanossomíase Humana Africana (*Human African Trypanosomiasis*, HAT), popularmente conhecida como Doença do Sono, é causada por um protozoário extracelular pertencente ao gênero *Trypanosoma* presente no continente africano. Apesar dos avanços no controle e tratamento da doença, o estudo dos tripanossomatídeos é interessante, pois esses organismos possuem mecanismos bioquímicos únicos e conservados entre si como a transcrição policistrônica e processamento da maioria do seu pré-RNA mensageiro por *trans-splicing*, que podem ser usados de modelo para o estudo de outros tripanossomatídeos. O *splicing* é o processo de retirada de regiões não codificantes do gene dos eucariotos, chamadas de íntrons, e subsequente ligação das regiões codificantes, os éxons. O *splicing* pode ocorrer de duas formas: *cis-splicing* e *trans-splicing*. Tanto no *cis* quanto no *trans-splicing*, as reações de transesterificação são catalisadas pelo spliceossomo, uma maquinaria dinâmica que se rearranja em cada passo da reação através da interação ordenada de cinco ribonucleoproteínas (*small nuclear ribonucleoproteins*, snRNPs) denominadas U1, U2, U4/U6 e U5 e inúmeros fatores não snRNPs. Esses complexos são compostos de pequenos RNAs ricos em uridina (U snRNAs) ligados a proteínas específicas de cada snRNP e proteínas comuns a todas as snRNPs (proteínas Sm). Esse trabalho tem como objetivo o estudo de duas proteínas específicas de U5 snRNP, U5-102K e U5-15K, bem como de uma helicase relacionada à liberação dos produtos do *splicing*, a Prp43. A U5-102K é formada por 44 hélices- $\alpha$  sendo 36 organizadas nas repetições chamadas TRPs. Essa proteína é essencial para a formação do tri-snRNP U5.U4/U6. Já a U5-15K possui uma estrutura semelhante a de uma tiorredoxina e tem uma localização central no tri-snRNP U5.U4/U6, participando da estabilização do U6 snRNA e sua saída do spliceossomo pode estar relacionada à ativação de um dos complexos formados durante a reação. A Prp43 é uma proteína da família DEAH-box e participa dos passos finais do processo de *splicing*, atuando na liberação do íntron excisado. Esse trabalho teve como objetivo a caracterização da atividade de autoclivagem da U5-15K, o estudo de sua estrutura terciária, interação com a U5-102K e a identificação dos parceiros de interação da U5-102K e da Prp43. Foram realizadas mutações sítio dirigidas na sequência da U5-15K, na cisteína 66 e em 4 serinas do C-terminal e os resultados mostram que o resíduo catalítico da autoclivagem seja a Ser 126 e o ponto de clivagem a Ala 142 e sugerem que a atividade pode

ocorrer quando a proteína faz parte do complexo B. Os testes de cristalização usando a U5-15K sem alça e a U5-15K S126A, resultaram em pequenos cristais para a segunda proteína. Para a interação entre U5-15K e U5-102K foi expresso o N-terminal da U5-102K, região onde a interação das duas proteínas ocorre em humanos e levedura, e sinais da interação foram vistos com uso da técnica de DLS. A purificação dos parceiros de interação da U5-102K não resultou em novas proteínas identificadas. Os experimentos com a Prp43 resultaram na identificação de uma proteína homóloga à NtrI, que em humanos e levedura, faz a ativação da atividade da Prp43.

Palavras-chave: Trans-splicing. snRNP. DimI. Prp6. Prp43.



## ABSTRACT

LIMA, A. L. **Structural and functional studies of spliceosomal proteins U5-15K, U5-102K, Prp43 of *Trypanosoma brucei***. 2020. 123 p. Tese (Doutorado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2020.

Human African Trypanosomiasis (HAT), commonly known as Sleeping Sickness, is caused by an extracellular protozoan belonging to the genus *Trypanosoma* spread through African continent. Despite the advances regarding control and treatment of the disease, the studies of trypanosomatids is of interest because of their unique and conserved biochemical mechanisms such as polycistronic transcription and processing of the majority of their pre-mRNA through trans-splicing that can be used to study other trypanosomatides. Splicing is the process of removal of non-coding regions of eucariotic gens, called introns, and subsequent linking of the coding regions, called exons. Splicing may occur in two ways cis-splicing and trans-splicing. Both in cis and trans-splicing the transesterification reactions are catalyzed by the spliceosomo, a dynamic machinery that rearranges itself in each reaction step through ordered interactions of five small nuclear ribonucleoproteins (snRNP) called U1, U2, U4/U6 e U5 and non snRNP factors. These complexes are composed by small RNAs rich in uridine (U snRNAs) bounded to snRNPs specific proteins and proteins common to all snRNPs (Sm proteins). This research aims the study of two U5 snRNP specific proteins, U5-102K e U5-15K, as well as the helicase associated to the release of splicing products, Prp43. U5-102K is composed by 44  $\alpha$ -helices, 36 of which are organized by repetitions called TRPs. This protein is essential to tri-snRNP U5.U4/U6 assembly. U5-15K has a thioredoxin like structure and locates itself in a central position of tri-snRNP U5.U4/U6, participating in U6 snRNA stabilization and its exit may be associated to the activation of a complex formed during the reaction. Prp43 is a protein from DEAH-box family and participates in the final steps of splicing. The objectives of this research were the characterization of U5-15K self-cleavage activity, study of its tertiary structure, study of the interaction between U5-15K and U5-102K and the identification of U5-102K and Prp43 partners. Site-directed mutations were performed at U5-15K's cistein 66 and the 4 serines closest to the C-terminus, the results showed that its self-cleavage catalytic residue is the Ser 126 and its cleavage site the Ala 142 and suggest that this activity can occur when the protein positions itself inside B complex. Crystallization tests were made using the truncated U5-15K and U5-15K S126A, what

resulted in small crystals for the second protein. To perform the interaction experiments between U5-15K and U5-102K, the N-term of U5-102K, region where the interaction occurs in human and yeast, was used and the interaction observed using DLS. The purification of interaction partners using U5-102K as bait did not result in new identifications. The same experiment, using Prp43 as bait resulted in the identification of a protein homologous to human and yeast NtrI that activates Prp43 function on splicing.

Keywords: Trans-splicing. snRNP. DimI. Prp6. Prp43

## LISTA DE FIGURAS

Figura 1 –	Missionário e explorador escocês David Livingstone (esquerda) e do patologista escocês responsável pela descoberta do <i>T. brucei</i> Sir David Bruce (direita).....	24
Figura 2 –	Casos totais de Tripanossomíase Humana Africana reportados a Organização mundial da saúde, de 1940 a 2013.....	25
Figura 3 –	(A) Distribuição geográfica dos casos reportados da Tripanossomíase Humana Africana de 2012 a 2016. (B) Representação das áreas de risco, onde as cores mais escuras representam um risco maior de contrair a doença e as cores mais claras um risco menor. Dados de 2012-2016.....	26
Figura 4 –	Forma do <i>Trypanosoma brucei</i> em esfregaço de sangue de um paciente com Tripanossomíase Humana Africana.....	28
Figura 5 –	Ciclo de vida do <i>Trypanosoma brucei</i> , mostrando a entrada da forma tripomastigota metacíclica na corrente sanguínea do hospedeiro mamífero, o desenvolvimento do parasito no organismo do hospedeiro mamífero, a ingestão da forma tripomastigota sanguínea pelo inseto vetor e desenvolvimento do parasito no organismo do vetor.....	28
Figura 6 –	As 6 principais morfologias dos tripanossomatídeos, sendo a amastigota a forma comum durante a fase intracelular no hospedeiro mamífero, a promastigota a forma comum no inseto vetor, a tripomastigota a forma sanguínea do hospedeiro mamífero e a forma infectante encontrada no inseto vetor.....	29
Figura 7 –	Representação gráfica simplificada da síntese proteica em eucariotos, desde a transcrição do DNA em pré-RNA mensageiro até a tradução do mRNA maduro em proteína.....	33
Figura 8 –	Posição do sítio de <i>splicing</i> 5', 3' (5'SS e 3'SS), flanqueando o íntron, e <i>branch point</i> (BP) em uma molécula de pré RNA mensageiro.....	33
Figura 9 –	Resultado da primeira (esquerda) e segunda (direita) reação de transesterificação para a formação do RNA mensageiro maduro.....	34
Figura 10 –	Esquema mostrando as principais diferenças entre as moléculas de RNA envolvidas no <i>cis</i> e <i>trans-splicing</i> , como os éxons pertencentes ao mesmo pré-RNA mensageiro (passo I) e íntron excisado na forma de um laço (passo III) no caso do <i>cis-splicing</i> e a presença do SL RNA (passo I) e íntron excisado na forma de Y (passo III) no caso do <i>trans-splicing</i> .....	34
Figura 11 –	Estrutura secundária do SL RNA, mostrando a localização do 5'ss, no	

	stem-loop I mais próximo a extremidade 5', e do sítio de ligação das proteínas SM, entre os stem-loops II e III.....	35
Figura 12 –	Esquema da formação do complexo U1 snRNP de mamíferos, mostrando o U1 snRNA, as proteínas SM, comuns às snRNPs, ao redor do sítio SM do U1 snRNA e as proteínas específicas.....	36
Figura 13 –	Representação esquemática do <i>cis-splicing</i> , mostrando as snRNPs envolvidas em cada passo, as modificações sofridas pelo spliceossomo durante o processo e as alterações no RNA mensageiro.....	37
Figura 14 –	Representação das interações entre U6 snRNA, U2 snRNA e o pré-mRNA(em azul).....	38
Figura 15 –	Representação do ataque nucleofílico da adenosina do BS no complexo B*, indicado pela seta vermelha, e formação do complexo C.....	38
Figura 16 –	Representação esquemática da partícula U1-snRNP humana, mostrando o anel formado pelas proteínas SM representado pelo círculo amarelo, o sítio SM do U1 snRNA selecionados em vermelho, as proteínas específicas U1C, em verde, U1 70K, em azul, e U1A, em vermelho, interagindo com o U1 snRNA.....	40
Figura 17 –	Comparação entre a partícula U2 de <i>T. brucei</i> (acima) e a de humano (abaixo), mostrando as proteínas SM-like 15K e 16,5K, em vermelho, substituindo as proteínas SM D3 e B, a diferença de enovelamento dos U2 snRNA dos dois organismos e a interação da proteína U2 40K de <i>T. brucei</i> interagindo com as proteínas SM-like, o que não acontece com a U2A' de humano.....	43
Figura 18 –	Representação esquemática do U4/U6 di-snRNP de levedura, mostrando as proteínas Prp31 (azul), Prp 3 (ciano) interagindo com os snRNAs U4 e U6, a proteína Prp4 (vermelha) e Snu13 (verde) posicionadas próximas ao U4 snRNA e o posicionamento do anel das proteínas Sm próximo a porção 3' do U4 snRNA e do anel LSm próximo a porção 3' do U6 snRNA.....	43
Figura 19 –	Representação esquemática do U4/U6.U5 tri-snRNP humano, onde os pontos pretos e brancos representam interações também observadas em levedura.....	44
Figura 20 –	a) Estrutura resolvida por RMN da U5-15K humana. b) Estrutura resolvida por difração de Raio-X da U5-15K humana. c) Envelope gerado por SAXS alinhado ao modelo gerado por homologia da U5-15 de <i>T. brucei</i> .....	46
Figura 21 –	Localização da DibI (vermelho) no U4/U6.U5 tri-snRNP de levedura, mostrando as regiões de interação com o U6 snRNA (amarelo), Prp8	

	(azul) e Prp31 (magenta). A área de interação entre a DibI e o U6 snRNA está circulada em preto.....	47
Figura 22 –	Estrutura cristalográfica da Prp43 de levedura, com o domínio NTD em azul escuro, RecA1 em verde, RecA2 em lilás, WHD em laranja, HBD em azul claro OB em marrom e CTD em vermelho.....	48
Figura 23 –	Representação gráfica do pareamento de bases e posição dos oligonucleotídeos (setas vermelhas) em relação DNA circular de interesse.....	56
Figura 24 –	Modelo estrutural da U5-15K de <i>T.brucei</i> com as serinas mutadas marcadas em vermelho e a cisteína em verde, o possível ponto de clivagem em amarelo e C terminal em azul.....	73
Figura 25 –	(a) Representação das diferenças de tamanho da proteína clivada, dependendo da posição dependendo da posição do ponto de clivagem, se próxima ao N ou C terminal. (b) SDS-PAGE 15% do teste de clivagem. Coluna 1. Marcador de massa molecular, 2. Alíquota da proteína logo após a purificação, 3. Alíquota da proteína após 12h após a purificação, 4. 24h após a purificação, 5.36h após a purificação e 6. 48h após a purificação. As setas vermelhas indicam a U5-15K não clivada e a U5-15K clivada. A diferença de massa das duas bandas indicam que a proteína é clivada no C-terminal.....	74
Figura 26 –	Localização da isoleucina 141 em amarelo, suposto ponto de clivagem calculado a partir do deslocamento das bandas presentes no marcador de massa molecular e da proteína clivada, em um gel SDS-PAGE 20% <sup>64</sup> , e dos demais resíduos substituídos por alaninas em vermelho.....	75
Figura 27 –	SDS-PAGE 15% do teste de expressão das mutantes U5-15K S120A, U5-15K S137A e U5-15K S144A. 1. Marcador de massa molecular, 2. U5-15K S120A antes da indução com IPTG, 3. U5-15K S120A depois de 16h da indução, 4. U5-15K S137A antes da indução, 5. U5-15K S137A 16h depois da indução, 6. U5-15K S144A antes da indução, 7. U5-15K S144A após a indução. A seta vermelha indica a posição das bandas de expressão das três mutantes.....	76
Figura 28 –	Cromatograma das purificações das proteínas U5-15K pc, U5-15KS120A, U5-15K S137A, U5-15KS144A, U5-15K C66A. A linha preta representa a absorbância a 280nm e a vermelha a porcentagem de tampão B no gradiente de imidazol (ponto mais baixo 0%, ponto mais alto 100%). Os picos indicados pelas setas vermelhas representam a eluição das proteínas de interesse.....	77
Figura 29 –	SDS_PAGE 20% dos testes de clivagem (a) da U5-15K não mutada, (b) mutante U5-15K pc, (c) mutante U5-15K C66A, (d) mutante U5-15K S137A, (e) mutante U5-15K S144A, (f) mutante U5-15K S120A, (g) mutante U5-15K S126A. Em todas as figuras, a coluna 1 mostra o	

	<p>marcador de massa molecular, a coluna 2 uma alíquota da proteína purificada, 3 uma alíquota 12h após a purificação, 4 24h após a purificação, 5 36h após a purificação, 6 48h após a purificação, 60h após a purificação e 7 72h após a purificação.....</p>	78
Figura 30 –	<p>Posicionamento da Ser120, ponto azul identificado com uma seta vermelha, em relação ao C-terminal clivado, estrutura em vermelho.....</p>	80
Figura 31 –	<p>SDS-PAGE 20% do teste de clivagem (a) das mutantes U5-15K S137A, U5-15K S144A e (b) da U5-15K nativa<sup>64</sup> na presença de 6 <math>\mu\text{g mL}^{-1}</math> de quimostatina. Onde a coluna 1 representa o marcador de massa molecular, a coluna 2 uma alíquota da proteína 12h após a purificação, a 3 24h após a purificação, a 4 36h após a purificação, a 5 48h após a purificação, a 6 60h após a purificação, a 7 72h após a purificação.....</p>	81
Figura 32 –	<p>Alinhamento da U5-15K de <i>T. brucei</i> (ciano) com a U5-15K de humano (vermelho) dentro de sua estrutura cristalina (verde). A alça flexível é evidenciada em preto.....</p>	83
Figura 33 –	<p>Alinhamento entre os modelos de estrutura terciária das proteínas U5-15K (ciano) e U5-15K truncada (verde), com a alça flexível em vermelho</p>	84
Figura 34 –	<p>Cromatograma da purificação da U5-15k truncada. A curva vermelha representa a porcentagem de tampão B no gradiente de imidazol e a curva preta a absorbância da amostra a 280nm.....</p>	84
Figura 35 –	<p>Representação gráfica de uma curva de desenovelamento térmico de uma proteína globular.....</p>	85
Figura 36 –	<p>a) Curva de fluorescência por temperatura das proteínas U5-15K (preta) e U5-15K (vermelha) e b) curva da sua derivada em relação a temperatura para a obtenção da temperatura de desenovelamento (<math>T_m</math>).....</p>	86
Figura 37 –	<p>Gel SDS-PAGE 20% do teste de clivagem da U5-15K sem a alça. A imagem da direita se refere ao teste controle e o da esquerda ao teste da U5-15K truncada na presença de 6 <math>\mu\text{g mL}^{-1}</math> de quimostatina. Em ambos géis a coluna 1 possui marcador de massa molecular, 2 alíquota da proteína após a purificação, 3 alíquota após 12h da purificação, 4 24h após a purificação, 5 36h após a purificação, 6 48h após a purificação, 7 60h após a purificação.....</p>	87
Figura 38 –	<p>Desenho esquemático da cauda PTP.....</p>	88
Figura 39 –	<p>Cristal da proteína U5-15K S126A na condição 1 da tabela 9.....</p>	90
Figura 40 –	<p>Posição da Prp6 (em vermelho) de levedura no U4/U6.U5 tri snRNP.....</p>	90
Figura 41 –	<p>a) Mapa do vetor pC-PTP-NEO e b) inserção da cauda PTP no DNA</p>	

	genômico do <i>T. brucei</i> .....	91
Figura 42 –	a) Confirmação por <i>western blot</i> da expressão da U5-102K com a causa PTP, onde 1 é o marcador de massa molecular, 2 controle positivo do <i>western blot</i> , 3 forma procíclica do <i>T. brucei</i> 427 não transfectada e 4 3 forma procíclica do <i>T. brucei</i> 427 expressando a U5-102K b) <i>Western blot</i> da seleção das colônias positivas, onde 1 é o marcador de massa molecular e 2 a 13 são as 12 colônias testadas. c) Eletroforese em gel de agarose da inserção da cauda PTP no gene da U5-102K, onde 1 é o marcador, 2 3 forma procíclica do <i>T. brucei</i> 427 não transfectada e 3 3 forma procíclica do <i>T. brucei</i> 427 expressando a U5-102K. As bandas relativas à proteína U5-102K (a e b) e ao C-terminal do gene da U5-102K (c) estão indicadas pelas setas vermelhas.....	92
Figura 43 –	SDS-PAGE 12% da purificação dos parceiros de interação da U5-102K, onde as coluna 1 e 2 possuem uma alíquota das proteínas que não se ligaram à coluna, 3 uma alíquota das proteínas eluídas da coluna contendo IgG (primeira etapa da purificação), 4 alíquota das proteínas que não se ligaram a segunda coluna (anti proteína C), 5 alíquota das proteínas eluídas da coluna contendo anti-proteína C (segunda etapa de purificação) e 6 marcador de massa molecular.....	93
Figura 44 –	A coluna da esquerda se refere à identificação das proteínas purificadas pela técnica de PTP-tag na U5-102K e a coluna da direita se refere às proteínas identificadas na purificação pela técnica de PTP-tag no C-terminal da U5-15K <sup>79</sup> .....	94
Figura 45 –	Estrutura do U4.U6/U5 tri snRNP de levedura com o C-terminal da Prp6 em evidência (círculo azul).....	95
Figura 46 –	SDS-PAGE 10% do teste de expressão da U5-102K, clonada em pET28a. A coluna 1 representa o marcador de massa molecular, a coluna 2 uma alíquota 1h após a indução com 0,5mM de IPTG, 3 2h após a indução, 4 3h após a indução, 5 4h após a indução, 6 16h após a indução e 7 uma alíquota antes da indução. A banda de expressão da U5-102K está indicada com a seta vermelha.....	96
Figura 47 –	a) Gel SDS-PAGE da purificação da U5-102K, onde 1 representa o marcador de massa molecular, 2 as proteínas solúveis totais na amostra, 3 as proteínas que não interagiram com a resina, 4 as proteínas que foram eluídas com tampão sem imidazol e os poços 5, 6, 7, 8 e 9 representam as frações de coleta do ÄKTA correspondentes ao pico de eluição no gradiente de concentração de imidazol. b) <i>Western blot</i> da purificação da U5-102K onde 1 representa a porção de proteínas que não interagiram com a coluna, 2 as proteínas eluídas em tampão sem imidazol e os poços 3, 4, 5, 6 e 7 são as frações de coleta do ÄKTA correspondentes ao pico de eluição do gradiente de concentração de imidazol.....	96
Figura 48 –	Mapa de vetor pETDuet-1.....	97

Figura 49 –	SDS-PAGE do teste de co-expressão das proteínas U5-15K e U5-102K em pET DUET a) 10% para a visualização da U5-102K, onde as setas vermelhas apontam as bandas que podem corresponder à expressão da proteína. b) 15% para a visualização da U5-15K, onde a chave vermelha indica a região do gel onde a banda de expressão apareceria. Em ambos géis a coluna 1 contem marcador de massa molecular, 2 uma alíquota da colônia 1 antes da indução com IPTG, 2 2h após a indução e 3 16h após a indução, 5 alíquota da colônia 2 antes da indução, 6 2h após a indução e 7 16h após a indução, 8 alíquota da colônia 3 antes da indução com IPTG, 9 2h após a indução e 10 16h após a indução.....	98
Figura 50 –	a) Interação entre as proteínas U5-15K (vermelha) e a porção N-terminal da proteína U5-102K (azul) no tri snRNP de levedura. b) Interação entre as proteínas U5-15K (verde) e a porção N-terminal da proteína U5-102K (vermelha) no complexo B do spliceossomo humano.....	99
Figura 51 –	Porção N-terminal do alinhamento das sequências das proteína U5-102K de <i>T. brucei</i> , U5-102K de <i>H. sapiens</i> e Prp6 de <i>S. cerevisiae</i> feito no ClustalW. As regiões de hélice alfa estão grifadas em rosa. As regiões de interação com a U5-15K e Dim1 em <i>H. sapiens</i> e <i>S. cerevisiae</i> respectivamente, então marcadas em vermelho.....	100
Figura 52 –	Sequencia de aminoácidos da proteína U5-102K de <i>T. brucei</i> , a região em vermelho indica a sequencia do N-terminal referida nesse trabalho como U5-102K N-term.....	101
Figura 53 –	Gel SDS-PAGE 15 % do teste de expressão em <i>E. coli</i> BL21(DE3) do N-terminal da proteína U5-102k clonado em pET SUMO, onde a coluna 1 contem marcador de massa molecular, 2 alíquota da colônia 1 antes da indução com IPTG, 3 alíquota da colônia 1 16h após a indução, 4 alíquota da colônia 2 antes da indução com IPTG, 5 alíquota da colônia 2 16h após a indução, 6 alíquota da colônia 3 antes da indução com IPTG, 7 alíquota da colônia 3 16h após a indução, 8 alíquota da colônia 4 antes da indução com IPTG, 9 alíquota da colônia 4 16h após a indução e a seta vermelha indica a banda de expressão da proteína.....	101
Figura 54 –	a) SDS-PAGE 15% do teste de solubilidade do N-terminal da U5-102K e <i>western blot</i> das mesmas amostras, onde 1 é a alíquota das proteínas insolúveis, 2 a alíquota das proteínas solúveis, 3 o marcador de massa molecular e 4 o controle positivo do <i>western blot</i> feito com anticorpo contra a cauda de histidinas, b) SDS-PAGE 15% das frações do ÄKTA correspondentes ao pico da purificação por afinidade do N-terminal da U5-102K, onde 1 é o marcador de massa molecular e 2 a 6 são alíquotas das frações do pico de purificação.....	102
Figura 55 –	Distribuição do diâmetro em nanômetros da U5-15K, da U5-102K e da mistura de 1:1 das duas proteínas, como assinalado no topo de cada quadro.....	103



Figura 56 –	Gel nativo com 20% de poliacrilamida do teste de interação entre U5-15K e o N-terminal da U5-102K, sendo que a coluna 1 contém o marcador de massa molecular, 2 uma alíquota da U5-15K purificada, 3 uma alíquota do N-terminal da U5-102K purificado e 4 uma alíquota da mistura de 1:1 das duas proteínas.....	104
Figura 57 –	(a) Digestão do pJET_Prp43 com EagI antes da mutação, a seta vermelha indica a porção que codifica o C-terminal digerida . (b) Digestão do pJET_Prp43, proveniente de três colônias, com EagI após a mutação sítio dirigida.....	105
Figura 58 –	<i>Western blot</i> da confirmação da transfecção da cepa 427 da forma procíclica de <i>T. brucei</i> com pC-PTP-NEO_Prp43, onde a coluna 1 contém o marcador de massa molecular, 2 uma alíquota de <i>T. brucei</i> PCF 427 transfectado, 3 a cepa selvagem de <i>T. brucei</i> PCF 427 e 4 o controle positivo para o <i>western blot</i> . A seta vermelha indica a altura da banda da Prp43 com a cauda PTP.....	106
Figura 59 –	<i>Western blot</i> da seleção dos clones da cepa 427 da forma procíclica de <i>T. brucei</i> transfectado com pC-PTP-NEO_Prp43. A coluna 1 contém marcador de massa molecular, a 2 o controle positivo do <i>Western blot</i> , feito do extrato proteico da cultura de <i>T. brucei</i> antes da seleção na placa de 96 poços e as colunas de 3 a 7 são colônias selecionadas na placa de 96 poços.....	106
Figura 60 –	SDS-PAGE 12% da purificação dos parceiros de interação da proteína Prp43, onde 1 são as proteína purificadas da coluna com anti-proteína C, 2 as proteínas que não interagiram com com a matriz contendo anti-proteína C, 3 as proteínas purificadas da coluna contendo IgG, 4 as proteínas que não interagiram com a matriz contendo IgG e 5 o marcador de massa molecular.....	107
Figura 61 –	Identificação por espectrometria de massas das proteínas co-purificadas com a Prp43.....	108
Figura 62 –	Localização de domínio G-Patch pelo servidor SMART.....	109
Figura 63 –	Alinhamento entre as sequências da suposta proteína Ntr1 de <i>T. brucei</i> e da Ntr1 de levedura.....	110
Figura 64 –	Alinhamento entre as sequências das supostas proteínas Ntr1 de <i>T. brucei</i> , <i>T. cruzi</i> e <i>L. major</i> .....	111



## LISTA DE TABELAS

Tabela 1 –	Correspondência da nomenclatura das principais proteínas envolvidas no <i>splicing</i> de humano, levedura e <i>T. brucei</i> .....	41
Tabela 2 –	Nome e sequência dos oligonucleotídeos iniciadores usados nesse trabalho, vetores onde as proteínas foram clonadas e enzimas de restrição utilizadas.....	53
Tabela 3 –	Nome e sequência dos oligonucleotídeos iniciadores usados para as mutações sítio dirigidas, região mutada das proteínas e nomes dados a elas nesse trabalho.....	55
Tabela 4 –	Ciclagem da PCR usada para as mutações descritas.	56
Tabela 5 –	Variação da temperatura para hibridização das fitas de DNA resultantes das mutações 2 e 3 da tabela 3.....	57
Tabela 6 –	Relação dos vetores usados nesse trabalho e em que cepas foram transformados e os antibióticos de seleção.....	59
Tabela 7 –	Vetores transformados nas cepas de expressão e respectivos antibióticos de seleção.....	61
Tabela 8 –	Relação das proteínas purificadas e composição do tampão A.....	63
Tabela 9 –	Especificações dos anticorpos usados nos <i>Western Blots</i> desse trabalho....	65
Tabela 10 –	Soluções de cristalização dos kits comerciais utilizados, que resultaram em cristais promissores.....	89
Tabela 11 –	Proteínas identificadas por espectrometria de massas.....	93
Tabela 12 –	Proteínas identificadas por espectrometria de massas.....	108
Tabela 13 –	Composição de aminoácidos na sequência da Ntr1 em porcentagem.....	109



## LISTA DE ABREVIATURAS E SIGLAS

BS	branch site
CATT	Teste de aglutinação em cartão, em inglês
CTD	Domínio C-terminal
DLS	Espalhamento dinâmico de luz, em inglês
DNA	Ácido desoxirribonucleico, em inglês
DSF	Fluorimetria diferencial de varredura, em inglês
EM	Meio de Eletroporação, em inglês
GST	Glutathione S-Transferase
HAT	Tripanosomíase Humana Africana, em inglês
IFI	Imunolocalização por imunofluorescência indireta
LB	Luria-Bertani
mRNA	RNA mensageiro
NTC	Nineteen complex
NTD	Domínio N-terminal
NTR	nineteen related
OD600	Densidade óptica à 600nm
ORF	Quadro aberto de leitura, em inglês
PBS	Tampão fosfato-salino
PBST	Tampão fosfato-salino com Tween
PCR	Reação em cadeia da polimerase, em inglês
RDT	Teste de diagnóstico rápido, em inglês
RNA	Ácido ribonucleico, em inglês
RNAi	RNA de interferência
RNM	Ressonância Magnética nuclear, em inglês
RRM	motivo de reconhecimento de RNA, em inglês
SAXS	Espalhamento de raio-X a baixo ângulo, em inglês
SL	Spliced leader
RNA	splice site RNA
sn RNP	small nuclear ribonucleoprotein
snRNA	small nuclear RNA
SS	splice site
SUMO	small ubiquitin-like modifier
TAP	Tandem Affinity Purification
TbPCF	Forma procíclica do <i>Trypanosoma brucei</i>
TEMED	Tetramethylethylenediamine
Tm	Temperatura média
TRP	<i>Tetra-trico Peptide Repeats</i>
TYM	Tripticase-extrato de levedo-maltose
VHb	Vitreocilla Haemoglobin
WHO	Organização Mundial da Saúde, em inglês



## SUMÁRIO

1	Introdução.....	23
1.1	Tripanossomíase Africana Humana (Doença do Sono).....	23
1.2	Ciclo de Vida e Transmissão.....	27
1.3	Sintomas, Diagnóstico e Tratamento.....	30
1.4	Importância do estudo do <i>T. brucei</i> .....	31
1.5	<i>Splicing</i> do RNA mensageiro.....	32
1.6	<i>Cis</i> e <i>Trans-Splicing</i> .....	36
1.7	As <i>small ribonucleoproteins</i> (sn RNPs).....	39
1.8	U5-15K e U5-102K.....	45
1.9	Prp43.....	47
2	Objetivos.....	51
3	Materiais e Métodos.....	53
3.1	Análise bioinformática das sequências de aminoácidos das proteínas U5-102K e Prp43.....	53
3.2	Manipulação do DNA.....	53
3.2.1	Extração do DNA genômico.....	53
3.2.2	Clonagem.....	53
3.3	Mutação sítio dirigida.....	54
3.4	Manipulação das cepas de <i>E. coli</i> .....	57
3.4.1	Preparo das células competentes.....	57
3.4.2	Transformação das bactérias competentes.....	59
3.5	Manipulação da cepa 427 da forma procíclica de <i>T. brucei</i> (TbPF 427).....	60
3.5.1	Cultura.....	60
3.5.2	Transfecção.....	60
3.6	Obtenção das proteínas recombinantes.....	61
3.6.1	Expressão das proteínas em <i>E. coli</i> .....	61
3.6.2	Lise celular.....	62
3.6.3	Cromatografia de afinidade.....	63
3.7	<i>Western blotting</i> .....	64
3.8	Caracterização da atividade de autoclivagem da proteína U5-15K.....	65
3.8.1	Teste de clivagem das proteínas mutantes.....	65
3.9	Ensaio de interação entre as proteínas U5-15K e U5-102K.....	66
3.9.1	DLS.....	66
3.9.2	Gel Nativo.....	66

3.10	Ensaio de cristalização.....	67
3.10.1	Fluorimetria Diferencial de Varredura (DSF).....	67
3.10.2	Ensaio de cristalização.....	67
3.11	Caracterização dos parceiros de interação das proteínas U5-102K e Prp43.....	67
3.11.1	Preparo do extrato proteico.....	67
3.11.2	Purificação em Tandem.....	68
3.11.3	SDS-PAGE.....	70
3.11.4	Espectrometria de massas.....	70
4	Resultados e Discussão.....	73
4.1	Estudo da atividade de autoclivagem.....	73
4.1.1	Mutações sítio dirigidas.....	73
4.1.2	Obtenção das proteínas recombinantes.....	75
4.1.3	Teste de clivagem.....	78
4.2	Estudo estrutural da U5-15K.....	82
4.2.1	Mutação sítio dirigida.....	82
4.2.2	Obtenção da proteína recombinante.....	84
4.2.3	Fluorimetria diferencial de varredura (DSF).....	85
4.2.4	Teste de clivagem.....	87
4.2.5	Teste de cristalização.....	88
4.3	Estudo do subcomplexo formado pelas proteínas U5-15K e U5-102K.....	89
4.3.1	Purificação dos parceiros de interação.....	89
4.3.2	Obtenção da proteína U5-102K recombinante.....	95
4.3.3	Obtenção do N-Terminal da proteína U5-102K.....	99
4.3.4	Ensaio de interação.....	102
4.4	Identificação dos parceiros de interação da proteína Prp43.....	105
5	Conclusão.....	113
6	Perspectivas Futuras.....	115
	Referências.....	117



# 1 INTRODUÇÃO

## 1.1 Tripanossomíase Humana Africana (Doença do Sono)

A Tripanossomíase Humana Africana (*Human African Trypanosomiasis*, HAT), popularmente conhecida como Doença do Sono, é causada por um protozoário extracelular pertencente ao gênero *Trypanosoma* presentes no continente africano. O termo Doença do Sono abrange doenças causadas por duas subespécies de *Trypanosoma brucei*, o *T. brucei gambiense* e o *T. brucei rhodensiense*, que possuem padrões epidemiológicos, distribuição geográfica e formas de tratamento distintos. Diferenças essas que são importantes para o planejamento do controle das duas formas da doença.<sup>1</sup>

A forma Gambiense da Doença do Sono é caracterizada por longos períodos assintomáticos e é predominante na África central. Já a forma Rhodensiense, que atinge os países do sul da África, é mais agressiva e de progressão mais acelerada (Figura 3).<sup>2</sup>

Além da HAT, algumas espécies de *Trypanosoma*, *T. congolense*, *T. vivax* e *T. brucei* spp, são causadores da Tripanossomíase Animal Africana ou Nagana. Essa forma da doença é severa e muitas vezes fatal em animais domésticos, enquanto que em animais selvagens ela provoca apenas infecções leves.<sup>3</sup>

Muito antes da medicina europeia entrar em contato com essa tripanossomíase, já existiam relatos de regiões afetadas pela doença no sono. Os primeiros que se tem notícia datam de 1200, quando um geógrafo Árabe descreveu seu encontro com uma vila cujos habitantes “eram osso e pele e estavam dormindo”. Os sintomas, tanto iniciais quanto avançados, são documentados desde o comércio de escravos iniciado no século XVI.<sup>3</sup>

A Doença do Sono e a Nagana foram consideradas por muito tempo obstáculos para o desenvolvimento da África Sub-Sahariana, especialmente da zona rural. Ambas as doenças reduzem a força de trabalho disponível, limitam a oferta de carne, leite e derivados e capacidade de manter lavouras que sustentem a população.<sup>4</sup>

A proposta de que a transmissão da Nagana era causada pela picada de uma mosca da família *Glossinidae*, conhecida popularmente como Tsétsé, foi feita em 1852 pelo missionário escocês David Livingstone<sup>3</sup> e o *T. brucei* só foi descoberto como causador da doença em 1895 pelo patologista escocês David Bruce (Figura 1).<sup>5</sup>

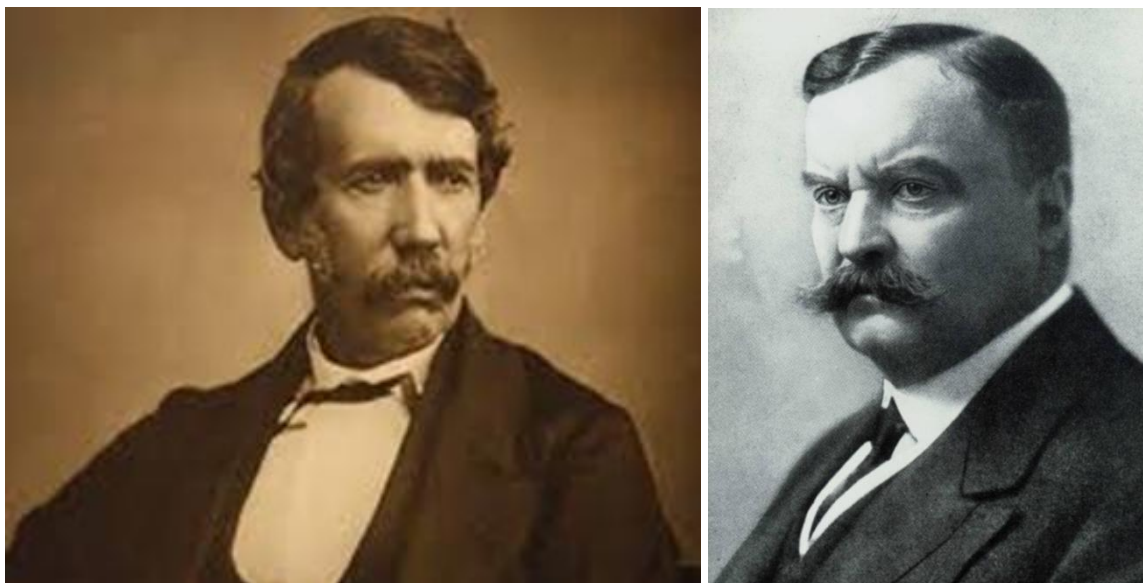


Figura 1 – Missionário e explorador escocês David Livingstone (esquerda) e do patologista escocês responsável pela descoberta do *T. brucei* Sir David Bruce (direita).  
Fonte: STEVERDING<sup>3</sup>.

Apesar desses relatos, a tripanossomíase africana só recebeu atenção do poder colonial no início do século 20, quando três grandes epidemias assolaram a África equatorial. A mais severa aconteceu na Uganda, na década de 1910. Em menos de 5 anos, 200 mil habitantes, um terço da população da região costeira das ilhas do Lago Victoria, haviam morrido devido a Doença do Sono.<sup>6</sup>

Em resposta a essas epidemias, os poderes imperiais da época organizaram expedições de cientistas, que contribuíram significativamente para a classificação dos protozoários causadores da doença, bem como seu ciclo de vida e forma de transmissão. A partir dessas descobertas, e da cooperação da comunidade científica, foram traçados planos para o combate da epidemia.<sup>6</sup>

Na época, as medidas de controle se baseavam em separar a população da Tsétsé, medida tomada pelos britânicos, que deslocou milhares de pessoas de suas terras e isolou os infectados em campos especiais. Os britânicos também promoveram a caça de animais silvestres, entendidos como reservatórios do *T. brucei* e tentaram eliminar o parasita usando Atoxyl, a primeira droga usada para tratar a Doença do Sono. Medidas que, contribuindo ou não para a redução dos números de casos, foram interrompidas durante a primeira guerra mundial e demoraram anos para serem retomadas.<sup>6</sup>

Até os anos 1960, a maioria dos países onde a Doença do Sono é endêmica havia conquistado sua independência e sem contar com o suporte financeiro dos colonizadores e

com a crescente instabilidade política e econômica, viram as políticas de controle das doenças tropicais serem deixadas de lado.<sup>7</sup>

Na década de 1970, o crescimento no número de casos registrados da Doença do Sono foi agravado pela proibição, necessária devido a questões ambientais, de aplicação de alguns inseticidas usados no controle do vetor, iniciando a epidemia mais recente, que afetou principalmente Angola, Congo, Sudão e Uganda.<sup>7</sup>

Entre as décadas de 1980 e 1990, o número de casos existentes foi estimado em cerca de 300 mil, entretanto, diversos territórios afetados não foram incluídos nas estimativas. Essa situação finalmente chamou a atenção de autoridades em saúde nos países afetados e a Organização Mundial da Saúde (WHO, World Health Organization) coordenou uma campanha envolvendo os governos locais e organizações não governamentais parceiras para a implementação de medidas de controle da doença (Figura 2).<sup>8-9</sup>

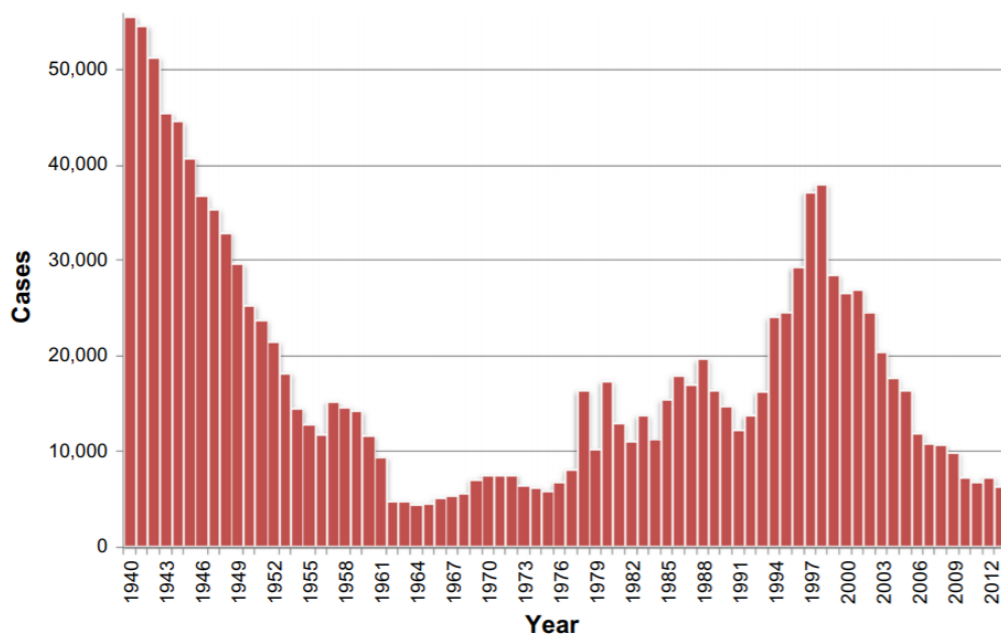


Figura 2 – Casos totais de Tripanossomíase Humana Africana reportados a Organização mundial da saúde, de 1940 a 2013.

Fonte: FRANCO<sup>8</sup>

Entre essas medidas implementadas se destacam o reforço do tratamento e diagnóstico, que contou com a ajuda de companhias farmacêuticas, como a Bayer e Sanofi, que trabalhavam com o desenvolvimento de medicamentos contra a Doença do Sono, e conscientização da população da zona rural, áreas mais afetadas.<sup>8-9</sup>

Todos esses esforços culminaram na redução dos casos reportados, sendo que em 2012 esse número ficou abaixo de 10 mil e em 2016 abaixo de 2500 (Figura 3). A meta da

Organização Mundial da Saúde é eliminar a Doença do Sono como um problema de saúde pública até 2020 e eliminação da transmissão da doença até 2030<sup>10</sup>.

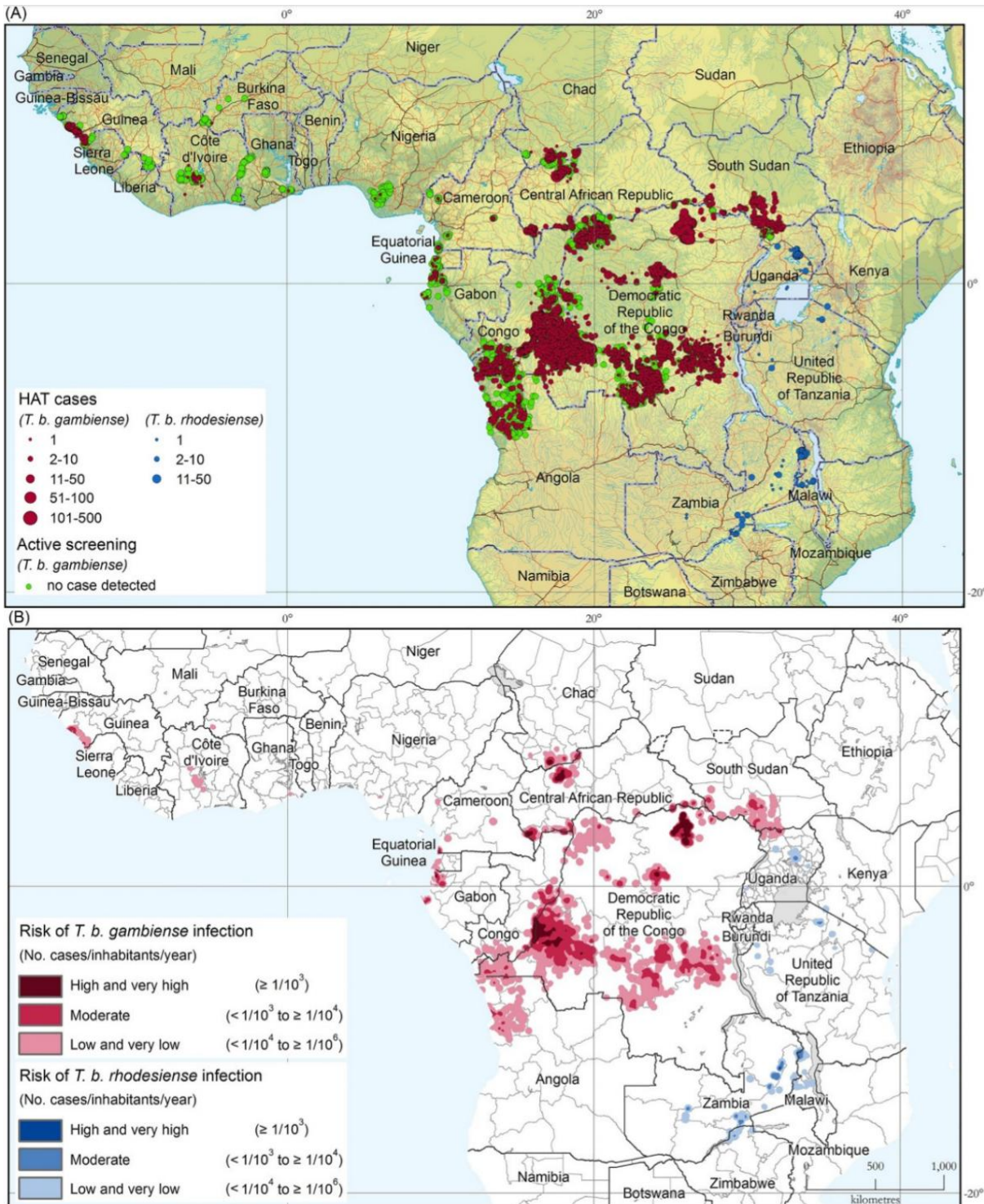


Figura 3 – (A) Distribuição geográfica dos casos reportados da Tripanossomíase Humana Africana de 2012 a 2016. (B) Representação das áreas de risco, onde as cores mais escuras representam um risco maior de contrair a doença e as cores mais claras um risco menor. Dados de 2012-2016.

Fonte: FRANCO<sup>10</sup>

Os dados mais recentes mostram que a meta de menos de 2000 casos reportados até 2020 foi atingida, sendo que 2017 o número ficou abaixo de 1500. Entretanto ainda existem áreas endêmicas onde a taxa de contágio é superior a 1 caso a cada 10000 pessoas. Essas são geralmente áreas de difícil acesso, ou áreas que tiveram epidemias recentes de outras doenças, como é o caso da Guiné, que teve o programa de controle da Doença do Sono interrompido entre 2014 e 2015 devido aos surtos de ebola.<sup>11</sup>

Apesar do controle obtido com décadas de vigilância e investimento, mudanças climáticas, crescimento populacional, urbanização e mesmo instabilidades políticas e sociais podem afetar a distribuição geográfica da Tripanossomíase Humana Africana.<sup>8</sup>

A temperatura ambiente tem uma grande influência sobre o desenvolvimento da Tsetsé e o aumento das médias anuais nas últimas décadas pode mudar a área viável à criação desse vetor. Estudos mostram que as taxas atuais do aumento de temperatura podem provocar um deslocamento dos criadouros de Tsetsé para o sul e leste africano, expondo milhares de pessoas à Doença do Sono.<sup>12</sup>

Quanto a crescente urbanização de alguns países africanos, o desmatamento e mudança no estilo de vida da população afeta também os hábitos da Tsetsé, que acaba por buscar áreas onde a disponibilidade de alimento é maior. O que torna as periferias das áreas urbanas particularmente favoráveis para sua proliferação.<sup>13</sup>

## 1.2 Ciclo de vida e transmissão

Como já foi dito, a tripanossomíase africana é causada por parasitas unicelulares pertencentes à família Trypanosomatidae, ordem Trypanosoma, espécies *Trypanosoma brucei*, que vivem no sangue, sistema linfático e espaço intersticial dos órgãos. Tanto na doença causada pela subespécie *T. brucei gambiense* quanto pela subespecie *T. brucei rhodensiense*, o parasita é transmitido por um inseto hematófago do gênero *Glossina*, conhecida popularmente como Tsetsé. Animais selvagens e domésticos servem como reservatório do parasita. Nas espécies de *Glossina*, tanto a fêmea quanto o macho se alimentam de sangue e podem carregar o *T. brucei* (Figura 4).<sup>14</sup>





Figura 4 – Forma do *Trypanosoma brucei* em esfregaço de sangue de um paciente com Tripanossomíase Humana Africana.

Fonte: KENNEDY <sup>15</sup>

O ciclo de vida do *T. brucei* tem dois estágios, um envolvendo o hospedeiro invertebrado (Tsetse) e outro envolvendo o hospedeiro mamífero (o homem no caso da Doença do Sono e gado e animais silvestres no caso da Nagana). O estágio do ciclo que se passa na Tsetse tem início quando essa se alimenta de um animal contaminado (representado com o número 5 na figura 5).<sup>15</sup>

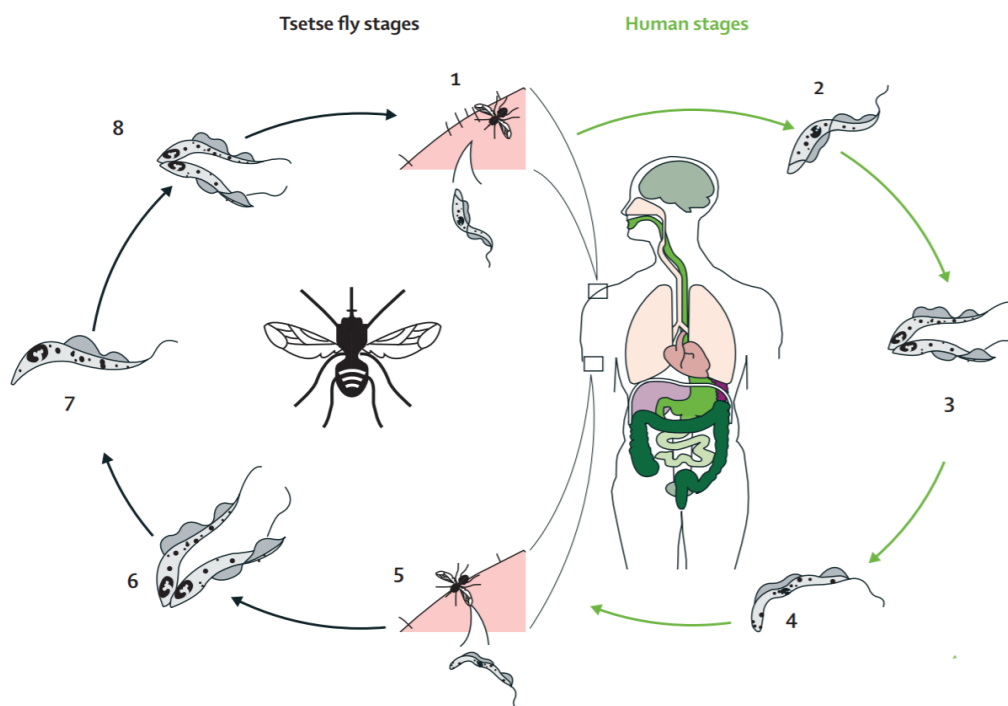


Figura 5 – Ciclo de vida do *Trypanosoma brucei*, mostrando a entrada da forma tripomastigota metacíclica na corrente sanguínea do hospedeiro mamífero, o desenvolvimento do parasito no organismo do hospedeiro mamífero, a ingestão da forma tripomastigota sanguínea pelo inseto vetor e desenvolvimento do parasito no organismo do vetor.

Fonte: Adaptada de KENNEDY <sup>15</sup>

No intestino da Tsetse, a forma tripomastigota sanguínea do parasita passa por uma série de mudanças anatômicas e bioquímicas dando origem a forma tripomastigota procíclica (número 6 na figura 5). Essa forma se multiplica, por divisão binária e deixa o intestino, se transformando na forma epimastigota (número 7 na figura 5).<sup>15</sup>

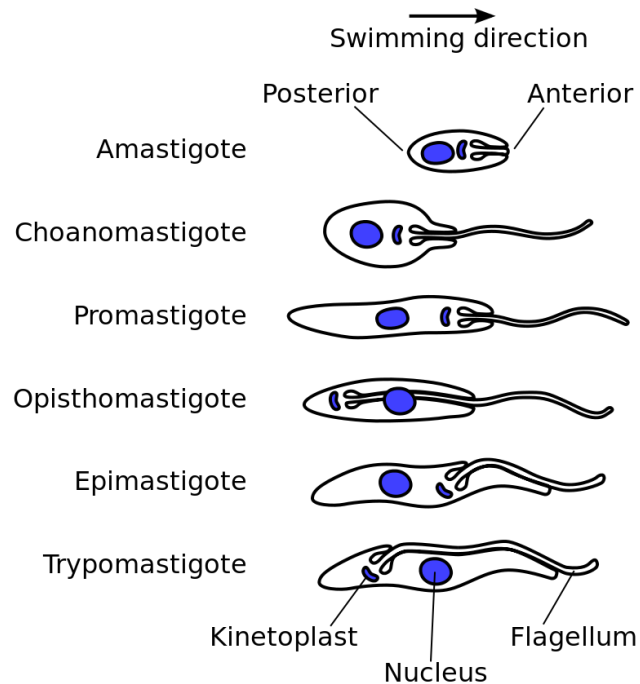


Figura 6 – As 6 principais morfologias dos tripanossomatídeos, sendo a amastigota a forma comum durante a fase intracelular no hospedeiro mamífero, a promastigota a forma comum no inseto vetor, a tripomastigota a forma sanguínea do hospedeiro mamífero e a forma infectante encontrada no inseto vetor.

Fonte: HAORE<sup>94</sup>.

A forma epimastigota se aloja nas glândulas salivares do inseto vetor, onde se multiplicam e se transformam em tripomastigota metacíclica, altamente infecciosa (número 8 na figura 5), que é transferida para o hospedeiro mamífero quando esse é picado pela Tsetse (número 1 na figura 5).<sup>15</sup>

No mamífero, o parasita se transforma em tripomastigota sanguíneo e é transportado para outras regiões do corpo (número 2 na figura 5), como nódulos linfáticos, baço, coração, fígado, olhos e sistema endócrino, onde se multiplicam (número 3 na figura 5). As células que permanecem no sistema sanguíneo são ingeridas pela Tsetse e continuam o ciclo.<sup>15</sup>

### 1.3 Sintomas, diagnóstico e tratamento

O primeiro sinal da infecção por tripanossomíase africana é o aparecimento de uma lesão na região da picada chamada de cancro tripanossomal. Essa lesão se desenvolve de 5 a 15 dias após a picada.<sup>16</sup>

O quadro clínico da doença do sono possui dois estágios, um inicial, chamado de hemolinfático, e um tardio, encefalítico. Esses estágios variam em duração dependendo da subespécie do parasito. No caso da doença do sono causada pelo *T. brucei gambiense*, a transição do primeiro para o segundo estágio pode durar até anos, já na tripanossomíase rhodensiense, essa transição é mais rápida, podendo durar alguns meses.<sup>16</sup>

O estágio inicial ocorre cerca de 1-3 semanas após a picada e é caracterizado por episódios de febre, que duram por volta de uma semana, e linfadenopatia generalizada (inflamação e inchaço dos gânglios linfáticos). Os sintomas iniciais tendem a ser não específicos, como mal-estar, dor de cabeça, perda de peso e fraqueza. O estágio tardio ocorre quando o parasita cruza a barreira hematoencefálica e é caracterizado por uma série de sintomas que podem variar entre os pacientes. Os sintomas gerais reportados contam com anormalidades psiquiátricas (mudança de comportamento, irritabilidade, tendências suicidas, etc), motoras, sensoriais e distúrbios do sono.<sup>17</sup>

Vale ressaltar que o *T. brucei* possui glicoproteínas de superfície responsáveis pela variação antigênica encontrada nas infecções por *T. brucei*. Durante o processo infeccioso, há alterações no silenciamento dos genes que codificam tais glicoproteínas, o que permite que o *T. brucei* evite a resposta imune do hospedeiro e torna difícil a produção de vacinas.<sup>18</sup>

É importante que o diagnóstico dessa tripanossomíase seja feito o mais cedo possível, uma vez que o tratamento disponível para os estágios avançados possui uma alta toxicidade. No caso da doença causada pelo *T. brucei rhodensiense*, o diagnóstico pode ser feito através de um esfregaço do sangue periférico, devido a sua alta parasitemia. No caso do *T. brucei gambiense* um dos primeiros métodos utilizados foi o teste de aglutinação em cartão (Card Agglutination Trypanosomiasis Test, CATT), um teste sorológico rápido desenvolvido em 1978. Apesar de barato, esse teste apresenta casos de falso positivo, que o tornam ineficiente em áreas onde a prevalência da doença é baixa, além de haver limitações para seu uso em instalações de saúde em locais remotos devido a ausência de agitadores, energia elétrica e refrigeradores.<sup>15,19,20</sup>

Para solucionar os problemas com o CATT, foram desenvolvidos dois testes rápidos para diagnóstico da HAT ( Rapid Diagnostic Test, RDT) que detectam anticorpos específicos contra duas glicoproteínas de superfície do *T. brucei*. Esses testes começaram a ser usados



em 2010 e apesar das melhorias, a produção dos antígenos nativos usados ainda é desafiadora e depende de processos caros e arriscados. Recentemente foi desenvolvido um RDT que usa antígenos recombinantes, se mostrando tão específico quanto os testes anteriores.<sup>95-96</sup>

Esses métodos funcionam para os estágios iniciais da doença. Para o estágio tardio, independente da subespécie de *T. brucei* causadora da doença, o único diagnóstico disponível ainda é o exame do fluido cerebrospinal e uma punção lombar é necessária.<sup>19</sup>

Já o tratamento da doença do sono, ainda depende em muito de medicamentos desenvolvidos anos atrás e que apresentam toxicidade elevada especialmente no estágio tardio. Para o estágio inicial da infecção por *T. brucei gambiense*, o tratamento consiste na administração intravenosa ou intramuscular de Pentamidina. Para o estágio inicial do *T. brucei rhodensiense*, é usada a Suramina, que pode ter como efeitos adversos disfunção renal, neuropatia periférica e reação anafilática.<sup>19</sup>

O único tratamento disponível para estágio tardio da doença do sono causada por *T. brucei rhodensiense* é a administração intravenosa de Melarsoprol, usado desde 1949. Esse medicamento é efetivo, mas sua administração é dolorosa e os efeitos adversos incluem uma grave encefalopatia reativa pós-tratamento em 5-10% dos pacientes.<sup>19</sup>

O tratamento do estágio tardio da doença do sono causada pelo *T. brucei gambiense* é menos tóxico que o *T. brucei rhodensiense* e conta com a combinação de dois medicamentos, a Eflornitina e o Nifurtimox. A Eflornitina é administrada de forma intravenosa e o Nifurtimox oralmente. Recentemente, novos medicamentos foram desenvolvidos para esse estágio e ainda estão em fase de teste. Um exemplo é o Fexinidazol, administrado por via oral. Testes mostraram que esse medicamento tratou com sucesso 91% dos pacientes, em contraste com 98% da taxa de sucesso do tratamento atual.<sup>19,21</sup>

Tendo em conta os aspectos aqui descritos, se mostra de grande importância o desenvolvimento de novos medicamentos, com maior eficácia no tratamento das diversas formas e subespécies de *T. brucei*, e com menores efeitos colaterais, bem como ferramentas de diagnósticos mais baratas e que possam ser levadas às áreas carentes.

#### **1.4 Importância do estudo do *T. brucei***

Além da tripanossomíase africana humana, também são doenças causadas por tripanossomatídeos a Doença de Chagas, endêmica em 21 países da América Latina, incluindo o Brasil, e as leishmanioses, cujos casos estão distribuídos pelo globo, sendo que aproximadamente 90% dos registros se concentram em 6 países, que incluem o Brasil<sup>22-23</sup>

A Doença de Chagas é causada pelo *Trypanossoma cruzi* e a transmissão se dá pelo contato com as fezes contaminadas do inseto vetor, os hemípteros do gênero *Triatoma*, com o sangue do hospedeiro mamífero. Já as leishmanioses são causadas por cerca de 20 espécies de *Leishmania* e a transmissão é feita por insetos do gênero *Phlebotomus* e *Lutzomyia*. Assim como a Doença do Sono, os tratamentos disponíveis são escassos, podem causar efeitos adversos graves e estão em uso há quase 50 anos, resultando no surgimento de formas resistentes.<sup>22,24</sup>

Recentemente, uma nova espécie de parasita semelhante aos organismos do gênero *Crithidia*, pertencente à família *Trypanosomatidae*, foi descoberta em pacientes do estado de Sergipe. Estudos indicam que a nova espécie está envolvida com a manifestação atípica semelhante à leishmaniose visceral, resistente aos tratamentos usados contra as espécies de *Leishmania*. Além disso, essa descoberta sugere a existência de outros potenciais vetores da doença, uma vez que a *Crithidia fasciculata*, parasita monoxênico de maior semelhança genética com a nova espécie, infecta apenas insetos dos gêneros *Culex* e *Anopheles*.<sup>25</sup>

Nos últimos anos houve avanços no controle dessas doenças e uma retomada do desenvolvimento de medicamentos, mas seu sucesso ainda depende do conhecimento do ciclo de vida do parasita. Devido à conservação dos processos biológicos existente entre os tripanossomatídeos e a maior acessibilidade laboratorial, o *Trypanosoma brucei* é muito usado como modelo no estudo dos mesmos. Isso ocorre, pois o cultivo de suas diferentes formas é possível *in vitro* e existem ferramentas genéticas para a caracterização da função de proteínas como expressão por indução e silenciamento gênico por interferência do RNA (RNAi).<sup>26</sup>

Além disso, o estudo dos tripanossomatídeos é interessante, pois esses organismos possuem mecanismos bioquímicos únicos e conservados entre si como a transcrição policistrônica e processamento da maioria do seu pré-RNA mensageiro por *trans-splicing*, objetivo central desse trabalho.<sup>27</sup>

## 1.5 *Splicing* do RNA mensageiro

A maior parte do RNA mensageiro (mRNA) dos eucariotos é sintetizada na forma de um precursor, pré-mRNA, e deve passar por uma série de modificações antes de ser transportada para fora do núcleo para que a tradução ocorra (figura 7). Uma dessas modificações é a retirada de regiões não codificantes do gene, chamadas de íntrons, e

subsequente ligação das regiões codificantes, os éxons. Esse processo recebe o nome de *splicing*.<sup>28-29</sup>

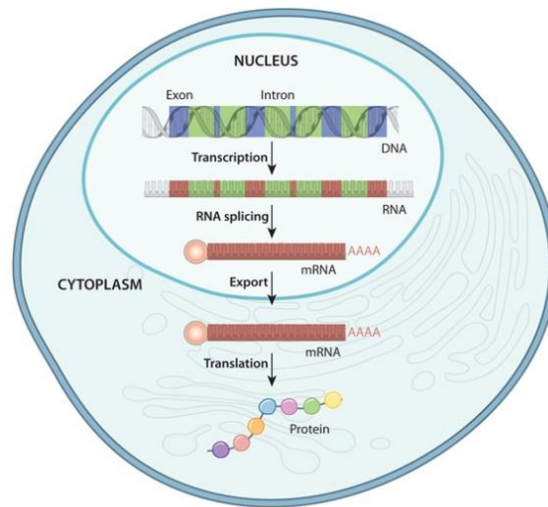


Figura 7 – Representação gráfica simplificada da síntese proteica em eucariotos, desde a transcrição do DNA em pré-RNA mensageiro até a tradução do mRNA maduro em proteína.  
Fonte: NATURE EDUCATION<sup>66</sup>

De uma forma simplificada, os íntrons são delimitados por pequenas seqüências conservadas na região 5' (5' *splice site* ou 5' ss), na região 3' (3' *splice site* ou 3' ss), como mostrado na figura 8. Essas regiões, juntamente com o *branch site* (BS) ou *branch point* (BP), são identificadas pelo spliceossomo durante o processo de *splicing*. O BS é normalmente localizado de 18 a 40 nucleotídeos antes do 3'ss e sua seqüência é conservada na levedura *Saccharomyces cerevisiae* e em mamíferos. Em leveduras, essa seqüência é altamente conservada (UACU AAC) enquanto que em mamíferos ela é degenerada (YNCURAY).<sup>29-30</sup>



Figura 8 - Ilustração da posição do sítio de *splicing* 5', 3' (5'SS e 3'SS), flanqueando o íntron, e *branch point* (BP) em uma molécula de pré RNA mensageiro.  
Fonte: Adaptada de WILL<sup>29</sup>

Em trypanossomatídeos, o mapeamento das seqüências do BS mostraram que as mesmas se encontravam mais distantes do 3'ss se comparadas às encontradas em outros organismos e nenhuma seqüência consenso foi determinada.<sup>31</sup>

A excisão dos íntrons e a ligação dos éxons ocorrem por meio de duas reações de transesterificação, onde o grupamento 2'OH da adenosina do BP, representada por um A na figura 8, realiza um ataque nucleofílico no 5'ss, o que resulta na clivagem do sítio 5' e ligação

entre 5'ss e BS, representado pela ligação entre A e p na figura 9. A segunda reação se dá com o ataque do 3'ss pelo grupamento 3'OH da porção 5' do éxon, E1 na figura 9, o que leva á ligação da região 5' de um éxon com a 3' do éxon adjacente, como mostrado na figura a seguir, formando o mRNA maduro e liberando o íntron.<sup>32</sup>

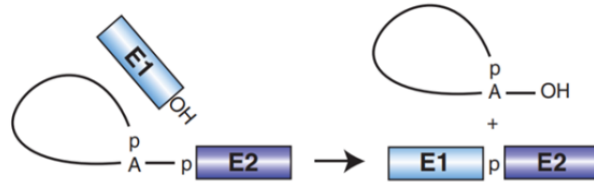


Figura 9 – Resultado da primeira (esquerda) e segunda (direita) reação de transesterificação para a formação do RNA mensageiro maduro.

Fonte: Adaptada de WILL<sup>29</sup>

Quando os éxons a serem unidos fazem parte da mesma molécula de pré-mRNA, o processo é denominado *cis-splicing* e tanto 5'ss quanto 3'ss fazem parte do mesmo íntron. Ao final das reações de transesterificação, o íntron é liberado na forma de laço ou *lariat*. Já o *trans-splicing* é uma reação intermolecular onde a molécula de pré-mRNA é unida a uma pequena molécula de RNA chamada *spliced leader RNA* (SL RNA) e o íntron é liberado na forma de um Y (figura 10).<sup>33-34</sup>

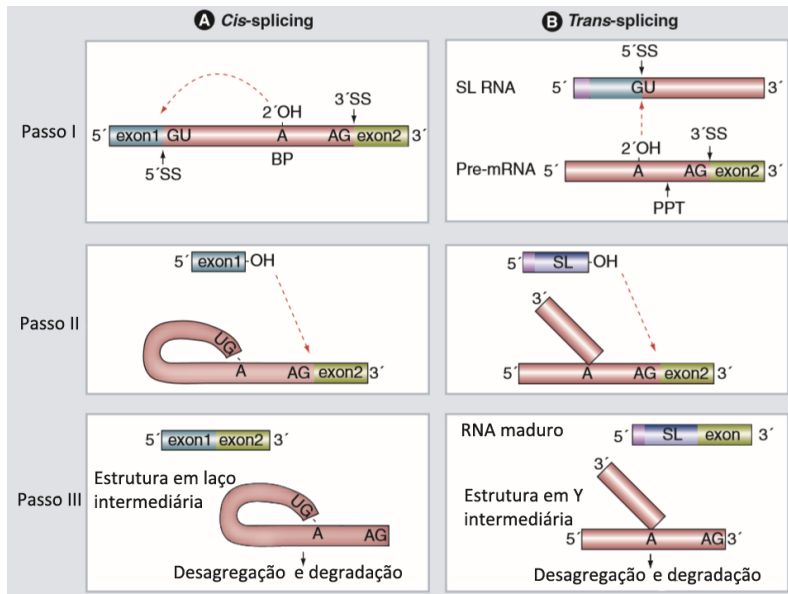


Figura 10 – Esquema mostrando as principais diferenças entre as moléculas de RNA envolvidas no *cis* e *trans-splicing*, como os éxons pertencentes ao mesmo pré-RNA mensageiro (passo I) e íntron excisado na forma de um laço (passo III) no caso do *cis-splicing* e a presença do SL RNA (passo I) e íntron excisado na forma de Y (passo III) no caso do *trans-splicing*.

Fonte: MICHAELI<sup>34</sup>

O SL RNA, presente em kinetoplastida, dinoflagelados, poríferos, cnidários, platelmintos, crustáceos, rotíferos, ctenóforos, quetognatas, nemátodos e urocordados, possui de 40 a 140 nucleotídeos dependendo do organismo e uma estrutura muito similar entre si, composta de 3 *stem-loops*. Em tripanossomatídeos a sequência SL possui 39-41 nucleotídeos seguidos de um íntron de tamanho variável que pode se dobrar em 2 *stem-loops* separados por uma região de fita simples. A essa região se liga o anel formado pelas proteínas Sm (figura 11).<sup>35</sup>

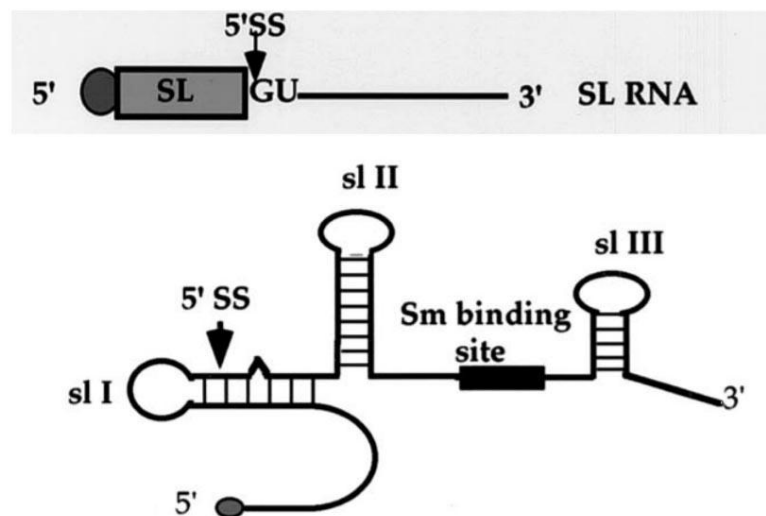


Figura 11 – Estrutura secundária do SL RNA, mostrando a localização do 5'ss, no stem-loop I mais próximo à extremidade 5', e do sítio de ligação das proteínas Sm, entre os stem-loops II e III.  
Fonte: Adaptada de LIANG<sup>35</sup>

Tanto no *cis* quanto no *trans-splicing*, as reações de transesterificação são catalisadas pelo spliceossomo, uma maquinaria dinâmica que se rearranja em cada passo da reação através da interação ordenada de cinco ribonucleoproteínas (*small nuclear ribonucleoproteins*, snRNPs) denominadas U1, U2, U4/U6 e U5 e inúmeros fatores não snRNPs.<sup>33-34</sup>

As snRNPs, por sua vez, são complexos compostos de pequenos RNAs ricos em uridina (U snRNAs) ligados a proteínas específicas de cada snRNP e proteínas comuns (proteínas Sm, B/B', D<sub>1</sub>, D<sub>2</sub>, D<sub>3</sub>, E, F e G). Assim como o SL RNA, os snRNAs possuem um sítio chamado Sm (AAUUUUUGA), em torno do qual essas proteínas comuns se agrupam, formando um anel (Figura 12).<sup>36-38</sup>

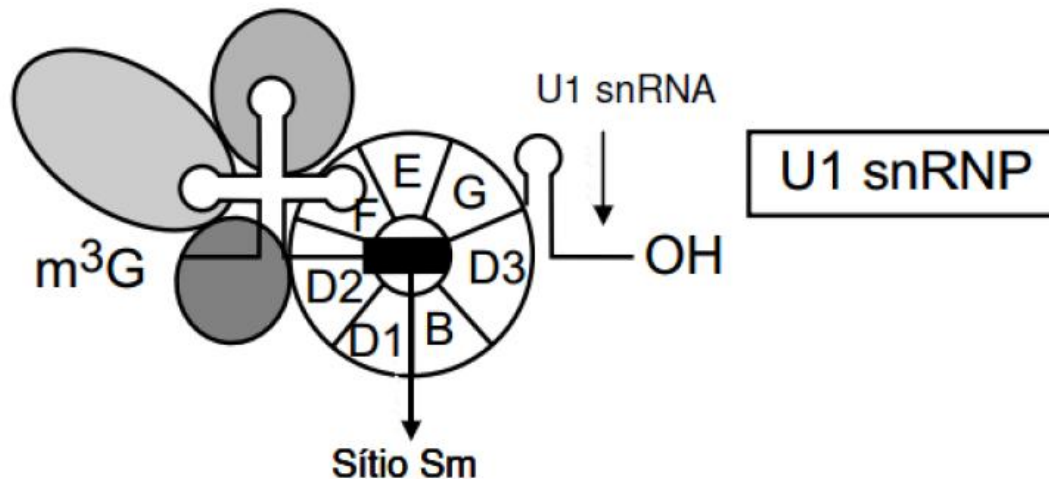


Figura 12 – Esquema da formação do complexo U1 snRNP de mamíferos, mostrando o U1 snRNA, as proteínas Sm, comuns às snRNPs, ao redor do sítio Sm do U1 snRNA e as proteínas específicas.  
Fonte: WILL<sup>29</sup>

Todas as proteínas Sm contém uma sequência assinatura, o domínio Sm, formado por dois segmentos conservados, Sm1 e Sm2, interrompidos por uma região não conservada de tamanho variável. Esse domínio também é encontrado nas proteínas LSm (Sm-like) associadas ao U6 snRNA (e ao SL RNA) e é essencial para a biogênese e função das snRNPs. Essas proteínas fazem parte do sistema de localização nuclear das snRNPs e influenciam na integração de algumas proteínas específicas aos respectivos snRNPs.<sup>36-37</sup>

Em tripanossomatídeos, mais especificamente nas partículas U2 e U4, algumas das proteínas Sm são distintas daquelas encontradas em leveduras e mamíferos. Na partícula U2, as proteínas Sm B e D3 são substituídas pela U2-15K (SSm2-1) e U2-16,5K (SSm2-2) respectivamente, e na partícula U4, a Sm D3 é substituída pela SSm-4.<sup>34</sup>

## 1.6 Cis e Trans-Splicing

O *cis-splicing* tem início quando o snRNA da partícula U1 interage, por pareamento de bases, com o 5'ss do íntron. Essa interação é estabilizada pelas proteínas específicas da U1 snRNP. Ao mesmo tempo, há o reconhecimento da sequência do *branch site* (BS) por pareamento com o snRNA da partícula U2 snRNP. Juntos, os complexos U1-U2 e mRNA formam o complexo A. Ao complexo A se associa o complexo U4/U6.U5 tri-snRNP, pré-formado pelas partículas U4/U6 snRNP e U5 snRNP. Esta associação, complexo A e U4/U6.U5 tri-snRNP, resulta no complexo B (figura 13).<sup>39</sup>



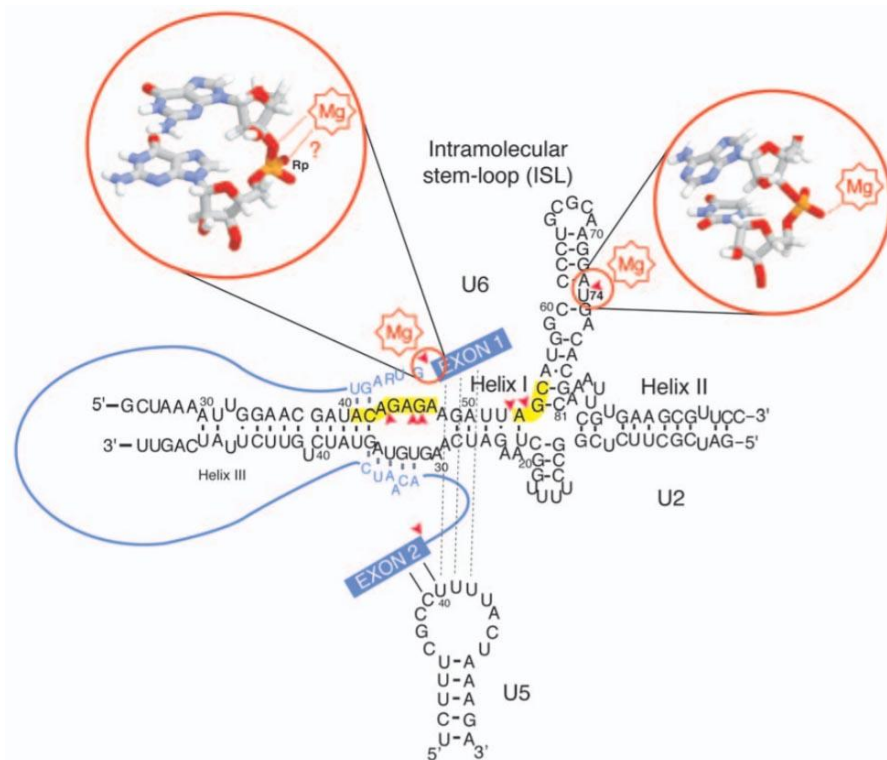


Figura 14 – Representação das interações entre U6 snRNA, U2 snRNA e o pré-mRNA (em azul) dentro do core catalítico do complexo B\*.  
Fonte: VALADKHAN<sup>40</sup>

Nesse passo, a adenosina do *branch site* (BS-A) ataca nucleofílicamente o 5'SS, resultando em um éxon 5' clivado e um éxon 3' ligado ao íntron em forma de laço, formando o complexo C (Figura 15). Logo após o primeiro passo catalítico, a hélice U2/BS, onde a adenosina do BS está ligada covalentemente à extremidade 5' do íntron, permanece no centro do complexo C até ser deslocada para fora do centro catalítico durante a ativação do complexo C em C\* mediada pela helicase Prp16. O complexo C\* catalisa o segundo passo do *splicing*. Em ambos os passos catalíticos, o correto posicionamento das extremidades 3' e 5' dos éxons é feita pelo loop I do U5 snRNA.<sup>42-43</sup>

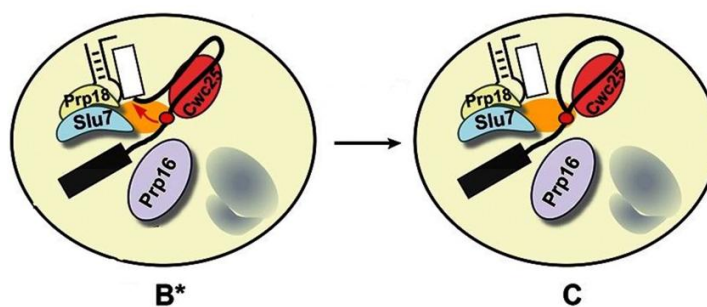


Figura 15 – Representação do ataque nucleofílico da adenosina do BS no complexo B\*, indicado pela seta vermelha, e formação do complexo C.  
Fonte: FICA<sup>32</sup>



A ligação dos éxons resulta no complexo P, de onde o RNA maduro é liberado pela proteína Prp22. Esse passo resulta no complexo ILS. A liberação do íntron em forma de laço é feita pela proteína Prp43, e as proteínas Snu144 e Brr2 promovem a desenovelamento do snRNA U4/U6, seguido da dissociação do complexo e reciclagem dos componentes spliceossomais U2, U5 e U6 (figura 13).<sup>44</sup>

No *SL trans-splicing*, a transferência da sequência SL ao pré-mRNA é análoga à retirada do íntron no *cis-splicing*, passando pelas mesmas reações de transesterificação. No entanto o pareamento do SL RNA com a extremidade conservada 5' do U1 snRNA não é necessária para a reunião das partículas que compõem o spliceossomo.<sup>45</sup>

Em nemátodos, foi visto que o SL RNA possui uma estrutura de snRNP, com proteínas Sm que formam um anel em torno do sítio Sm e proteínas específicas, como as SL175 e SL30. Essas proteínas fazem a ponte entre o 5'ss do SL RNA e o BS do éxon, papel que no *cis-splicing* é desempenhado pelas proteínas da U1 snRNP. Em tripanossomatídeos, nenhuma proteína específica do SL RNP foi descoberta, e ainda não se sabe ao certo como o SL RNP ativa seu próprio 5'ss. Os passos restantes da reação transcorrem como no *cis-splicing*.<sup>46-47</sup>

## 1.7 As small ribonucleoprotein (snRNPs)

A partícula U1 snRNP humana é formada por um snRNA (U1 snRNA), sete proteínas Sm (SmB/SmB', SmD1, SmD2, SmD3, SmE, SmF e SmG) e três proteínas específicas (U1-70K, U1-A e U1-C). Já a de levedura possui um snRNA mais complexo e se associa com fatores que não foram encontrados em humano (Prp39, Snu71, Prp40, Prp42, Snu56, Nam8, Urn1 e Prp5), mas conserva o motivo dedo de zinco (*Zinc finger*) da U1-C (Yhc1 em levedura) e a região 5' de fita simples do U1 snRNA (Figura 16).<sup>48-49</sup>

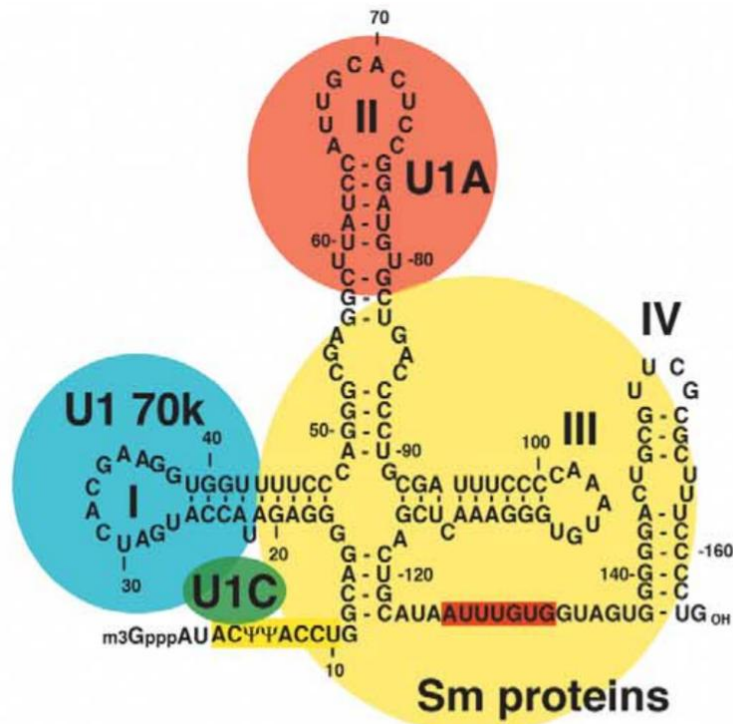


Figura 16 – Representação esquemática da partícula U1-snRNP humana, mostrando o anel formado pelas proteínas SM representado pelo círculo amarelo, o sítio SM do U1 snRNA selecionados em vermelho, as proteínas específicas U1C, em verde, U1 70K, em azul, e U1A, em vermelho, interagindo com o U1 snRNA.

Fonte: MUTO<sup>49</sup>

Em humanos, o motivo de reconhecimento de RNA (RNA recognition Motif, RRM) da proteína U1-A interage com o *stem-loop* II do U1 snRNA e o RRM da U1-70K interage com o *stem-loop* I. O N-terminal da U1-70K também interagem com as interfaces dos subcomplexos formados pelas proteínas SmD2/SmF e SmB/SmD3, estabilizando a interação da U1-C com a SmD3. A proteína U1-C, por sua vez, faz o contato da extremidade 5' do U1 snRNA e o 5'ss do íntron.<sup>50</sup>

Em tripanossomatídeos, foram identificadas proteínas específicas da partícula U1 sendo elas U1-A, U1-C, U1-70K e uma proteína específica do organismo denominada U1-24K. Essa partícula é necessária para o *cis-splicing* dos pré-mRNAs que codificam para a PAP (*poly(A) polymerase*) e para uma RNA helicase putativa.

Tabela 1 – Correspondência da nomenclatura das principais proteínas envolvidas no *splicing* de humano, levedura e *T. brucei*.

	Humano	Levedura	Tripanossomatídeos
<b>Proteínas Sm</b>	SmD1	SmD1	SmD1
	SmD2	SmD2	SmD2
	SmD3	SmD3	SmD3/ Sm16,5K*/ Ssm4**
	SmE	SmE	SmE
	SmF	SmF	SmF
	SmG	SmG	SmG
	SmB	SmB	SmB/ Sm15K***
<b>Proteínas Sm-like (U6-snRNP)</b>	LSm2	LSm2	LSm2
	LSm3	LSm3	LSm3
	LSm4	LSm4	LSm4
	LSm5	LSm5	LSm5
	LSm6	LSm6	LSm6
	LSm7	LSm7	LSm7
	LSm8	LSm8	LSm8
	<b>U1 snRNP</b>	U1-70K	U1-70K
U1-A		U1-A	U1-A
U1-C		U1-C	U1-C
-		-	U1-24K
<b>U2 snRNP</b>	U2-A'	Msl1p	U2-40K
	U2-B''	Lea1p	U2-B''
<b>SF3a</b>	SF3a120	Prp21	SF3a120
	SF3a66	Prp11	SF3a66
	SF3a60	Prp9	SF3a60
<b>SF3b</b>	SF3b155	Rse1p	SF3b155
	SF3b145	Hsh155p	SF3b145
	SF3b130	Cus1p	SF3b130
	SF3b49	Hsh49p	SF3b49
	SF3b14a		SF3b14a
	SF3b14b	Rds3p	SF3b14b
	SF3b10	Ysf3p	SF3b10
<b>U4/U6 disnRNP</b>	15,5K	Snu13	15,5K
	61K	Prp31	61K
	90K	Prp3	90K
	60K	Prp4	60K
	20K	-	-
	U6-110p	Prp24	-
<b>U5 snRNP</b>	U5-220K	Prp8	U5-220K
	U5-200K	Brr2	U5-200K
	U5-116K	Snu114	U5-116K
	U5-102K	Prp6	U5-102K
	U5-100K	Prp28	-
	U5-52K	Lin1	-
	U5-40K	Snu40	U5-40K
	U5-15K	Dib1	U5-15K
	-	-	Cwc21
	Prp43	Prp43	Prp43

Continua

## Continuação

<b>Outros</b>	110K	Snu66	110K
<b>Fatores de <i>Splicing</i></b>	65K	Sad1	65K
	27K	-	27K
	Prp17	Prp17	Prp17
	Prp18	Prp18	-
	Prp16	Prp16	-
<b>Complexo NTC e NTR</b>	Prp22	Prp22	-
	Slu7	Slu7	-
	hPrp19	Prp19	Prp19
	-	Ntr1	-
	-	Ntr2	-

\*Proteínas Sm-like encontradas no anel Sm da partícula U2 snRNP, substituindo a proteína SmD3.

\*\*Proteína Sm-like encontrada no anel Sm da partícula U4 snRNP, substituindo a proteína SmD3.

\*\*\* Proteínas Sm-like encontradas no anel Sm da partícula U2 snRNP, substituindo a proteína SmB.

Fonte: Elaborada pela autora

A partícula U2 snRNP é composta pelo U2 snRNA, sete proteínas Sm comuns e duas proteínas específicas, a U2-A' e a U2-B''(Msl1p e Lea1p em levedura). Em tripanossomatídeos duas das proteínas Sm são distintas, a SmD3 é substituída pela Sm16,5K e a SmB pela Sm15K (figura 17). A partícula U2 se associam outros dois subcomplexos: o SF3a, formado pelas proteínas SF3a60, SF3a66 e SF3a120, parte da extremidade 3' do U2 snRNA, e o SE3b, formado pelas proteínas SF3b130, SF3b155, SF3b145, SF3b49, SF3b14a, SF3b14b e SF3b10 (Rse1p, Hsh155p, Cus1p, Hsh49p, Rds3p e Ysf3p em levedura), integrado à extremidade 5' do U2 snRNA.<sup>51-52</sup>

As subunidades SF3a e SF3b são necessárias para ancorar a partícula U2 snRNP ao pré-mRNA. O SF3a faz contato direto com o U1 snRNP através da interação entre a proteína SF3a120 e o *stem-loop* IV do U1 snRNA. Já o SF3b está relacionado à proteção da extremidade 5' do U2 snRNA, até que este esteja pronto para formar a hélice dupla com o U6 snRNA. Essa proteção provavelmente é feita pelas proteínas SF3b49 e SF3b145.<sup>51-53</sup>

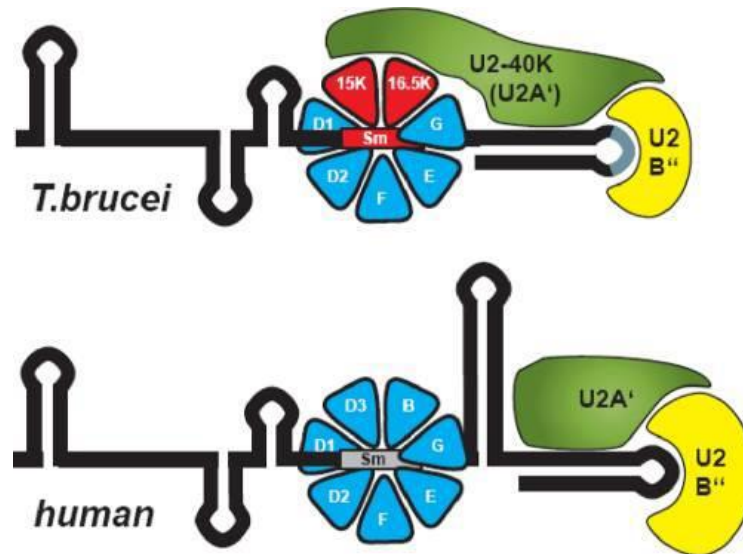


Figura 17 – Comparação entre a partícula U2 de *T. brucei* (acima) e a de humano (abaixo), mostrando as proteínas SM-like 15K e 16,5K, em vermelho, substituindo as proteínas SM D3 e B, a diferença de enovelamento dos U2 snRNA dos dois organismos e a interação da proteína U2 40K de *T. brucei* interagindo com as proteínas SM-like, o que não acontece com a U2A' de humano.

Fonte: PREUßER<sup>47</sup>

Em *S. cerevisiae*, o U4/U6 di-snRNP é composto de 2 snRNAs e 18 proteínas: Snu13, Prp31, Prp3, Prp4 (15,5K, 61K, 90K, 60K e 20K em humano), sete proteínas Sm que se ligam à extremidade 3' do U4 snRNP e sete proteínas LSM que se ligam à extremidade 3' do U6 snRNA. Supõe-se que as proteínas Sm e LSM interagem com os respectivos snRNAs antes da formação do complexo U4/U6 snRNA. A proteína Snu13 se liga ao *stem-loop* 5' (motivo *k-turn*) do U4 snRNA, facilitando a ligação da Prp31 (Figura 18).<sup>54</sup>

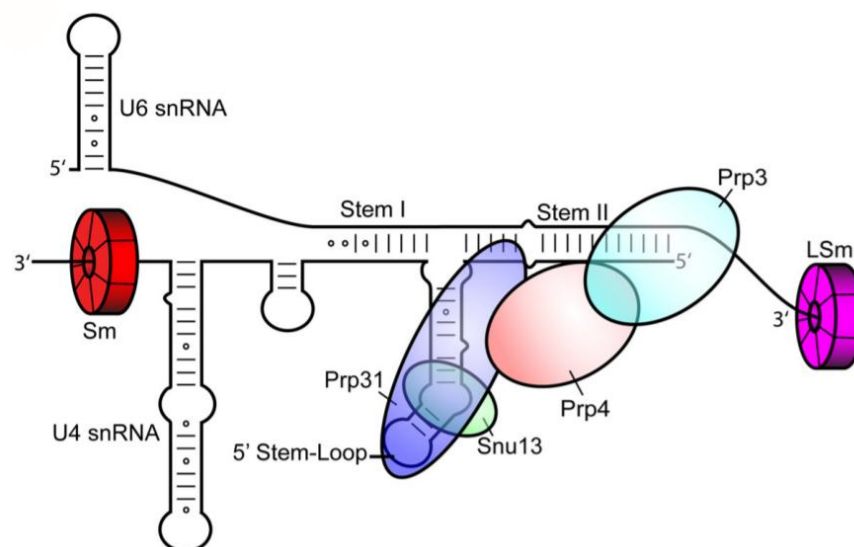


Figura 18 – Representação esquemática do U4/U6 di-snRNP de levedura, mostrando as proteínas Prp31 (azul), Prp 3 (ciano) interagindo com os snRNAs U4 e U6, a proteína Prp4 (vermelha) e Snu13 (verde) posicionadas próximas ao U4 snRNA e o posicionamento do anel das proteínas Sm próximo a porção 3' do U4 snRNA e do anel LSM próximo a porção 3' do U6 snRNA.

Fonte: HARDIN<sup>54</sup>

A partícula U5 snRNP, por sua vez, contém proteínas essenciais para a remodelagem do *spliceosomo* durante os passos do *splicing* e sua composição é altamente conservada de levedura à humano. As proteínas comuns entre os dois organismos são Prp8, Snu114, Brr2, Prp6, Prp28, Lin1, Snu40 e Dbi1(220K, 116K, 200K, 102K, 100K, 52K, 40K e 15K em humanos). Com exceção da 52K (Lin1 em levedura), todas as proteínas estão presentes no tri-snRNP.<sup>55</sup>

A ausência da 52K no tri-snRNP U4/U6.U5 sugere que ela esteja associada a sua formação, mas que não tenha função no *splicing* propriamente dito. A deleção da Lin1 em levedura não se mostrou letal ao organismo, sugerindo que a função da U5-52K e sua hortóloga Lin1 são auxiliares ou redundantes para o processo de *splicing*.<sup>57</sup>

Apenas as proteínas Prp8, Snu114 e Brr2 continuam interagindo com o U5 snRNA durante os passos catalíticos. As proteínas Brr2 e Snu114 também são as responsáveis pela desestabilização, anterior à ativação do *spliceosomo*, da hélice formada pelo U4 e U6 snRNAs (Figura 19).<sup>55-56</sup> Em *T. brucei* todas essas proteínas estão presentes, exceto 52K e 100K. Também foi descoberta uma proteína chamada Cwc21 que parece ser essencial para o primeiro passo do *splicing*.<sup>58</sup>

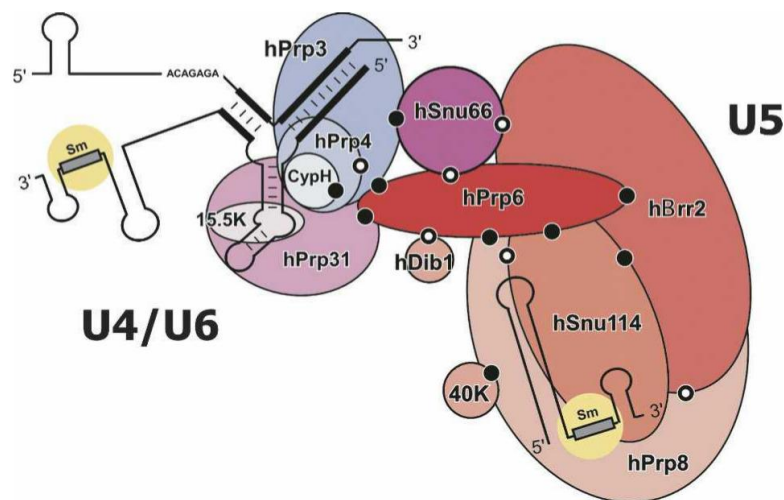


Figura 19 – Representação esquemática do U4/U6.U5 tri-snRNP humano, onde os pontos pretos e brancos representam interações também observadas em levedura.

Fonte: LIU<sup>56</sup>

## 1.8 U5-15K e U5-102K

Em vertebrados, insetos, fungos, nemátodos e plantas, a ligação entre as proteínas U5-15K (Dib1 em levedura) e U5-102K (Prp6) é feita pela proteína U5-52K (Lin1), sendo que dois domínios distintos da 52K estão envolvidos na interação, o N-terminal, que se liga a 102K, e o C-terminal, que se liga a 15K<sup>57</sup>. Em *T. brucei*, as proteínas U5-15K e U5-102K interagem fortemente, sem necessidade de uma proteína homóloga a U5-52K.<sup>59</sup>

A U5-102K é importante para a formação de um U4/U6.U5 tri-snRNP estável através da interação dessa proteína com a proteína específica de U4/U6 snRNP Prp31. A estrutura atômica do tri-snRNP de *S. cerevisiae* mostra que a proteína super helicoidal Prp6 possui 44 hélices- $\alpha$ , sendo que 36 estão organizadas em repetições chamadas TRP (*Tetra-trico Peptide Repeats*), e apresenta o C-terminal conservado, responsável por mediar interações com outros componentes do *spliceosomo*. Mutações no N-terminal mostraram que essa região não é essencial para a associação do *spliceosomo*, mas sim para a sua ativação.<sup>60</sup>

U5-15K é uma proteína altamente conservada entre os eucariotos e sua conexão com o *splicing* foi inicialmente descrita em *S. cerevisiae* como Dib1, onde experimentos mostraram sua presença no U4/U6.U5 tri-snRNP. A falha no *splicing* em levedura, decorrente da deleção da Dib1 (homóloga de U5-15K em *S. cerevisiae*), confirmou sua essencialidade nesse processo,<sup>61</sup> o que também foi demonstrado em *T. brucei*.<sup>59</sup>

A estrutura da U5-15K humana, analisada por Ressonância Magnética Nuclear (RMN)<sup>62</sup> e difração de raio-X<sup>63</sup>, tem enovelamento semelhante ao das tiorredoxinas, caracterizado por quatro fitas  $\beta$  flanqueadas por três hélices- $\alpha$  (figura 20). A partir dessa estrutura, foi construído um modelo da U5-15K de *T. brucei* por homologia (figura 19 c).<sup>64</sup>

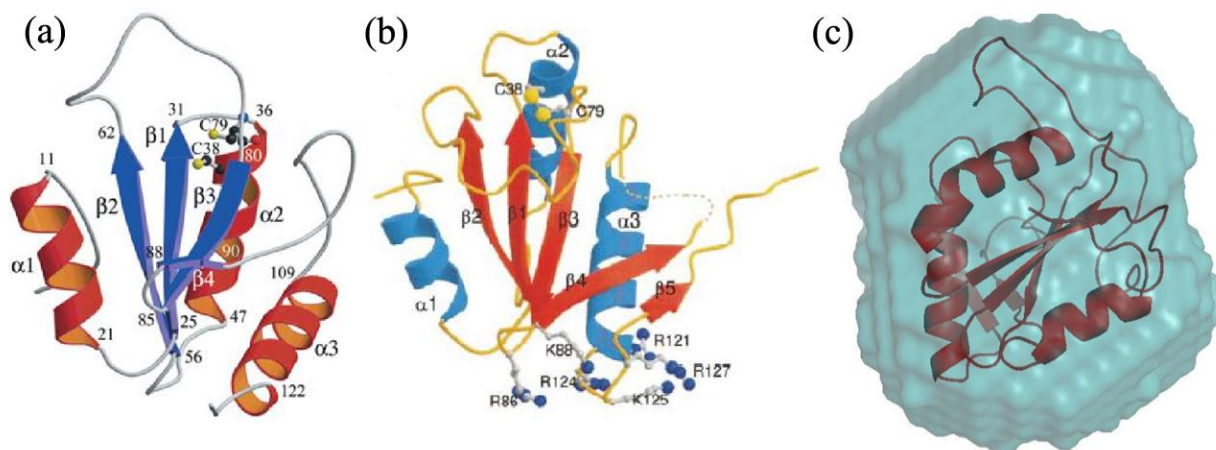


Figura 20 – (a) Estrutura resolvida por RMN da U5-15K humana. (b) Estrutura resolvida por difração de Raio-X da U5-15K humana. (c) Envelope gerado por SAXS alinhado ao modelo gerado por homologia da U5-15 de *T. brucei*.

Fonte: ZHANG<sup>62</sup>; REUTER<sup>63</sup>; LIMA<sup>64</sup>

De acordo com o modelo, a estrutura de *T. brucei* possui uma alça flexível condizente com envelope gerado por espalhamento de raio-X a baixo ângulo (SAXS). Alinhando esse modelo à estrutura cristalográfica da U5-15K de humano, a alça parece estar em regiões importantes para a formação do cristal. Dessa forma, a sua deleção aumentaria as chances de obtenção de cristais.<sup>64</sup>

Também foi visto que a U5-15K de humano possui uma atividade de autoclivagem na extremidade C-terminal, inibida por quimostatina, EDTA e ácido 6-hexanóico<sup>65</sup>. Essa atividade também foi vista na U5-15K de *T. brucei*, sendo impedida por inibidores de serino ou cisteíno proteases.<sup>64</sup>

Experimentos usando a sequência de 20 aminoácidos no C-terminal da U5-15K como ligação entre Vhb (*Vitreoscilla haemoglobin*) e GST (*Glutathione S-Transferase*) mostram que essas duas proteínas são separadas na presença da U5-15K truncada (hDim11-128), o que sugere que a U5-15K é capaz de clivar o mesmo sítio em diferentes contextos e que a região central semelhante a tioredoxina retém a atividade de peptidase.<sup>65</sup>

A estrutura do U4/U6.U5 tri-snRNP de levedura, resolvidas recentemente por criomicroscopia eletrônica, mostram que a Dib1 tem uma localização central no complexo, associando-se ao U5 snRNA, ao pré-mRNA, à proteína Prp31 e ao domínio N-terminal da Prp8. Também foi visto que tanto Dib1 quanto Prp8 estão associadas com a estabilização do *hairpin* formado pela parte invariante do U6 snRNA (ACAGAGA) e a ativação do complexo catalítico pode ocorrer quando mudanças na conformação da Prp8 expulsam a Dib1 do complexo, liberando a sequência ACAGAGA para se ligar ao 5'ss (Figura 21).<sup>67-68</sup>



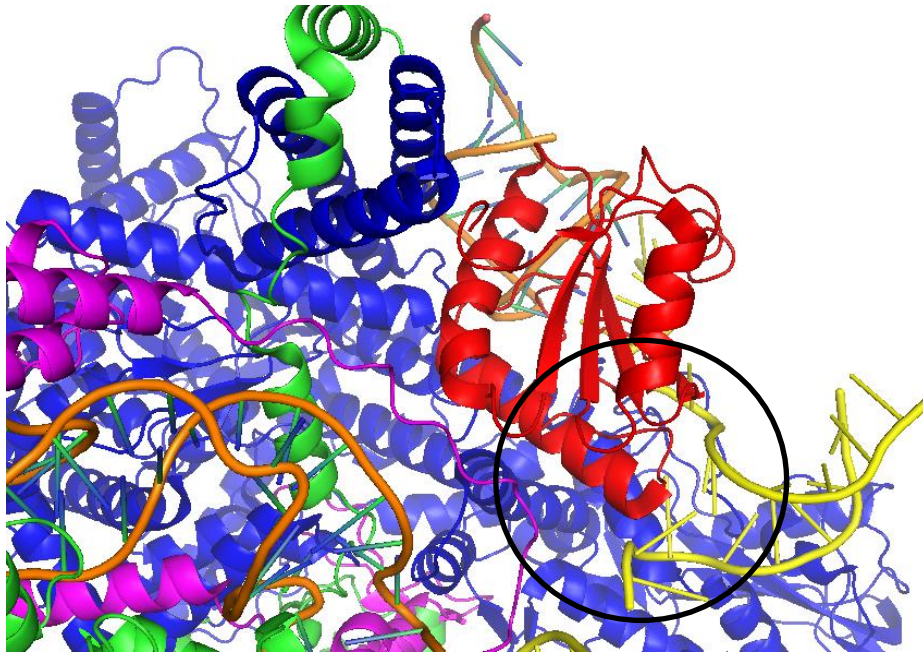


Figura 21 – Localização da Dibi (vermelho) no U4/U6.U5 tri-snRNP de levedura, mostrando as regiões de interação com o U6 snRNA (amarelo), Prp8 (azul) e Prp31 (magenta). A área de interação entre a Dibi e o U6 snRNA está circulado em preto.

Fonte: NGUYEN<sup>68</sup>

A posição da cauda C-terminal, que parece interagir com o U6 snRNA no U4/U6.U5 tri snRNP, e a necessidade da retirada da Dibi1 para ativação do complexo B, aliadas ao fato de a forma truncada da U5-15K, hDim11-128, agir como dominante negativo da proteína selvagem em levedura e a conservação do C-terminal, podem sugerir que a atividade de autoclivagem possa ser um ponto de regulação da ativação do complexo catalítico.<sup>62,69</sup>

## 1.9 Prp43

A Prp43 é uma RNA helicase da família DEAH-box, cuja estrutura cristalográfica revelou 6 domínios: domínio N-terminal (NTD), RecA1, RecA2, WH, HB e um domínio C-terminal (CTD). Os domínios RecA fazem o posicionamento do motivo de ligação de RNA e os domínios WH e HB formam um canal que encerra o RNA (Figura 22).<sup>70</sup>

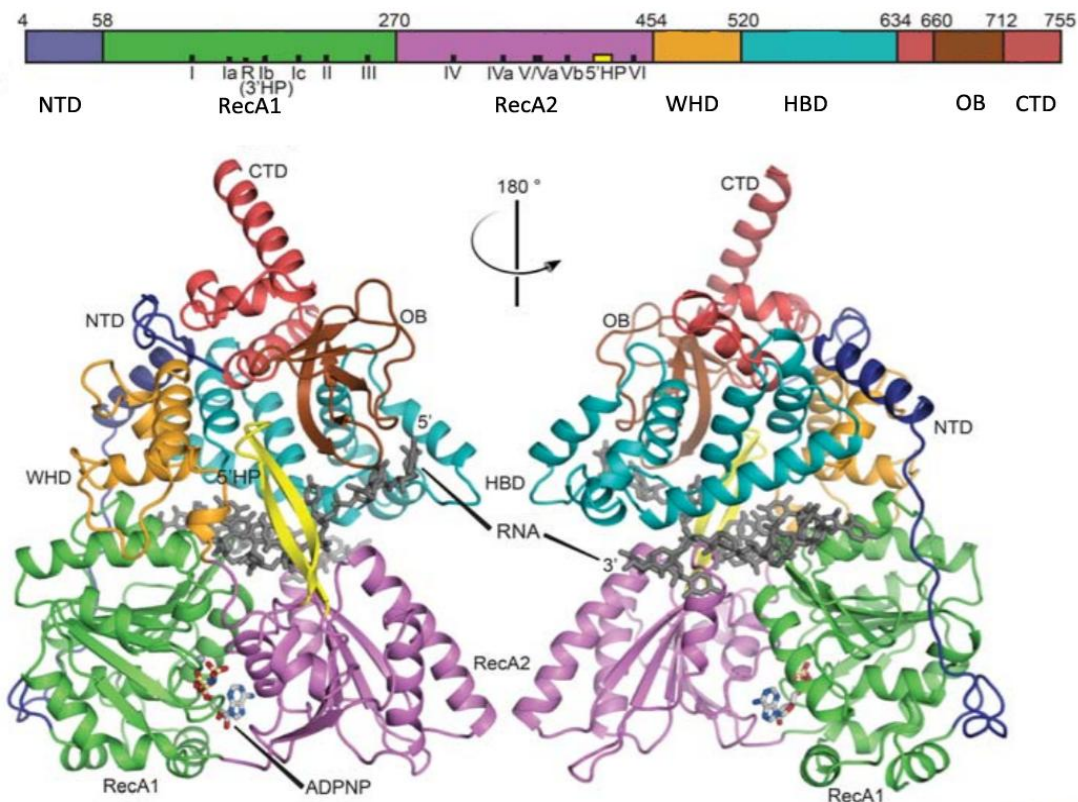


Figura 21 – Estrutura cristalográfica da Prp43 de levedura, com o domínio NTD em azul escuro, RecA1 em verde, RecA2 em lilás, WHD em laranja, HBD em azul claro OB em marrom e CTD em vermelho. Fonte: HE <sup>70</sup>

A Prp43 tem como função a liberação do íntron excisado e desmontagem do spliceossomo após a finalização da reação. Além disso, experimentos sugerem que ela esteja relacionada ao descarte de pré-mRNAs rejeitados pelo spliceossomo. Em humanos, a Prp43 foi identificada como sendo parte do U2 snRNP, o que levanta a possibilidade de sua atuação ter como alvo a interação U2-pré-mRNA.<sup>71</sup>

Foi observado que, para a ativação completa da Prp43, é necessária a interação do seu domínio C-terminal com o motivo G-patch da proteína Ntr1. A associação dessas duas proteínas estimula tanto a atividade helicase quanto ATPase da Prp43.<sup>72</sup>

Em tripanossomatídeos, foi encontrada uma proteína homóloga à Prp43 e mutações em suas regiões conservadas resultaram no acúmulo de Y-íntron, indicando sua importância também o *trans-splicing*.<sup>97</sup>

Tendo em vista o que foi exposto sobre o que se conhece das proteínas envolvidas no SL *trans-splicing* de *T. brucei*, esse trabalho tem foco em proteínas que participam do complexo B, como é o caso da U5-15K e U5-102K e uma proteína que atua no complexo ISL, como é o caso da Prp43, buscando mais informações sobre a função e os parceiros de interação dessas proteínas em *T. brucei*.

## 2 OBJETIVOS

Esse trabalho tem por objetivo estudar as proteínas U5-15K, U5-102K e Prp43 de *Trypanosoma brucei*, com foco no seu envolvimento no *trans-splicing*, e estudo das proteínas U5-15K e U5-102K recombinantes:

Os objetivos específicos são descritos a seguir:

1. Investigar as condições de cristalização da proteína U5-15K.
2. Caracterizar a atividade de autoclivagem da proteína U5-15K, bem como seu sítio catalítico.
3. Estudar a estrutura do subcomplexo formado pelas proteínas U5-15K e U5-102K.
4. Estudar as interações entre subcomplexo formado pelas proteínas U5-15K e U5-102K e as proteínas pertencentes ao U5.U4/U6 tri snRNP.
5. Estudar as interações da proteína Prp43.



### 3 MATERIAIS E MÉTODOS

#### 3.1 Análise bioinformática das sequências de aminoácidos das proteínas U5-102K e Prp43

A análise dos parâmetros físico-químicos teóricos da proteína U5-102K de *Trypanosoma brucei* foi feita a partir da sua sequência de aminoácidos (XP\_829626.1) usando a ferramenta PROTPARAM<sup>73</sup> do servidor ExPASy.

A predição da estrutura secundária da porção N-terminal da U5-102K foi feita através do servidor PSIPRED<sup>74-75</sup> e a análise de aminoácidos conservados das proteínas U5-102K, Prp43 (XP\_844789.1) e NtrI (XP\_846187.1) foram feitas através do alinhamento múltiplo com sequências homólogas, gerado pelo programa ClustalW.<sup>76</sup>

#### 3.2 Manipulação do DNA

##### 3.2.1 Extração do DNA genômico

Para a extração de DNA, 1ml da cultura de *T. brucei*, forma procíclica, na fase logarítmica ( $5 \times 10^6$  células) foi centrifugado a 3000 x g, a 4 °C por 7 minutos e foi usado o kit de extração DNeasy® Plant Mini (50) (QIAGEN).

##### 3.2.2 Clonagem

Os oligonucleotídeos iniciadores usados neste trabalho foram desenhados no programa *SnapGene Viewer* a partir da sequência de DNA das proteínas U5-15K (XM\_841996.1), U5-102K (XM\_824533.1) e Prp43 (XM\_839696.1). Esses oligonucleotídeos estão representados na tabela abaixo.

Tabela 2 – Nome e sequência dos oligonucleotídeos iniciadores usados nesse trabalho, vetores onde as proteínas foram clonadas e enzimas de restrição utilizadas.

Nome	Enzima de Restrição	Vetor	Sequência
15KBglII sense	<i>Bgl</i> II	pETDuet-1	5'-AGATCTCATGACGGAGTTGACACGCC-3'
15K anti-sense	<i>Kpn</i> I	pETDuet-1	5'-GGTACCCTAACGGCGTACAGCAGCTG-3'
102KBamHI sense	<i>Bam</i> HI	pETDuet-1	5'-GGATCCGATGAGCGCTGCGGCACCTGTC-3'
102K anti-sense	<i>Not</i> I	pETDuet-1	5'-GCGGCCGCTTACAATTTGATCATTTC-3'
102K AS EagI	<i>Eag</i> I	pC-PTP-NEO	5'-CGGCCGCAATTTGATCATTTCACGA-3'
102K S ApaI	<i>Apa</i> I	pC-PTP-NEO	5'-GGGCCCTGCTTCAGCAACTTCCTAAG-3'

continua

### continuação

<b>ProtC-Tap3</b>	-	-	5'TCGCTAGCAGTAGTTGGAATATCATATTTCCCATCAA TAAGACGAGG3'	AAGACGAGG3'
<b>N-Term NdeI sense</b>	<i>NdeI</i>	pETSUMO	5'- CATATGCCTGTCCTTTCCACTAAG-3'	
<b>N-Term BamHI anti-sense</b>	<i>BamHI</i>	pETSUMO	5'- GGATCTAACGAAAGCGTATCCGTCC-3'	
<b>Prp43 AS EagI</b>	<i>EagI</i>	pC-PTP-NEO	5'- CGGCCGAGACGAATCATCATCATC-3'	
<b>Prp43 S ApaI</b>	<i>ApaI</i>	pC-PTP-NEO	5'- GGGCCCGATCAGCGCGGATGCGCCATG-3'	
<b>Prp43 out</b>	-	-	5'- ATGGCCGCGATGCTTTTG-3'	

Fonte: Elaborada pela autora.

A amplificação das sequências de interesse foi feita por meio da Reação em Cadeia da Polimerase (PCR), empregando polimerase termorresistente recombinante de alta precisão *Phusion (Thermo scientific)*, nas condições descritas pelo fabricante<sup>90</sup>. O resultado foi analisado em gel de agarose 1% com 0,5µgml<sup>-1</sup> de *SyberSafe (Thermo scientific)* em tampão TAE 1x (8mM Tris-acetato pH 8,0; 0,2mM EDTA).

Os fragmentos amplificados foram purificados usando o *PCR Purification kit (Cellco)* seguindo as instruções do fabricante e clonados no vetor de clonagem pJET1.2 (*Thermo scientific*). O produto foi transformado em *Escherichia coli* DH10b competentes como será descrito na seção 3.4.2.

As colônias positivas foram inoculadas em 5mL de meio LB e incubadas a 37°C por aproximadamente 16 horas. O DNA plasmidial foi extraído com o kit *Fast-n-Easy Plasmid Mini-Prep (Cellco)* seguindo as recomendações do fabricante. Os insertos foram digeridos com as respectivas enzimas de restrição (tabela 2) e clonados nos vetores de expressão compatíveis (tabela 2) com T4 DNA ligase (*Thermo scientific*) seguindo as recomendações do fabricante.<sup>91</sup> O produto das clonagens foi transformado em *E. coli* DH10b competente seção 3.4.2.

A construção contendo a sequência de DNA codificante para a proteína U5-102K de *T. brucei* no vetor de expressão pET28a foi gentilmente cedida pelo Dr. Marco Tulio A. da Silva.

Após a confirmação da clonagem, as construções pET28a\_102K, pETDUET\_15K\_102K e pET SUMO\_102k, foram transformadas em cepas de expressão seção 3.4.2.

### 3.3 Mutações sítio dirigida

Nesse trabalho foram realizadas sete mutações sítio dirigidas no gene que codifica a proteína U5-15K, sendo seis substituições de aminoácidos específicos por alaninas e uma

deleção de 11 aminoácidos da alça flexível compreendida entre as treoninas 36 e 48, e uma mutação na sequência da Prp43 para permitir sua inserção no vetor pC-PTP-NEO. Os aminoácidos mutados e os pares de oligonucleotídeos usados estão representados na tabela 3. Cada mutação foi realizada de forma independente.

Tabela 3 – Nome e sequência dos oligonucleotídeos iniciadores usados para as mutações sítio dirigidas, região mutada das proteínas e nomes dados a elas nesse trabalho.

Número	Nome do oligonucleotídeo	Sítio da mutação	Sequência	Nome dado proteína mutante
1	15k C66A sense	Cisteína 66	5'-CAAAGGTACGGAAGTATGCCACAATTTATGTTGTG-3'	U5-15K C66A
	15k C66A anti-sense		5'-CACAACATAAATTGTGGCATACTTCCGTACCTTTG-3'	
2	15k alça sense	Alça (ácido glutâmico 37 até serina 47)	5'-GTGAAGCGACGACACGCCAAATAGATGAAGTACTTGT CGCACTCGCACCAAAG-3'	U5-15K sem alça
	15k alça anti-sense		5'-CTATTTGGCGTGTGCTGCTTACCCTAATGACTGAAG CGGATAAGTACAAG-3'	
3	15k pc sense	Ponto de Clivagem (lisina 139 até tirosina 143)	5'-GGGAGAAGTGGAGCTGCTGCTGCTTCCGAAAAGA AATTCACAACAGCTGCTGTACGC-3'	U5-15K pc
	15k pc anti-sense		5'-GAATTTCTTTTCCGAGCTGCTGCTGCTTCCACTTCT CCCCACGGTACGCTGCGTC-3'	
4	15k S120A sense	Serina 120	5'-CAATTTTGTGTAGCGGAAGATGAACTGTTATC-3'	U5-15K S120A
	15k S120A anti-sense		5'-GATAACAGTTCATCTTCCGCTACAACAAAATTG-3'	
5	15k S126A sense	Serina 126	5'-GAAGATGAACTGTTAGCCATCGCTGACGCAGCGTAC-3'	U5-15K S126A
	15k S126A anti-sense		5'-GTACGCTGCGTCAGCGATGGCTAACAGTTCATCTTC-3'	
6	15k S137A sense	Serina 137	5'-GTACCGTGCAGGGAGAGCTGGAAAGACGATAG-3'	U5-15K S137A
	15k S137A anti-sense		5'-CTATCGTCTTCCAGCTCTCCCCGACGGTAC-3'	
7	15k S144A sense	Serina 144	5'-GAAAGACGATAGCATACGCGAAAAGAAATTC-3'	U5-15K S144A
	15k S144A anti-sense		5'-GAATTTCTTTTCCGCGTATGCTATCGTCTTTC-3'	
8	Prp43 A702A sense	Alanina 702	5'-CCGCAATAACCGGCAGAGAGGGACCAAAAATAAAG-3'	---
	Prp43A702A antisense		5'-CTTATTTTGGTCCCTCTGCCGCGTATTGCGG-3'	

Fonte: Elaborada pela autora.

Os pares de oligonucleotídeos, representados pelos números 1, 4, 5, 6, 7 e 8 na tabela C, foram desenhados de forma que suas sequências se sobrepusessem perfeitamente (Figura 23 b), já os pares 2 e 3 possuem uma região de aproximadamente 12 nucleotídeos onde há sobreposição apenas com a sequência da proteína alvo<sup>77</sup> mas não entre os oligonucleotídeos (figura 23 a).

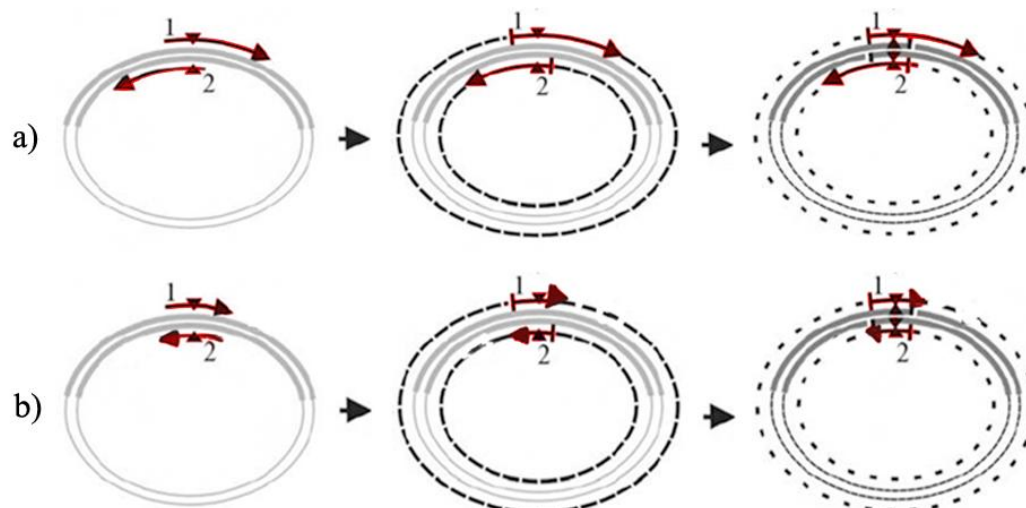


Figura 23 - Representação gráfica do pareamento de bases e posição dos oligonucleotídeos (setas vermelhas) em relação ao DNA circular de interesse.

Fonte: LIU<sup>77</sup>

Para as mutações 1, 4, 5, 6, 7 e 8, a PCR foi feita utilizando 45ng de DNA plasmidial (pETSUMO\_U515K), extraído com o kit *Fast-n-Easy Plasmid Mini-Prep (Cellco)*, 0,3  $\mu\text{L}$  oligonucleotídeo sense (0,3  $\mu\text{g}\cdot\mu\text{L}^{-1}$ ), oligonucleotídeo anti-sense (0,3  $\mu\text{g}\cdot\mu\text{L}^{-1}$ ), 0,3  $\mu\text{L}$  DNA polimerase *Phusion* e água tipo1 para completar 25  $\mu\text{L}$  de reação. Também foi preparada uma reação controle contendo os mesmos reagentes exceto a DNA polimerase. A ciclagem foi usada como descrito na tabela 4. Os tempos de extensão das mutações na U5-15K e Prp43 foram de 8 e 4 minutos respectivamente.

Tabela 4 - Ciclagem da PCR usada para as mutações descritas.

Passo	Temperatura (°C)	Tempo
<b>Desnaturação inicial</b>	98	30 segundos
<b>Desnaturação</b>	98	10 segundos
<b>Anelamento</b> 30 vezes	62	30 segundos
<b>Extensão</b>	72	1 minuto por kbase*
<b>Extensão final</b>	72	10 minutos

\*8 minutos para U5-15K em pET SUMO e 4 minutos para Prp43 em pJET1.2.

Fonte: Elaborada pela autora.

Após a PCR, 12,5 $\mu\text{L}$  da reação foram digeridos com a enzima de restrição *DpnI*, a 37°C por aproximadamente 16 horas e transformados em *E. coli* DH10b. Foram escolhidas cinco colônias para cada transformação, inoculadas em 5mL de meio LB incubado a 37°C por aproximadamente 16 horas e o DNA plasmidial foi extraído com o kit *Fast-n-Easy Plasmid Mini-Prep (Cellco)*, para posterior análise por sequenciamento.

Para as mutações 2 e 3, foram usados 500ng de DNA plasmidial, 0,3 $\mu\text{L}$  de DNA polimerase *Phusion*, 4  $\mu\text{L}$  de oligonucleotídeo (0,3  $\mu\text{g}\cdot\mu\text{L}^{-1}$ ). Além disso, foram feitas duas



reações (ciclagem da tabela 4), uma para cada oligonucleotídeo do par, bem como dois controles (reação sem DNA polimerase). Em seguida, 20µl do produto de cada PCR foram hibridizados como descrito na tabela 5 e digeridas com a enzima de restrição *DpnI*, a 37°C por aproximadamente 16 horas.

Tabela 5 – Variação da temperatura para hibridização das fitas de DNA resultantes das mutações 2 e 3 da tabela 3.

Temperatura (°C)	Tempo
95	5 minutos
90	1 minuto
80	1 minuto
70	30 segundos
60	30 segundos
50	30 segundos
40	30 segundos
37	∞

Fonte: Elaborada pelo autor.

O produto foi transformado em *E. coli* DH10b e foram escolhidas cinco colônias, como citado anteriormente. O DNA plasmidial foi extraído com o kit *Fast-n-Easy Plasmid Mini-Prep (Cellco)* e procedeu-se com o sequenciamento.

### 3.4 Manipulação das cepas de *E. coli*

#### 3.4.1 Preparo das células competentes

Foram usadas as cepas DH10b, BL21(DE3), BL21(DE3) *codon plus* e Rosetta Gami B(DE3) de *E. coli*, sendo que as duas primeiras foram preparadas seguindo o protocolo para células ultracompetentes e duas últimas o de eletrocompetentes.

Para preparar as células ultracompetentes, primeiramente foram preparados dois Erlenmeyers (125mL e 500mL) e tubos de ensaio com 5 e 10ml de meio TYM (20gL<sup>-1</sup> triptona; 5gL<sup>-1</sup> extrato de levedura; 5,8gL<sup>-1</sup> NaCl e 2,5 gL<sup>-1</sup> MgSO<sub>4</sub>; pH 7,5).

Os tubos de ensaio contendo 5ml de meio TYM foram inoculados com *E. coli* e incubados a 37°, sob agitação por aproximadamente 16h.

No dia seguinte, a cultura foi diluída (1:100) em 10 ml de meio TYM e incubada a 37°, 150rpm, até atingir uma densidade óptica à 600nm (OD600) entre 0,2 e 0,6, para então ser transferida para 40 ml de meio TYM (preparado no erlenmeyer de 125ml) e incubada a 37°, 150rpm, até atingir OD600 entre 0,5 e 0,9. A cultura foi então transferida para 200ml de meio

TYM (preparado no erlenmeyer de 500ml) e incubada a 37°C, 150rpm, até atingir uma OD<sub>600</sub> de 0,6.

As células foram resfriadas rapidamente em gelo e centrifugada a 4000 rpm por 15 minutos a 4°C e ressuspensas gentilmente em 100ml de tampão TBF1 (0,03M acetato de potássio, 0,05M MnCl<sub>2</sub>, 0,1M KCl, 0,01M CaCl<sub>2</sub> e 15% m/v de glicerol) gelado e esterilizado por filtração.

As células foram novamente centrifugadas, 4000 rpm, 4°C, por 8 minutos e ressuspensas gentilmente com 10ml de tampão TBF2 (0,01M MOPS pH7, 0,075M CaCl<sub>2</sub>, 0,01M KCl e 15% m/v de glicerol) gelado e esterilizado por filtração.

Dessas células, foram separadas alíquotas de 100µl em microtubos de 1,5ml, congeladas em nitrogênio líquido e estocadas a -80°C.

Para o preparo das células eletrocompetentes, uma única colônia de *E. coli* foi inoculada em 25ml de meio 2xYT (16gL<sup>-1</sup> de triptona, 10gL<sup>-1</sup> de extrato de levedura e 5gL<sup>-1</sup> de NaCl, pH7) e incubado por aproximadamente 16 horas a 37°C, sob agitação.

Desse inóculo, 2,5ml foram transferidos para 250 ml de meio 2xYT e incubados a 37°C, 150rpm, até atingir a OD<sub>600</sub> entre 0,2 e 0,25.

As células foram então resfriadas em gelo, transferidas para oito garrafas de 40ml (32 ml em cada garrafa), centrifugadas a 10000 rpm, 4°C, por 10 minutos, ressuspensas em 8ml de glicerol 10% estéril e previamente gelado e transferidas para dois tubos (32ml cada).

Esses tubos foram novamente centrifugados (10000 rpm, 4°C) por 15 minutos, as células foram ressuspensas em 35ml de glicerol 10% estéril e gelado e centrifugados por 15 minutos (10000 rpm, 4°C).

O sobrenadante foi descartado, as células ressuspensas com 15ml de glicerol 10% estéril e gelado, transferidas para um único tubo e centrifugadas pela última vez por 20 minutos (10000 rpm, 4°C).

O sobrenadante foi descartado e as células ressuspensas em 1ml de glicerol 10%. Desses 1ml, 25µl foram diluídos em 10ml de glicerol 10% e a OD<sub>600</sub> foi medida. Continuou-se a adicionar 25µl da suspensão de células até uma OD<sub>600</sub> de 0,15.

Essas células foram então alíquotadas em microtubos de 1,5 ml estéreis e congeladas em nitrogênio líquido. O armazenamento foi feito a -80°C.

### 3.4.2 Transformação das bactérias competentes

Para a transformação das cepas quimiocompetentes (DH10b e BL21(DE3)) foram usados 20µl da solução de ligação (seção 2.3.2), uma alíquota de 100µl de *E. coli* competente, previamente descongelada, e 80µl de tampão de transformação (100 mM de KCl; 30 mM de CaCl<sub>2</sub>; 50 mM de MgCl<sub>2</sub> e 1% de PEG 6000 ou PEG 4000). A adição da reação de ligação e do tampão à *E. coli* foi feita no fluxo laminar para evitar contaminações. Manteve-se essa reação em gelo por 30 minutos, em seguida foi dado o choque térmico a 42°C por dois minutos. Foram então adicionados 500µl de meio LB e foi feita a incubação por aproximadamente 1 hora a 37°C, sob agitação.

Por fim, as células foram decantadas por centrifugação (aproximadamente 10s a 10000g), ressuspensas com 100µl do próprio meio contido no microtubo e espalhadas em uma placa de Petri contendo 20ml de LB-ágar (10 gL<sup>-1</sup> triptona; 5 gL<sup>-1</sup> extrato de levedura; 10 gL<sup>-1</sup> NaCl; 1,5% ágar) contendo o antibiótico de seleção pertinente (tabela 6). As placas foram incubadas a 37°C por aproximadamente 16 horas.

Tabela 6 – Relação dos vetores usados nesse trabalho e em que cepas foram transformados e os antibióticos de seleção.

Vetor	Cepa de <i>E.coli</i>	Antibiótico de seleção (concentração)
pJET1.2		Ampicilina (100 µgml <sup>-1</sup> )
pET SUMO		Canamicina (50 µgml <sup>-1</sup> )
pET 28a	DH10b	Canamicina (50 µgml <sup>-1</sup> )
pETDuet-1		Ampicilina (100 µgml <sup>-1</sup> )
pc-PTP-neo		Ampicilina (100 µgml <sup>-1</sup> )
pET 28 <sup>a</sup>		Canamicina (50 µgml <sup>-1</sup> )
pETDuet-1	BL21(DE3)	Ampicilina (100 µgml <sup>-1</sup> )
pETSUMO		Canamicina (50 µgml <sup>-1</sup> )
pETDuet-1	BL21(DE3) <i>codon plus</i>	Ampicilina (100 µgml <sup>-1</sup> )
		Cloranfenicol (34 µgml <sup>-1</sup> )
pETDuet-1	Rosetta Gami B(DE3)	Ampicilina (100 µgml <sup>-1</sup> )
		Canamicina (50 µgml <sup>-1</sup> )
		Cloranfenicol (34 µgml <sup>-1</sup> )
		Tetraciclina (12,5 µgml <sup>-1</sup> )

Fonte: Elaborada pela autora.

Após a incubação as colônias foram retiradas com auxílio de uma ponteira de 10µl e diluídas em 10µl de meio LB. Dessa solução, 3 µl foram retirados para a realização de uma PCR para a confirmação da clonagem. Essa PCR foi realizada com a enzima Taq DNA Polymerase recombinante (*Thermo Scientific*) seguindo as recomendações do fabricante<sup>92</sup> e a visualização foi feita em gel de agarose 1% com 0,5µgml<sup>-1</sup> de *SyberSafe* (*Thermo scientific*).

Para a transformação das bactérias eletrocompetentes (BL21(DE3) *codon plus* e Rosetta Gami B(DE3)) foram usados 100µl das células e 20µl de DNA plasmidial.

A alíquota foi descongelada em gelo e a ela adicionou-se de 1 a 2µg de DNA. As células foram então transferidas para uma cubeta de eletroporação (Gene Pulser Cuvette Bio-Rad). O DNA usado deve ter o mínimo de sal possível, para não estourar as células durante o pulso elétrico. A cubeta foi levada a um eletroporador (Gene Pulser® II, Bio-Rad) programado para 200Ω, 25µF e 1,7kV. Após o pulso, a constante de tempo deve estar entre 4 e 4,7ms.

Imediatamente após o pulso, foi adicionado 1ml de meio LB na cubeta, para retirada das células e transferência para um microtubo de 1,5ml estéril e em seguida o microtubo foi levado ao agitador para incubação a 37°C por aproximadamente uma hora.

Por fim as células foram plaqueadas em LB sólido com os antibióticos de seleção (tabela 6) e incubadas a 37°C por aproximadamente 16h.

### **3.5 Manipulação da cepa 427 da forma procíclica de *T. brucei***

#### **3.5.1 Cultura**

Nesse trabalho foi usada a cepa 427 de *T. brucei* em sua forma procíclica (TbPCF 427). Seu cultivo foi feito em meio SDM-79<sup>78</sup>, a 26°C. A manutenção das culturas foi feita em garrafas contendo 10ml do meio com repiques de 1:10 uma vez por dia e 1:50 a cada 3 dias, dependendo dos experimentos conduzidos.

#### **3.5.2 Transfecção**

Para a transfecção foram usados 10 ml ( $1 \times 10^6$ - $10^7$  parasitos por ml) da cultura de TbPCF 427. Essas células foram centrifugadas por 7 minutos a 3000xg, temperatura ambiente, o sobrenadante foi descartado e as células ressuspensas em meio de eletroporação (*Electroporation Media*, EM em inglês) contendo 0,15 mM CaCl<sub>2</sub>; 120 mM KCl; 10 mM K<sub>2</sub>HPO<sub>4</sub>; 25 mM HEPES; 2 mM EDTA; 5 mM MgCl<sub>2</sub>; pH 7,6.

As células foram novamente centrifugadas por 7 minutos a 3000xg, temperatura ambiente e ressuspensas com 1,5 mL de EM. À cubeta de eletroporação (Gene Pulser Cuvette Bio-Rad), foram adicionados 450µL das células ressuspensas e 100µL de pc-PTP-neo\_102k e pc-PTP-neo\_Prp43 digeridos com BbsI, ambos a uma concentração de 10µg.

As cubetas foram submetidas a pulsos de 1,5 kV, 25 $\mu$ F em um eletroporador Gene Pulser® II (Bio-Rad) e imediatamente transferidas para uma garrafa com 10 mL de meio SDM-79 contendo 20  $\mu$ gml<sup>-1</sup> do antibiótico de seleção G418.

Desses 10ml, 2,5ml foram adicionados a um tubo de plástico contendo 2,5ml de uma cultura de TbPDF427 não transfectado e 7,5ml de meio SDM-79. Em seguida, 120 $\mu$ l foram distribuídos em 96 poços de uma placa de poliestireno estéril. A placa foi mantida em câmara úmida a 26°C para a seleção dos clones, que foram posteriormente transferidos para garrafas de 10 mL com 40  $\mu$ gml<sup>-1</sup> G418.

Para a confirmação da transfecção, 1ml das culturas foi centrifugado a 1000xg por 5 minutos, as células foram ressuspensas com tampão de amostra e aplicadas em SDS-PAGE 10% para que o extrato proteico fosse analisado por *Western Blot* (seção 3.7). Além disso, extraiu-se o DNA genômico de 10 mL de cultura e foi feita uma PCR com os oligonucleotídeos iniciadores 102K 5'out e ProtC-Tap3 (tabela 1).

### 3.6 Obtenção das proteínas recombinantes

#### 3.6.1 Expressão das proteínas em *E. coli*

Para a expressão das proteínas recombinantes foram usadas as cepas BL21(DE3), BL21(DE3) *codon plus* e Rosetta Gami B(DE3). As construções que cada cepa recebeu e os antibióticos de seleção estão representados na tabela 7. Para todas as colônias, células foram inicialmente inoculadas em 5ml de meio LB contendo o antibiótico de seleção correspondente e incubados a 37°C por aproximadamente 16 horas, sob agitação.

Tabela 7 – Vetores transformados nas cepas de expressão e respectivos antibióticos de seleção.

Cepa ( <i>E. coli</i> )	Construção	Antibióticos (concentração)
BL21(DE3)	pET-SUMO_15K C66A	Canamicina (50 $\mu$ gml <sup>-1</sup> )
	pET-SUMO_15k S120A	Canamicina (50 $\mu$ gml <sup>-1</sup> )
	pET-SUMO_15k S126A	Canamicina (50 $\mu$ gml <sup>-1</sup> )
	pET-SUMO_15k S137A	Canamicina (50 $\mu$ gml <sup>-1</sup> )
	pET-SUMO_15k S144A	Canamicina (50 $\mu$ gml <sup>-1</sup> )
	pET-SUMO_15k truncada	Canamicina (50 $\mu$ gml <sup>-1</sup> )
	pET-SUMO_15k PC	Canamicina (50 $\mu$ gml <sup>-1</sup> )
	pET-SUMO_102K N-term	Canamicina (50 $\mu$ gml <sup>-1</sup> )
	pET 28a_102k	Canamicina (50 $\mu$ gml <sup>-1</sup> )
	pET duet_102K_15K	Ampicilina (100 $\mu$ gml <sup>-1</sup> )
	BL21(DE3) <i>codon plus</i>	pET duet_102K_15K
		Cloranfenicol (34 $\mu$ gml <sup>-1</sup> )
Rosetta B(DE3) <i>Gami</i>		Ampicilina (100 $\mu$ gml <sup>-1</sup> )
		Cloranfenicol (34 $\mu$ gml <sup>-1</sup> )
	pET duet_102K_15K	Canamicina (50 $\mu$ gml <sup>-1</sup> )

Fonte: Elaborado pela autora.

Os testes de expressão iniciais foram feitos com aproximadamente 3 colônias resultantes das transformações descritas na seção 3.4.2. Nesses testes, 100µl do inóculo inicial foi transferido para um tubo de ensaio contendo 10ml de meio LB e o antibiótico correspondente. As culturas foram incubadas à 37°C, sob agitação de 150rpm, por aproximadamente 1 hora. Em seguida a temperatura foi reduzida para 20°C e a expressão da proteína de interesse foi induzida com 0,5mM de IPTG quando as culturas atingiram uma DO<sub>600</sub> de 0,6. Foram recolhidas alíquotas imediatamente antes da indução e depois de 2, 4, 6 e 16 horas de indução. Essas alíquotas foram centrifugadas por 5 minutos à 10000xg em centrífugas de bancada e as células foram ressuspensas em tampão de amostra (250mM Tris-HCl, pH6,8; 10% SDS, 30% v/v Glicerol, 10mM DTT, 0,05% w/v Bromofenol *Blue*) e aplicadas em gel SDS-PAGE 15%.

Para a expressão em maior escala, inicialmente foram feitos dois pré-inóculos de 5ml para cada litro de meio LB (autoclavado em erlenmeyer de 2,5L). Esses 10 mL da cultura de bactéria foram adicionados ao meio com o antibiótico de seleção correspondente e incubado a 37°C, 150rpm, por aproximadamente uma hora.

Após esse período a temperatura foi reduzida para 20°C e a expressão das proteínas foi induzida pela adição de 0,5 mM de IPTG quando as culturas atingiram a DO<sub>600</sub> de 0,6. As culturas foram mantidas a 20°C, 150 rpm por aproximadamente 16 horas.

### **3.6.2 Lise celular**

As células foram centrifugadas por 45 minutos a 4000 rpm, 4°C e ressuspensas em 50ml de tampão. Para as sete mutantes da U5-15K (nomes às quais serão referidas estão na tabela 2) o tampão utilizado é composto por 20mM Tris-HCl pH 7,5 ; 100mM NaCl. Para a proteína U5-102k expressa com pET 28a, foi usado um tampão composto por 20mM Tris-HCl pH 7,5 ; 150mM NaCl e um comprimido de inibidor de protease (cOmplete EDTA-free, Roche). Para a porção N-terminal da U5-102K foi usado um tampão composto por 50mM Tris-HCl pH 7,5 ; 150mM NaCl.

A lise foi feita por ultrassonicação com 10 ciclos de 30s de pulso ligado intercalados por 45s desligado. As células foram mantidas em gelo para que não houvesse um aquecimento excessivo e consequente desnaturação das proteínas.

O lisado foi centrifugado por 45 minutos, a 4°C, 13000rpm, para a separação das proteínas solúveis e insolúveis, sendo a fração solúvel foi submetida a uma cromatografia de afinidade para obtenção das proteínas de interesse (seção 3.6.3).

### 3.6.3 Cromatografia de afinidade

Os vetores de expressão em *E. coli* usados nesse trabalho (pET28a, pETDUET-1 e pETSUMO) possuem uma sequência de 6 histidinas (HIS-tag) fusionada a porção N-terminal das proteínas de interesse, sendo que esta extensão possui afinidade com íons cobalto ( $\text{Co}^{2+}$ ) ou níquel ( $\text{Ni}^{2+}$ ).

Para a separação das proteínas recombinantes foi usada uma coluna TALON™ *Metal Affinity Resin*, com íons  $\text{Co}^{2+}$  imobilizados na resina, acoplada ao sistema ÄKTA *explorer10* (GE).

A purificação das proteínas expressas em pETSUMO foi realizada em 3 etapas, a cromatografia de afinidade inicial, a clivagem com SUMO protease e retirada do imidazol e uma segunda cromatografia por afinidade.

Tabela 8 – Relação das proteínas purificadas e composição do tampão A.

Proteína Recombinante	Vetor de expressão	Tampão A
U5-15K truncada	pETSUMO	20mM Tris-HCl pH 7,5 100mM NaCl
U5-15K pc		
U5-15K C66A		
U5-15K S120A		
U5-15K S126A		
U5-15K S137A		
U5-15K S144A		
U5-102K N-Term	pET 28 <sup>a</sup>	50mM Tris-HCl pH 7,5 150mM NaCl
U5-102K		20mM Tris-HCl pH 7,5 150mM NaCl

Fonte: Elaborada pela autora.

Na primeira etapa (afinidade 1) a coluna foi equilibrada com aproximadamente 5 volumes de tampão A, que varia para cada proteína, como demonstrado na tabela 8, em seguida foi aplicado o sobrenadante proveniente da lise de 1L de cultura de *E. coli* (seção 3.6.2). A coluna foi novamente lavada com tampão A, aproximadamente 8 volumes, e aplicou-se um gradiente de concentração de imidazol (de 0 a 500mM), a eluição das proteínas recombinantes se deu a partir da presença de aproximadamente 150mM de imidazol no tampão A.

A essa fração das proteínas, foram adicionados 1mg de SUMO protease para cada 10mg de proteína recombinante para a remoção da HIS-tag e proteína de fusão SUMO. Durante essa etapa, a proteína foi mantida em gelo por aproximadamente duas horas. O

imidazol foi retirado usando uma coluna de dessalinização *HiTrap Desalting* (GE) acoplada ao *ÄKTA explorer10* (GE).

Na terceira e última etapa (afinidade 2) a solução proteica dessalinizada foi novamente aplicada em uma coluna *TALON™* para a separação das proteína recombinante, agora sem HIS-tag, da proteína de fusão SUMO e da SUMO protease, ambas contendo a cauda de 6 histidinas.

A proteína U5-102K, expressa com pET28a, foi purificada em duas etapas, uma cromatografia de afinidade e uma diálise a 4°C, por aproximadamente 16h para a retirada do imidazol. Essa cromatografia, segue o mesmo protocolo da afinidade 1 mencionada anteriormente.

As proteínas foram concentradas por meio de ultra-filtração em membrana Milipore com corte de 10kDa, 2500xg a 4°C. A concentração de proteína foi determinada utilizando espectrofotômetro Nanodrop™ 1000 (Thermo Scientific).

### **3.7 Western Blot**

A técnica de *western blot* foi usada tanto para a confirmação da expressão (3.6.1) e purificação da proteína U5-102K (3.6.3) quanto para a confirmação da transfecção da forma procíclica da cepa de *T. brucei* 427 com pc-PTP-neo\_102k e pc-PTP-neo\_Pr43 (3.5.2).

Para isso, foi feita a eletroforese em gel SDS-PAGE 10% das amostras de proteína e em seguida a transferência das bandas para uma membrana de nitrocelulose Hybond ECL (Amersham Bioscience) usando o sistema Mini Trans-Blot® Electrophoretic Transfer Cell (Bio-Rad) por 1 hora, 100V e 400 mA em tampão de transferência (2,5mM Tris base; 19,2mM glicina). A membrana foi corada com Ponceau (0,2%) para confirmar a transferência e lavada com PBST (PBS; 0,1% Tween 20).

Após a retirada do corante, a membrana foi bloqueada por uma hora a temperatura ambiente com PBST contendo 5% de leite em pó desnatado. Em seguida, a membrana foi lavada e incubada em PBST com 2,5% de leite e o anticorpo primário. Foram feitas três lavagens de 5 minutos com PBST e a membrana foi incubada com o anticorpo secundário em PBST com 2,5% de leite. A relação dos anticorpos usados, bem como seus títulos e os tempos de incubação estão representados na tabela 9.



Tabela 9 – Especificações dos anticorpos usados nos *Western Blots* desse trabalho.

Vetores usados nas clonagens	Anticorpo primário	Título	Tempo de incubação	Anticorpo secundário	Título	Tempo de incubação
<b>pET28a</b>	Contra cauda de histidinas (anti-HIS)	1:15000	2 horas	Contra camundongo	1:10000	45 minutos
<b>PETSUMO</b>	Contra cauda de histidinas (Anti-HIS)	1:15000	2 horas	Contra camundongo	1:10000	45 minutos
<b>pC-PTP-NEO</b>	Contra proteína A	1:5000	2 horas	Contra coelho	1:5000	1 hora

Fonte: Elaborada pela autora.

Foram feitas três lavagens com PBST e a membrana foi revelada com uma solução de luminol (10mM TrisHCl pH8,5; 90 mM ácido p-cumárico em DMSO; 250mM luminol em DMSO) e peróxido de hidrogênio usando o aparelho C-Digit (Li-COR) e analisadas com o software Image Studio Digs 3.1.

### 3.8 Caracterização da atividade de autoclivagem da proteína U5-15K

#### 3.8.1 Teste de clivagem das proteínas que sofreram mutação

Para o teste de clivagem, as proteínas U5-15KC66A, U5-15K truncada, U5-15K pc, U5-15K S120A, U5-15K S126A, U5-15K S137A e U5-15K S144A foram concentradas até aproximadamente 1 mgml<sup>-1</sup>. As mutantes U5-15KC66A, U5-15K truncada, U5-15K S120A, U5-15K S137A e U5-15K S144A concentradas foram separadas em alíquotas de 500µl em microtubos de 1,5µl. Dessas alíquotas, quatro receberam inibidores de protease: Leupeptina (10µM), quimostatina (6µgml<sup>-1</sup>), EDTA (1mM) e PMSF (2mM), e uma foi mantida apenas no tampão de purificação (tabela 7) como controle da atividade de autoclivagem.

Essas microtubos foram mantidos a temperatura ambiente e alíquotas foram retiradas de 12 em 12 horas, fervidas com tampão de amostra (188mM Tris HCl pH 6,8; 6% SDS; 30% glicerol; 15% β-mercaptoetanol; 0,01% azul de bromofenol), e aplicadas em gel SDS-PAGE 20%, para a visualização das bandas, o gel foi corado com solução corante (0,2% Coomassie brilliant blue R-250; 50% metanol; 10% ácido acético) e descorado com solução de metanol (45%) e ácido acético (10%).

Para as mutantes U5-15K S126A e U5-15K pc, foi realizado apenas o teste com a alíquota controle (sem adição de inibidores).

### **3.9 Ensaios de interação entre as proteínas U5-102K e U5-15K**

#### **3.9.1 Espalhamento dinâmico de luz (DLS)**

Para esse experimento, utilizou-se o equipamento *Zetasizer UV (Malvern)*. As amostras das proteínas U5-15K, U5-102K e uma mistura 1:1 das duas proteínas foram concentradas até  $1 \text{ mgmL}^{-1}$  em dois tampões diferentes (50mM Tris HCl, pH7,5; 150mM NaCl e 150mM KCl; 20mM Tris HCl; pH7,5; 3mM  $\text{MgCl}_2$ ; 0,1% Tween). Foi usada uma cubeta de quartzo para a realização das medidas. A seleção das medidas representativas e o cálculo dos raios hidrodinâmicos e estado oligomérico das proteínas foi realizado pelo software *ZETASIZER*.

#### **3.9.2 Gel nativo**

Outra análise da formação do subcomplexo entre as proteínas U5-15K e U5-102K foi feita por eletroforese em gel nativo 20%. Para isso foram preparados tampões e géis de poliacrilamida não desnaturantes (tampão de corrida: 25mM Tris HCl, pH 8,3; 192mM glicina; tampão de amostra 2x concentrado: 62,5mM TrisHCl, pH6,8; 25% glicerol; 0,5% azul de bromofenol; gel nativo 20%: 20% acrilamida; 120mM Tris HCl, pH 8,8; 0,01% persulfato de amônio; 0,1% TEMED) .

O marcador de massa molecular foi preparado com  $2 \text{ mgmL}^{-1}$  de Citocromo C (aproximadamente 12,4kDa) e  $5 \text{ mgmL}^{-1}$  de albumina bovina (aproximadamente 66kDa). Foram aplicadas amostras da U5-15K e U5-102K a  $1 \text{ mgmL}^{-1}$  e uma alíquota da mistura 1:1 das duas proteínas. A eletroforese foi realizada a 5mA, 4°C, por aproximadamente 12h. Para a visualização das bandas, o gel foi corado com solução corante (0,2% Coomassie brilliant blue R-250; 50% metanol; 10% ácido acético) e descorado com solução de metanol (45%) e ácido acético (10%).

### **3.10 Ensaios de cristalização**

#### **3.10.1 Fluorimetria diferencial de varredura (DSF)**

Para os experimentos iniciais de DSF, foram usadas concentrações de 1  $\mu\text{M}$ , 2  $\mu\text{M}$ , 4  $\mu\text{M}$  e 6  $\mu\text{M}$  de U5-15K e U5-15K sem alça em tampão fosfato de potássio (20mM pH 7,4). Também foram testadas três concentrações da sonda Sypro Orange (25x, 50x e 75x). As medidas foram feitas em duplicada em uma placa de 96 poços no aparelho CFX96™ Real Time System (Bio-Rad) e a análise das curvas foi feita com o software BioRad CFX manager. A partir das curvas foi calculada a temperatura de desnaturação média ( $T_m$ ).

### 3.10.2 Ensaios de cristalização

Os ensaios de cristalização desse trabalho foram feitos com as mutantes U5-15K sem alça e U5-15K S126A, em placas de 96 poços, através do método de difusão de vapor.

Para a U5-15K sem alça na concentração de 3  $\text{mgml}^{-1}$  foram utilizados os kits comerciais Crystal Screen HT (Hampton Research), MORPHEUS HT, SG1 Screen™ e JCSG-plus HT (Molecular Dimensions) e para a U5-15K S126A na concentração de 4  $\text{mgml}^{-1}$  foram utilizados os kits comerciais MORPHEUS HT, SG1 Screen™, BCS Screen™ e JCSG-plus HT (Molecular Dimensions) e Index HT (Hampton Research).

Para ambas as proteínas recombinantes, empregou-se a técnica da gota apoiada (*sitting drop*) com auxílio do robô Honeybee 393 (Molecular Dynamics), no caso da U5-15K truncada, e pelo robô *Crytal Gryphon* (Art Robinson Instruments), no caso da U5-15k S126A. As placas foram mantidas em uma sala a 17°C.

## 3.11 Caracterização dos parceiros de interação das proteínas U5-102K e Prp43

### 3.11.1 Preparo do extrato proteico (Y4)

Para o preparo de extrato proteico, 4 litros de cultura de TbPCF 427 foram centrifugados por 10 minutos a 2700 x g, 2 °C. O sobrenadante foi descartado e as células transferidas para um tubo de plástico de 50ml.

As células foram novamente centrifugadas por 7 minutos a 2700 x g, 2°C e o sobrenadante foi descartado. O sedimento foi lavado com 40 mL de tampão *Tryps Wash* (100mM NaCl; 3mM MgCl<sub>2</sub>; 20mM Tris-HCl pH 7,5) gelado e centrifugadas por 7 minutos a 2700 x g, 2°C. Esse procedimento foi realizado mais uma vez, o sobrenadante descartado e as células ressuspensas em tampão de transcrição gelado (150 mM sucrose; 20 mM ácido

glutâmico; 20 mM HEPES-KOH pH 7,7; 3 mM MgCl<sub>2</sub>) em uma razão de 1:1 do volume de células.

As células foram mantidas no gelo por 20-30 minutos, adicionou-se 1  $\mu\text{ml}^{-1}$  de DTT (1M), leupeptina (10  $\text{mgml}^{-1}$ ) e aprotinina (10  $\text{mgml}^{-1}$ ) e procedeu-se para lise celular por pressão em *French Press Cell Disrupter* (Thermo Electron Corporation). Foram utilizados 6 ciclos de de 1 minuto cada, com uma pressão de 2.000 psi.

Alíquotas de 1ml do extrato foram divididas em microtubos de 1,5ml, congeladas em nitrogênio líquido e armazenadas a -80°C. Cada fração foi chamada de Y4. Esse procedimento foi realizado tanto para as culturas transfectadas com pc-PTP-neo\_102K quanto pc-PTP-neo\_Prp43.

### 3.11.2 Purificação em *Tandem*

A purificação das proteínas de *T. brucei* usando a técnica de PTP-tag foi feita em dois dias.

No primeiro dia, foram adicionados 100  $\mu\text{l}$  de tampão de extração (1,5M KCl; 19,5mM Tris HCl pH 7,7; 3mM MgCl<sub>2</sub>; 0,5mM DTT; 1% Tween 20) a dez microtubos de 1,5 mL previamente gelados. Dez alíquotas de Y4 (seção 3.11.1) foram descongeladas e transferidas para esses microtubos, com cuidado para não lizar o núcleo, que foram mantidos em gelo por 20 minutos, sendo agitados a cada 5 minutos.

Em seguida, foram centrifugados por 15 minutos a 14000 rpm, 2 °C e o sobrenadante foi transferido para novos microtubos previamente gelados e foram novamente centrifugados por 3 minutos à 21.000 x g, temperatura ambiente. Os sobrenadantes foram reunidos em um tubo cônico de 15 mL livre de RNase.

Em outro microtubo, foi dissolvido um comprimido de inibidor de protease (cOmplete mini EDTA-free da Roche) em 1ml de PA-150 (150mM KCl; 20mM TrisHCl pH 7,7; 3mM MgCl<sub>2</sub>; 0,5mM DTT; 0,1% Tween 20) e 500  $\mu\text{l}$  foram adicionados ao sobrenadante, o restante foi reservado para uso nas etapas subsequentes.

Um suporte foi preparado com aproximadamente 260  $\mu\text{l}$  de resina IgG Sepharose <sup>TM</sup> (Amersham Biosciences) equilibrada com 50ml de tampão PA-150. Em seguida, o lisado com inibidor de protease foi aplicado à resina, o suporte foi devidamente fechado e mantido sob agitação branda a 4°C por aproximadamente 2 horas.

Após este período, esperou-se que a resina decantasse, o eluato coletado em um novo tubo cônico de 15mL. Foram separados 5  $\mu$ L da amostra, adicionou-se 45  $\mu$ l de tampão de amostra, e fervido para desnaturação da proteínas por 5 minutos.

Adicionou-se 10 mL de tampão PA-150 gelado, o sistema foi novamente devidamente fechado e mantido sob agitação branda por 10 minutos a 4°C. Após decantação da resina foram adicionados 15 ml de tampão PA-150. Após estas etapas de lavagem acrescentou-se 2ml de tampão TEV (1,5 M KCl; 200mM Tris HCl pH 7,7; 15mM MgCl<sub>2</sub>; 5mM EDTA pH 8; 10mM DTT; 1% Tween 20) e 300U de ProTEV (Promega). A coluna foi devidamente fechada e mantida sob agitação a 4°C por aproximadamente 16 horas.

No segundo dia, a coluna foi retirada da agitação e esperou-se que a resina decantasse. O sobrenadante foi coletado e 20  $\mu$ l foram separados e fervidos com 30  $\mu$ l de tampão de amostra. A esse sobrenadante foram adicionados 4ml de PC-150 (150mM KCl; 20mM Tris HCl pH 7,7; 3mM MgCl<sub>2</sub>; 1mM CaCl<sub>2</sub>; 0,5mM DTT; 0,1% Tween 20), 7,5  $\mu$ l de CaCl<sub>2</sub> 1M e 500  $\mu$ l da solução de inibidor de protease previamente preparada.

Enquanto isso, preparou-se uma coluna contendo aproximadamente 260  $\mu$ l anti-proteína C Affinity Matrix (Roche), que foi lavada com 50 ml de PC-150 e em seguida foram adicionados os 6mL de sobrenadante. O suporte foi devidamente fechado e mantido sob agitação por aproximadamente 2 horas a 4°C.

A coluna foi então retirada da agitação e esperou-se que a resina decantasse para então coletar o novo sobrenadante, 20  $\mu$ l foram separados e fervidos com 30  $\mu$ l de tampão de amostra. Em seguida, foram adicionados mais 10 mL de PC-150 à resina e o suporte fechado foi novamente mantido sob agitação por 10 minutos, a 4°C.

Ainda na sala à 4°C, o suporte foi retirado da agitação, esperou-se que a resina decantasse e foram adicionados mais 20 mL de PC-150. Essa lavagem foi repetida mais duas vezes.

O suporte da resina foi então cortado aproximadamente na marca de 1ml e foram adicionados 600  $\mu$ l de tampão de eluição EGTA/EDTA (5mM Tris HCl pH 7,7; 10mM EGTA; 5mM EDTA pH 8; 0,01mgml<sup>-1</sup> Leupeptina); a coluna foi devidamente fechada e agitada por 5 minutos a temperatura ambiente. O eluído foi coletado em um microtubo de 2ml. Esse procedimento foi repetido duas vezes e o eluído foi armazenado a -80°C.

As proteínas do eluído foram concentradas por centrifugação à vácuo até um volume de aproximadamente 800  $\mu$ l e foram adicionados 15  $\mu$ l de StrataClean™ Resin (Agilent Technologies) e agitadas por 5 minutos a temperatura ambiente. Em seguida, essa suspensão

foi centrifugada por 5 minutos a 3000 x g e o sobrenadante foi separado. À resina foram adicionados 30 µL de tampão de amostra e fervida por 5 minutos.

### 3.11.3 SDS-PAGE

As alíquotas reservadas durante o processo da seção anterior foram aplicadas em SDS-PAGE 12%, tomando-se o cuidado de não deixar a frente de corrida sair do gel. As placas e cuba de corrida foram tratadas com metanol para evitar a contaminação com queratina. O gel foi corado com Sypro Ruby20 de acordo com as recomendações do fabricante e as bandas correspondentes ao eluato foram cortadas e armazenadas a -20°C para análise por espectrometria de massas.

### 3.11.4 Espectrometria de massas

A espectrometria de massas e o preparo das amostras analisadas foram realizados no Laboratório Especial de Toxinologia Aplicada do Instituto Butantan em parceria com a Dra Júlia Chagas Pinheiro da Cunha.

Antes da aplicação das amostras, as bandas previamente cortadas (2.12.3) foram descoradas com uma solução contendo 50% de metanol e 5% de ácido acético por 3 horas a temperatura ambiente. Em seguida foram desidratadas com acetonitrila 100% por 5 minutos.

Após esse tempo, adicionou-se 10mM de DTT e os fragmentos de gel foram incubados por 30 minutos a temperatura ambiente. Em seguida foram adicionados 30µl de solução de iodoacetamina (50mM iodoacetamina, 100mM bicarbonato de sódio) por 30 minutos.

Os fragmentos foram lavados com 100mM de bicarbonato de amônio por 10 minutos e novamente desidratados com acetonitrila 100% e hidratados com bicarbonato de amônio. O passo de desidratação repetido mais uma vez. A proteólise foi realizada com 50ngµl<sup>-1</sup> de Tripsina (Sigma) em 50mM de bicarbonato de amônio, a 4°C por aproximadamente 16h.

No dia seguinte, foi adicionado ácido fórmico 5% até recobrir todo o gel e prosseguiu-se com incubação por 10 minutos a temperatura ambiente. O sobrenadante foi coletado e adicionou-se solução contendo 50% de acetonitrila e 5% de ácido fórmico por 10 minutos a temperatura ambiente.

As amostras foram secas a vácuo e ressuspensas com ácido fórmico 0,5% e dessalinizadas em uma resina do tipo zip-tip (Milipore).

A análise dos fragmentos de proteína vindos da purificação da 102k foi feita em ESI-TRAP e dos fragmentos da purificação da PRP43 em LTQ-Orbitrap Velos (Thermo Scientific) a 1,9kV e 200°C, em modo de aquisição dependente dos dados, com cobertura de 200 a 20000m/z e resolução de 3000Da. Os dados foram analisados no Mascot ([www.matrixscience.com/](http://www.matrixscience.com/)) e a sequência de proteínas obtidas foram identificadas por busca em banco de dados de sequência de *T. brucei* (TriTrypDB).





## 4 RESULTADOS E DISCUSSÃO

### 4.1 Estudo da atividade de autoclivagem

#### 4.1.1 Mutação sítio dirigida

Estudos anteriores<sup>79</sup> mostraram que a proteína U5-15K de *T. brucei* apresenta uma atividade de autoclivagem na sua porção C-terminal. Essa atividade também foi observada em sua homóloga humana<sup>66</sup> e testes realizados com inibidores de protease mostraram que essa atividade é abolida, na proteína de *T. brucei*, por inibidores de serino e cisteíno proteases.<sup>64,66</sup>

Partindo dessas informações, foram feitas 4 mutações buscando identificar o resíduo potencialmente envolvido na clivagem e inativa-lo. Essas proteínas mutantes foram denominadas U5-15K C66A, U5-15K S120A, U5-15K S126A, U5-15K S137A, U5-15K S144A, como demonstrado na figura 24. Também foi feita uma mutação para inativação dessa atividade alterando a sequência de aminoácidos em torno do possível resíduo de clivagem, essa proteína mutante foi denominada U5-15K pc.

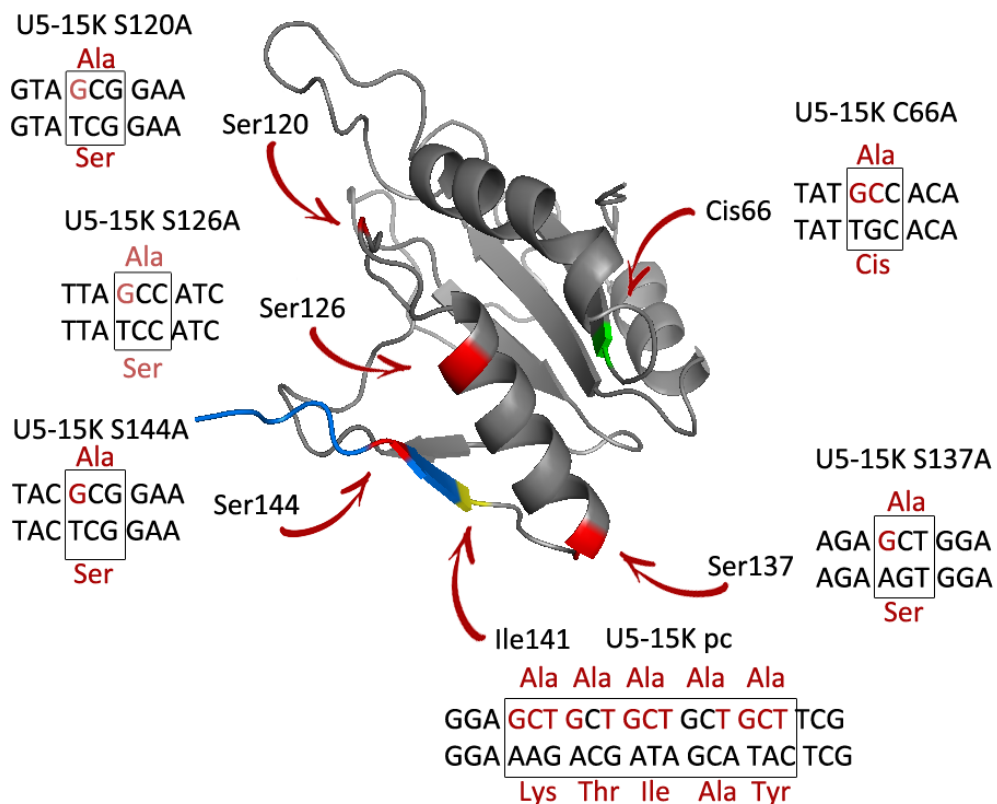


Figura 24- Modelo estrutural da U5-15K de *T. brucei* com as serinas mutadas marcadas em vermelho e a cisteína em verde, o possível ponto de clivagem em amarelo e a C terminal em azul.

Fonte: Elaborada pela autora.

O cálculo do ponto de clivagem foi feito por meio de dois experimentos<sup>64</sup> um para definir se ela ocorria no C-terminal ou N-terminal, ilustrado na figura 25, e um para apontar aproximadamente em que aminoácido.

No primeiro experimento, foram coletadas alíquotas da U5-15K purificada, ainda fusionada à proteína SUMO, e foi feita a comparação dos tamanhos das bandas dessa construção com as da proteína clivada, o que mostrou que a clivagem ocorria no C-terminal.<sup>64</sup>

No segundo experimento, foi feita a curva do deslocamento das proteínas em SDS-PAGE 20% em função da massa e a partir dessa curva foi calculada a massa da proteína clivada, chegando em um ponto de clivagem aproximadamente em torno da isoleucina 141.<sup>64</sup>

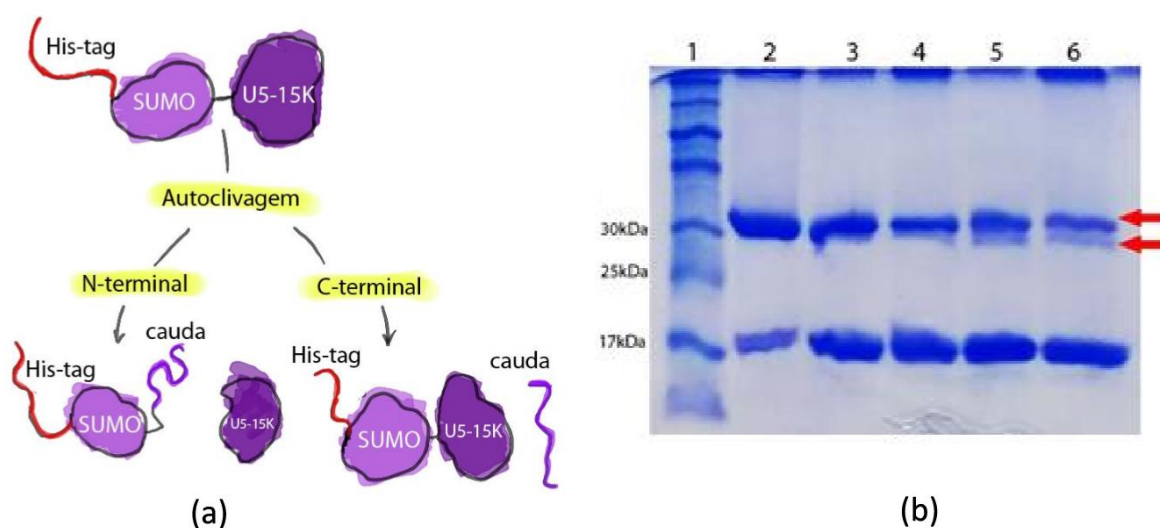


Figura 25 – (a) Representação das diferenças de tamanho da proteína clivada, dependendo da posição do ponto de clivagem, se próxima ao N ou C terminal. (b) SDS-PAGE 15% do teste de clivagem. Coluna 1. Marcador de massa molecular, 2. Alíquota da proteína logo após a purificação, 3. Alíquota da proteína após 12h após a purificação, 4. 24h após a purificação, 5. 36h após a purificação e 6. 48h após a purificação. As setas vermelhas indicam a U5-15K não clivada e a U5-15K clivada. A diferença de massa das duas bandas indica que a proteína é clivada no C-terminal.

Fonte: Adaptadas de LIMA<sup>64</sup>

Para estudar a hipótese da U5-15K possuir potencial de auto-clivagem, atividade similar a de uma cisteína protease, foi feita a substituição do resíduo Cys66, única cisteína presente na proteína, por uma alanina. Se essa hipótese se mostrasse verdadeira, seria pouco provável que a auto-clivagem desempenhasse algum papel no *splicing*, uma vez que esse resíduo está distante de ponto de clivagem.

Quanto à possibilidade de U5-15K se comportar como uma serino protease, das 9 serinas presentes em sua sequência foram escolhidas as serinas 120, 126, 137 e 144, por estarem localizadas nas proximidades do ponto de clivagem (figura 24).

Para a construção da U5-15k pc, foram dois aminoácidos anteriores à isoleucina 141 e dois posteriores, como mostrado nas figuras 24 e 26. Desses 5 resíduos, um era uma alanina (Ala 142) e permaneceu sem alterações na proteína mutante. Essas mutações foram realizadas como descrito na seção 3.3 desse trabalho.

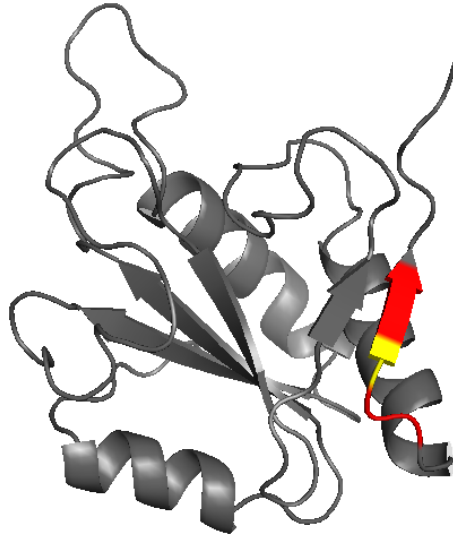


Figura 26 – Localização da isoleucina 141 em amarelo, suposto ponto de clivagem calculado a partir do deslocamento das bandas presentes no marcador de massa molecular e da proteína clivada, em um gel SDS-PAGE 20%<sup>64</sup>, e dos demais resíduos substituídos por alaninas em vermelho.

Fonte: Elaborada pela autora.

#### 4.1.2 Obtenção das proteínas recombinantes

Os DNAs mutantes foram clonados no vetor pETSUMO e transformados em *E. coli* BL21(DE3), como descrito na seção 3.4.2. Das placas resultantes de clonagem, foram escolhidas algumas colônias para a realização dos testes de expressão iniciais, feitos em 5mL de meio LB. Como se tratavam de mutações pontuais na sequência da U5-15K a primeira condição de expressão testada foi igual a usada na expressão da proteína nativa descrita na seção 3.6.1.

Para a confirmação da expressão, alíquotas da cultura foram retiradas antes da indução com IPTG e depois de aproximadamente 16h da indução, como descrito na seção 3.6.1. A análise das proteínas expressas foi feita em SDS-PAGE 15% (figura 27) corado com coomassie blue e a confirmação por *western blot* (seção 3.7) usando anticorpo contra a cauda de 6 histidinas (His-Tag).

Na figura 26 é possível observar um aumento na espessura das bandas indicadas pela seta vermelha, pouco acima de 31 kDa, nas 3 culturas após 16 horas da indução com IPTG

(colunas 3, 5 e 7), essas bandas também foram marcadas no western blot, indicando a expressão das proteínas de interesse.

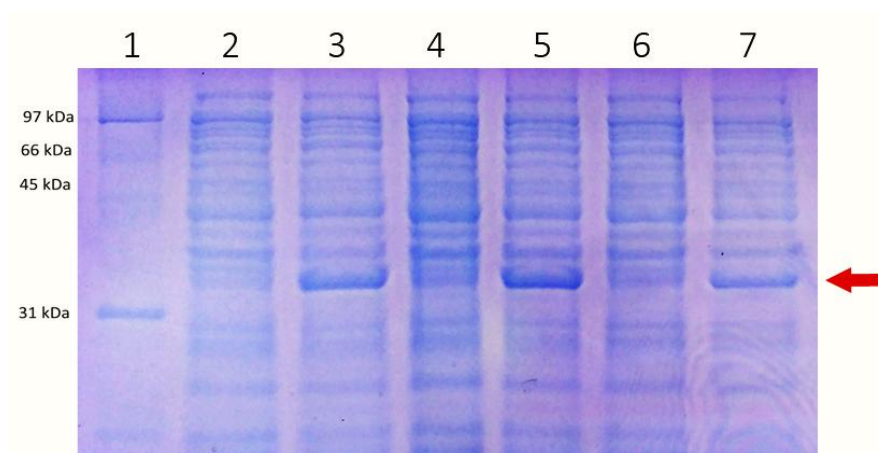


Figura 27 – SDS-PAGE 15% do teste de expressão das mutantes U5-15K S120A, U5-15K S137A e U5-15K S144A. 1. Marcador de massa molecular, 2. U5-15K S120A antes da indução com IPTG, 3. U5-15K S120A depois de 16h da indução, 4. U5-15K S137A antes da indução, 5. U5-15K S137A 16h depois da indução, 6. U5-15K S144A antes da indução, 7. U5-15K S144A após a indução. A seta vermelha indica a posição das bandas de expressão das três mutantes.

Fonte: Elaborada pela autora.

A purificação das proteínas mutantes foi feita por cromatografia de afinidade como descrito na seção 3.6.3 9, e os cromatogramas estão representados na figura 28.

Os cromatogramas da figura 28 foram gerados no sistema ÄKTA a partir da medida da absorbância da solução que sai da coluna em função do tempo (curva preta) e a curva vermelha indica a porcentagem de tampão B (Tampão A com 300mM de imidazol) no tampão de eluição. Os primeiros picos a aparecerem, com 0% de tampão B, representam as proteínas de *E. coli* que não tem afinidade pelos íons de cobalto imobilizados na coluna. Após a passagem dessas proteínas, lava-se a coluna com tampão A até que não haja proteínas presentes no tampão que passa pelo sensor do equipamento.

Ao aumentar a concentração de imidazol no tampão de eluição (concentração de tampão B), a proteína de interesse, que interagem com os íons cobalto através da cauda de histidina, é eluída. Esses picos estão indicados pelas setas vermelhas.

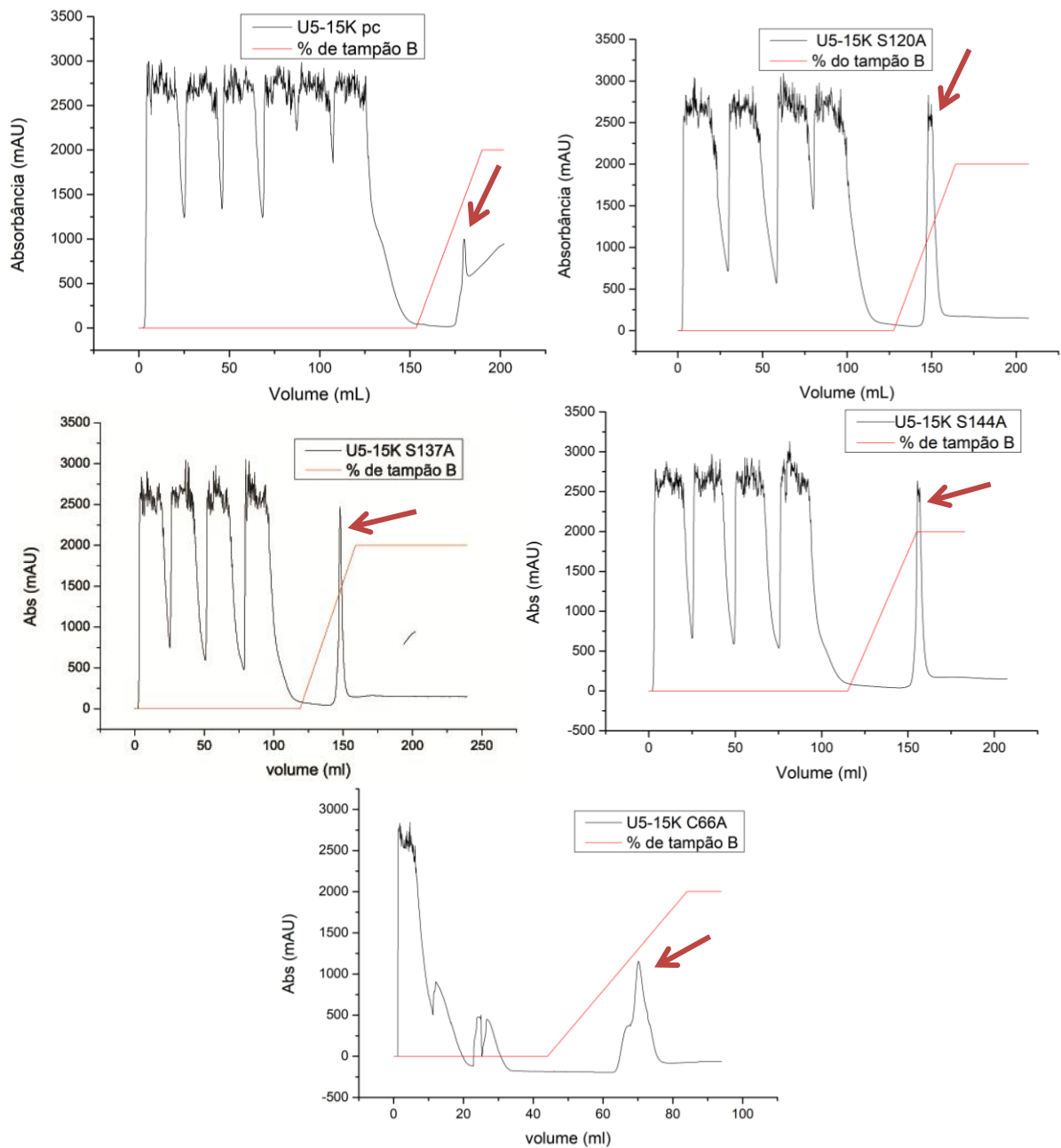


Figura 28 – Cromatograma das purificações das proteínas U5-15K pc, U5-15KS120A, U5-15K S137A, U5-15KS144A, U5-15K C66A. A linha preta representa a absorbância a 280nm e a vermelha a porcentagem de tampão B no gradiente de imidazol (ponto mais baixo 0%, ponto mais alto 100%). Os picos indicados pelas setas vermelhas representam a eluição das proteínas de interesse.

Fonte: Elaborada pela autora.

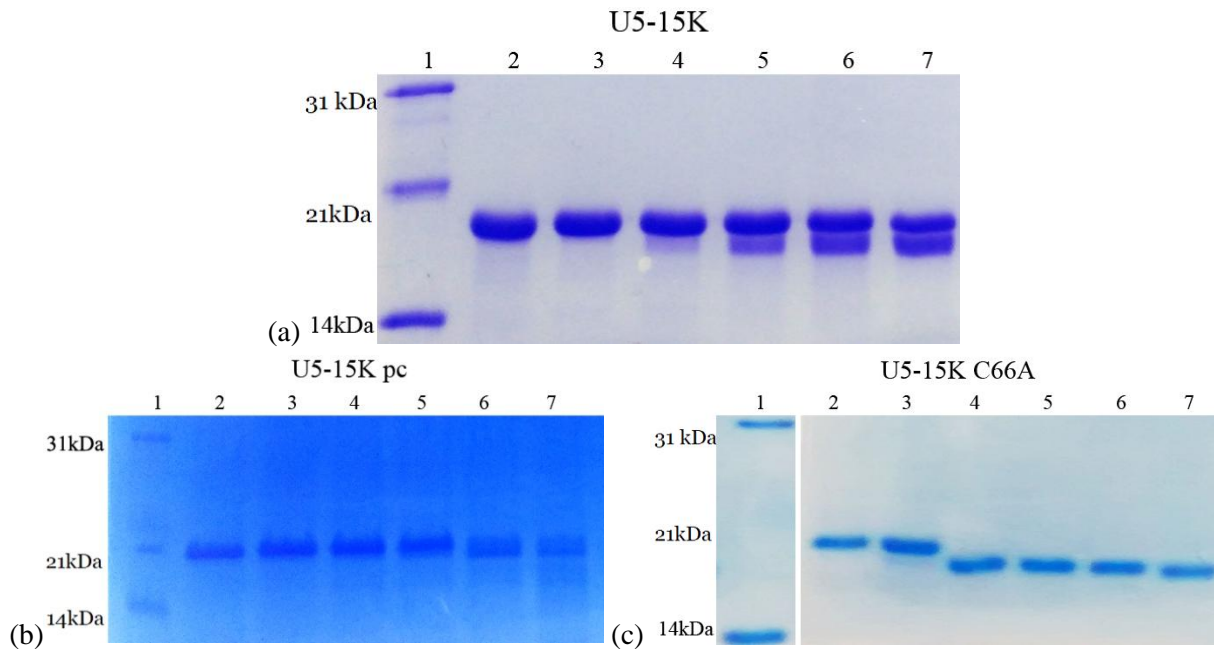
Com base nesses resultados, pode-se dizer que as proteínas mutantes foram eluídas a partir de aproximadamente 50% de tampão B no tampão de eluição, o que equivale ao tampão A contendo 150mM de imidazol. Após esse passo foi feita a diálise para a retirada do imidazol, clivagem da proteína de fusão com SUMO protease e nova afinidade para a separação das proteínas de interesse da SUMO e SUMO protease.

É normal que a proteína eluída após a primeira etapa de purificação seja acompanhada de proteínas expressas normalmente pela *E. coli* e que tem afinidade pelos íons de cobalto, a terceira etapa retêm essas proteínas juntamente com a SUMO e SUMO protease, reduzindo a presença de contaminantes na solução final, tornando uma etapa de purificação por exclusão de tamanho desnecessária.

#### 4.1.3 Teste de clivagem

Como já mencionado, tanto a U5-15K de humanos quanto a de *T. brucei* possuem uma atividade de autoclivagem. No caso da U5-15K de *T. brucei*, o início dessa atividade é observado após 24 horas de sua purificação a partir de um extrato proteico de *E. coli*. Ambas proteínas tem essa atividade abolida por inibidores de protease como ácido 6-aminoexanóico e quimostatina. Diferente da proteína de humanos, a clivagem da U5-15K de *T. brucei* não foi inibida por EDTA.<sup>64,66</sup>

Com isso em mente, e tendo as proteínas purificadas em mãos, foram iniciados os testes de clivagem como descrito na seção 3.8.1. Os resultados são mostrados a seguir (figura 29).



continua



**continuação**

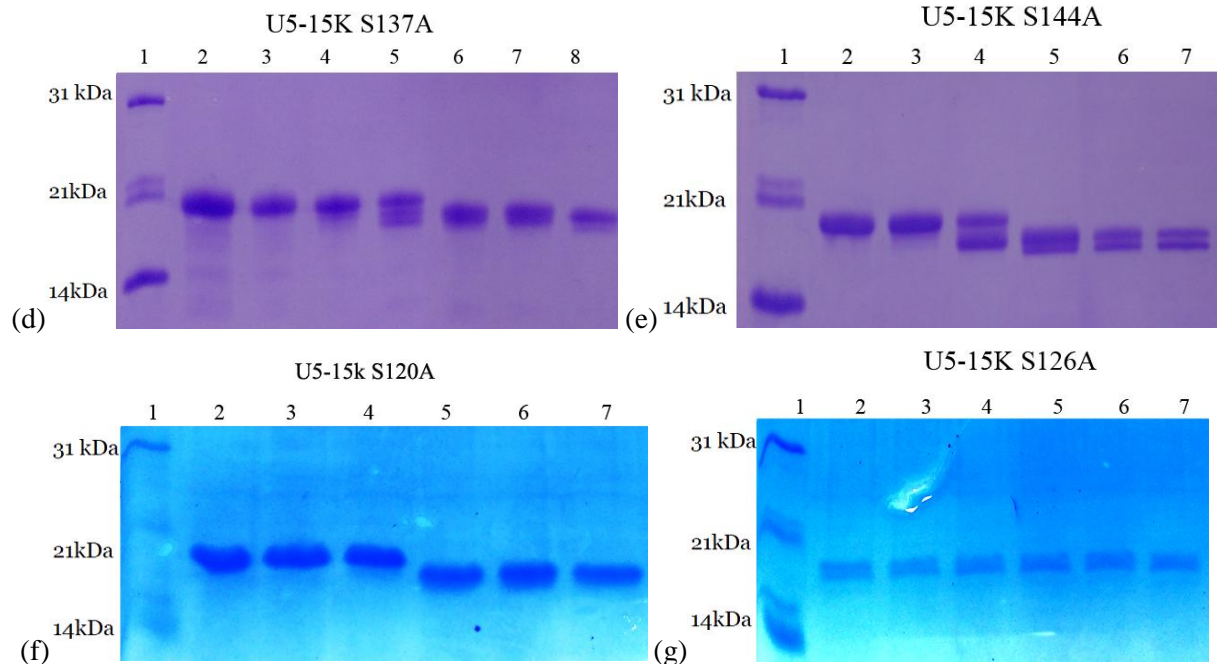


Figura 29 – SDS\_PAGE 20% dos testes de clivagem (a) da U5-15K não mutada, (b) mutante U5-15K pc, (c) mutante U5-15K C66A, (d) mutante U5-15K S137A, (e) mutante U5-15K S144A, (f) mutante U5-15K S120A, (g) mutante U5-15K S126A. Em todas as figuras, a coluna 1 mostra o marcador de massa molecular, a coluna 2 uma alíquota da proteína purificada, 3 uma alíquota 12h após a purificação, 4 24h após a purificação, 5 36h após a purificação, 6 48h após a purificação, 60h após a purificação e 7 72h após a purificação.

Fonte: Elaborada pela autora.

Nos testes feitos com a proteína mutante U5-15K C66A é possível observar que há o aparecimento de uma banda de massa inferior a partir de 24h (coluna 4 da figura 29 (c)), indicando que a mutação na cisteína 66 não afeta a atividade de autoclivagem. Esse resultado demonstrando que a autoclivagem não se assemelha à catálise exercida por uma cisteína protease.

Resultados similares foram descritos para a U5-15K S120A (surgimento de uma banda de massa inferior a da U5-15K na coluna 4 da figura 29 (f)), o que pode ser decorrente da posição dessa serina em relação ao ponto de clivagem, ou seja, ela se encontra mais distante e fora da área de contato com esse sítio, como pode ser observado na figura 30.

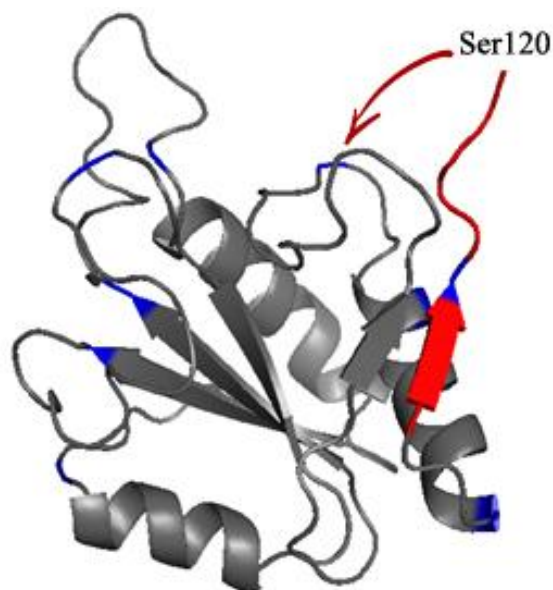


Figura 30 – Posicionamento da Ser120, ponto azul identificado com uma seta vermelha, em relação ao C-terminal clivado, estrutura em vermelho.

Fonte: Elaborada pela autora.

Quanto às mutações U5-15K S144A e U5-15K S137A, observou-se que a atividade de autoclivagem não foi inativada, além disso, comparando-se as bandas presentes no SDS-PAGE dessas variantes com as da U5-15K nativa, pode-se notar o aparecimento de uma terceira banda. Esse fenômeno é evidente no tempo de 36h para a U5-15K S137A e para os tempos de 36h, 48h, 60h e 72h da U5-15K S144A (figura 29 (d) e (e)).

Dada a posição calculada anteriormente para o ponto de clivagem, Ile 141, o aparecimento de novos pontos de clivagem na U5-15K S137A e U5-15K S144A sugere que a autoclivagem se dá em uma alanina, especificamente Ala 142, a alanina mais próxima da Ile141, de modo que a inserção de novas alaninas ao redor dela permite a clivagem em mais de um ponto. Outro fato que corrobora essa afirmação é o arraste de banda no SDS-PAGE 20% da U5-15K pc. Nessa proteína mutante, 4 novas alaninas são adicionadas ao redor da Ala 142, resultando em uma variedade de tamanhos de proteínas clivadas, muito próximos uns dos outros, o que no SDS-PAGE aparece como uma arraste (figura 29 (b)).

Somado a isso observamos também o efeito que a quimostatina tem sobre a autoclivagem, na proteína selvagem, essa molécula age como um inibidor da atividade proteolítica, mas se observarmos os efeitos da sua adição às proteínas mutantes U5-15K S137A e U5-15K S144A (figura 31 (a)), notamos o aparecimento de apenas uma banda, ao invés de duas como aconteceu no teste anterior.



O que permite concluir que a clivagem no ponto original (Ala 142) continua a ser inibida pela adição da quimostatina, embora tal inibição não ocorra para os novos pontos de clivagem das mutantes.

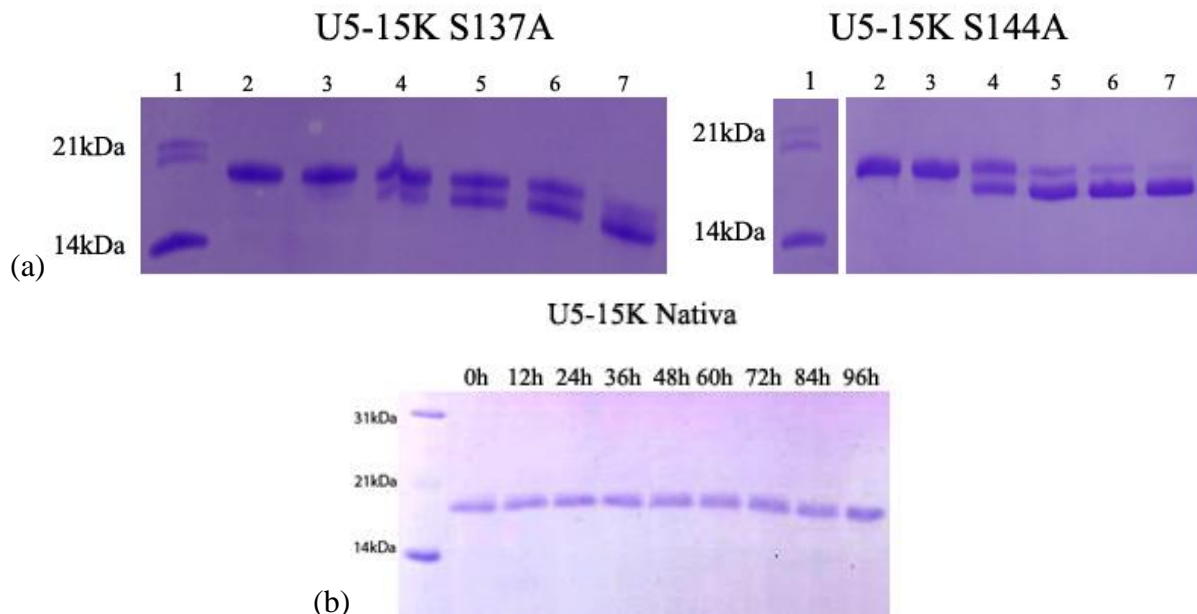


Figura 31 – SDS-PAGE 20% do teste de clivagem (a) das mutantes U5-15K S137A, U5-15K S144A e (b) da U5-15K nativa<sup>64</sup> na presença de  $6 \mu\text{g mL}^{-1}$  de quimostatina. Onde a coluna 1 representa o marcador de massa molecular, a coluna 2 uma alíquota da proteína 12h após a purificação, a 3 24h após a purificação, a 4 36h após a purificação, a 5 48h após a purificação, a 6 60h após a purificação, a 7 72h após a purificação.

Fonte: Elaborada pela autora.

Por fim, a mutação presente na U5-15K S126A aboliu a atividade de autoclivagem, pois, comparando as bandas no SDS-PAGE com as obtidas na autoclivagem da U5-15K selvagem, não há o aparecimento de bandas de menor massa molecular a partir de 24h, como é possível observar na figura 29. Isso sugere que a serina 126 seja essencial para a atividade de autoclivagem, semelhante à de uma serino protease da U5-15K.

Considerando que a estrutura tridimensional da proteína, usada para a escolha das serinas candidatas à mutação, é um modelo obtido através das estruturas da U5-15K de humano, é possível que haja uma variação na posição do C-terminal da proteína de *T. brucei*, aproximando a serina 126 do sítio de autoclivagem, bem como das serinas 137 e 144, o que explicaria o aparecimento dos dois novos sítios de clivagem descritos anteriormente.

A posição da serina 126 também permite que a autoclivagem aconteça enquanto a U5-15K faz parte do tri-snRNP e do complexo B. Trabalhos anteriores mostram que a forma truncada no C-terminal da U5-15K de humano age como inibidor negativo dominante da proteína selvagem em levedura e que há um acúmulo de complexo B na célula<sup>62</sup>, somado a

isso o fato de que a U5-15K não está presente no complexo B catalítico, é possível que a autoclivagem possua um papel no *splicing*, mais especificamente na ativação do complexo B em B\*.

Para avaliar essa hipótese foram planejados experimentos para a inserção da U5-15K S126A em um dos alelos de *T. brucei*, usando o vetor pC-PTP-puro e a deleção do alelo selvagem usando a técnica de nocaute condicional. Essa técnica se baseia na inserção de um cassete de expressão, contendo a ORF da proteína de interesse e marcadores de seleção, flanqueado pelo loxP em um dos alelos selvagens e sua subsequente deleção dependente de Cre recombinase.<sup>93</sup>

Os oligonucleotídeos foram desenhados e foi dado início às clonagens, estes experimentos se encontram em andamento.

## 4.2 Estudo estrutural da U5-15K

### 4.2.1 Mutação sítio dirigida

Em trabalhos anteriores<sup>64</sup>, as tentativas de cristalizar a U5-15K de *T. brucei* a fim de determinar sua estrutura terciária foram infrutíferas. Entretanto, análises do modelo gerado no servidor SWISS-MODEL<sup>80-84</sup> e alinhamento entre as sequências da U5-15K de *T. brucei* e de *H. sapiens* demonstraram a existência de uma alça flexível na proteína de *T. brucei*.

Pelo posicionamento das moléculas de U5-15K de humanos na formação do cristal e alinhando sua estrutura terciária com o modelo da proteína de *T. brucei*, foi possível detectar que a alça, uma estrutura altamente flexível representada em vermelho na figura 32, se encontra em uma região de contato cristalino entre as moléculas, o que impede a formação dos cristais.

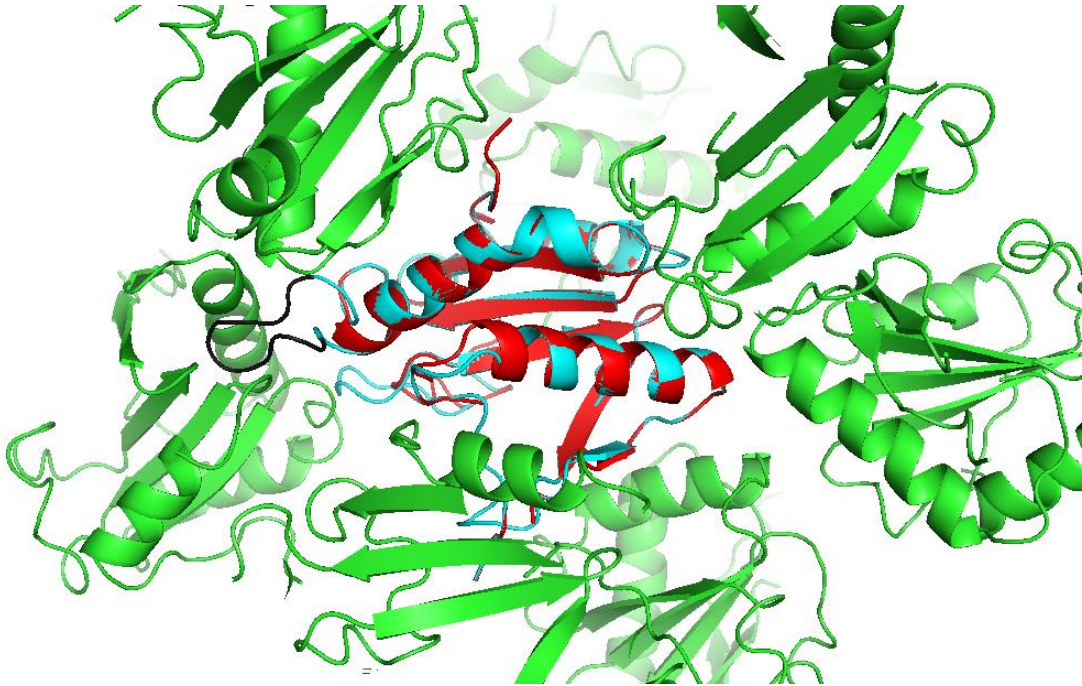


Figura 32 – Alinhamento da U5-15K de *T. brucei* (ciano) com a U5-15K de humano (vermelho) dentro de sua estrutura cristalina (verde). A alça flexível é evidenciada em preto.  
 Fonte: Elaborada pela autora.

Além disso, não é possível descartar que a atividade de autoclivagem, com o passar do tempo, diversifica a população de proteínas presentes na solução, fator adicional que dificulta a cristalização da proteína de interesse.

Para contornar o problema da presença de uma região pouco estruturada e muito flexível, optou-se pela deleção de 11 aminoácidos compreendidos entre as Treoninas 36 e 48.

Foi feita uma nova modelagem por homologia da U5-15K truncada, com intuito de avaliar se essa deleção poderia ter algum efeito negativo na estrutura terciária da proteína. Observando o alinhamento entre as duas estruturas (figura 33) é possível dizer que a deleção da alça poderia resultar no acréscimo de um aminoácido na hélice- $\alpha$  que aparece em primeiro plano, indicada pelo círculo preto na figura abaixo.

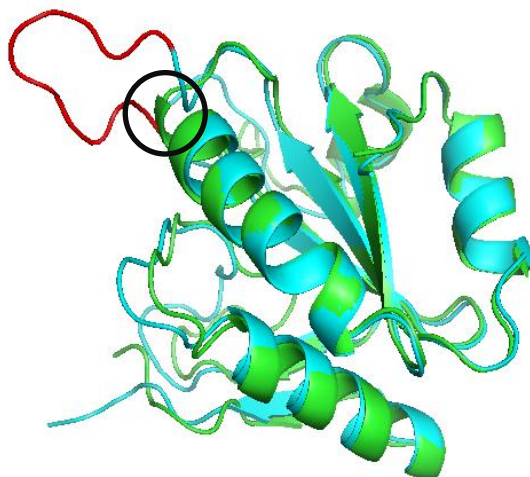


Figura 33 – Alinhamento entre os modelos de estrutura terciária das proteínas U5-15K (ciano) e U5-15K truncada (verde), com a alça flexível em vermelho.

Fonte: Elaborada pela autora.

A mutação sítio dirigida foi feita como descrito na seção 3.3 e o DNA mutante foi clonado em pETSUMO e transformada em *E. coli* BL21(DE3) como descrito na seção 3.4.2.

#### 4.2.2 Obtenção da Proteína recombinante

Da mesma forma que já foi descrito para as outras mutantes na seção 4.1.2, foram feitos testes iniciais de expressão (resultado não mostrado). Após o teste inicial, a proteína foi expressa em 1L de meio LB, como descrito na seção 2.6.1 e purificada por cromatografia de afinidade, o pico de eluição da proteína de interesse está indicado pela seta vermelha (figura 34).

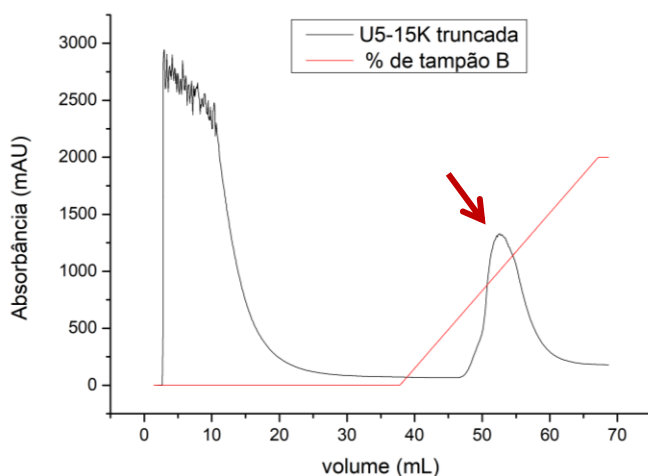


Figura 34 – Cromatograma da purificação da U5-15k truncada, com o pico de eluição indicado pela seta vermelha. A curva vermelha representa a porcentagem de tampão B no gradiente de imidazol e a curva preta a absorbância da amostra a 280nm.

Fonte: Elaborada pela autora.

Com a proteína pura em mãos foram feitos alguns testes para a verificação do estado de enovelamento da proteína.

#### 4.2.3 Fluorimetria diferencial de varredura (DSF)

A fluorimetria diferencial de varredura tem como princípio a diferenciação de uma proteína enovelada e uma desenovelada pela exposição das suas regiões hidrofóbicas a uma sonda fluorescente hidrofóbica. Durante o desenovelamento térmico da proteína, é possível medir os níveis de fluorescência do meio.<sup>85</sup>

Para uma proteína globular, a curva típica de desenovelamento térmico está representada na figura 35, com um mínimo de fluorescência enquanto a proteína se encontra enovelada e um máximo de fluorescência quando os bolsões hidrofóbicos estão expostos à sonda.

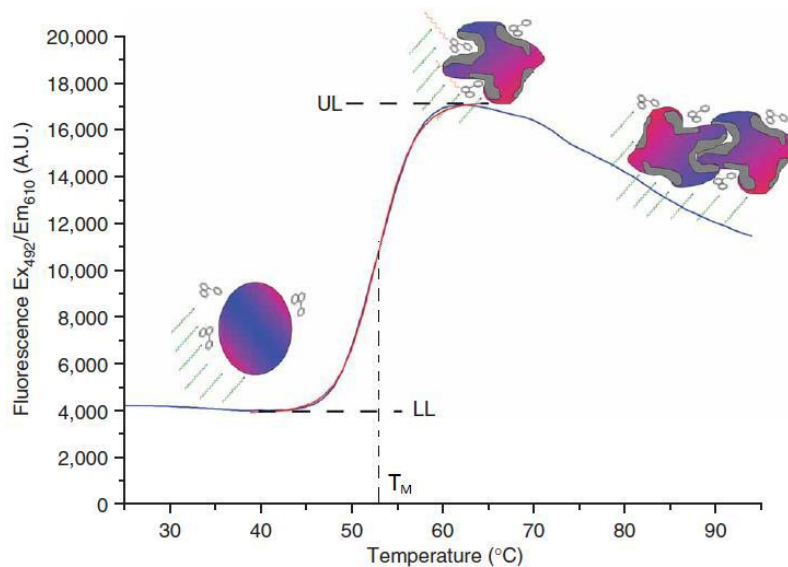


Figura 35 – Representação gráfica de uma curva de desenovelamento térmico de uma proteína globular.  
Fonte: Adaptada de NILSEN<sup>85</sup>

Para avaliar a formação de estrutura terciária da U5-15k truncada, foi realizado o DSF como descrito na seção 3.10.1. Também foi feita a comparação da curva resultante com a encontrada para a U5-15K selvagem.

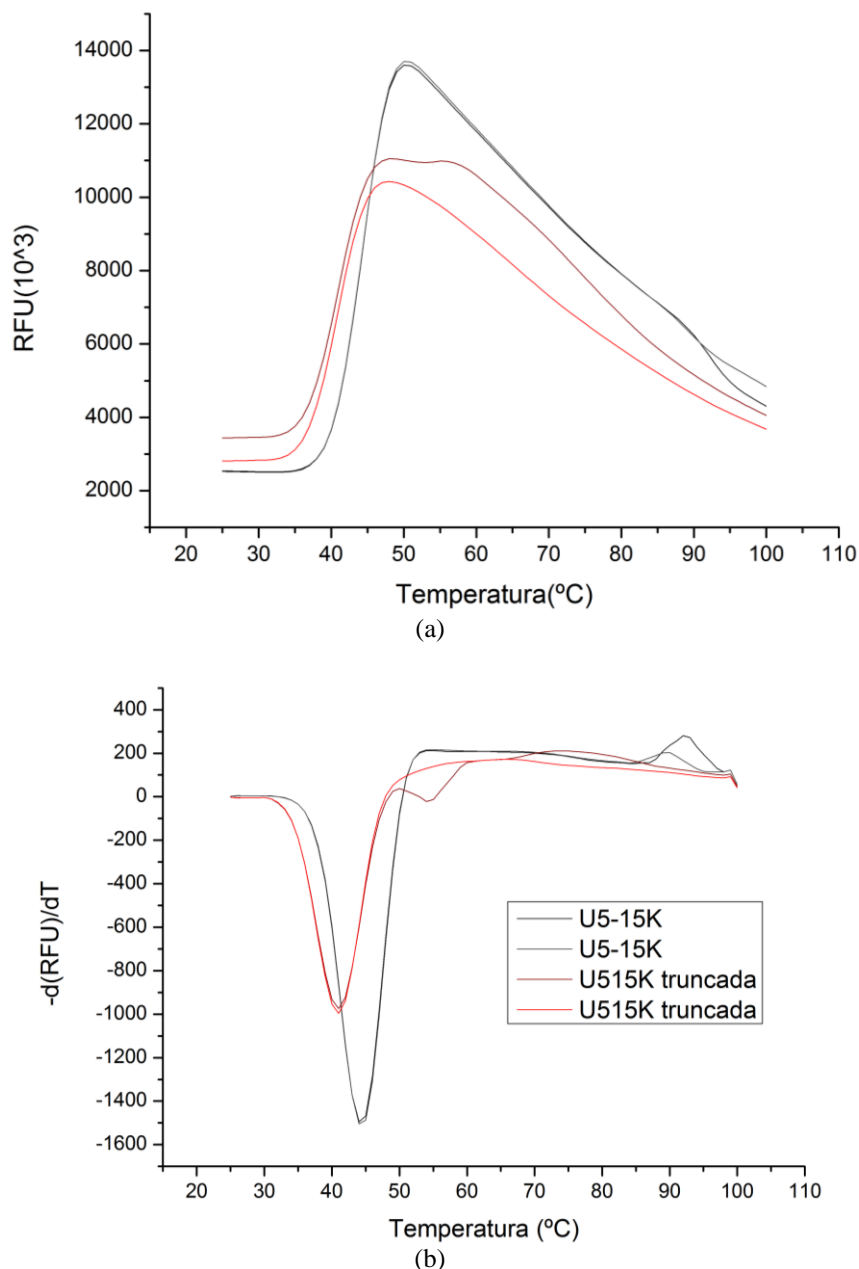


Figura 36 – a) Curva de fluorescência por temperatura das proteínas U5-15K (preta) e U5-15K (vermelha) e b) curva da sua derivada em relação à temperatura para a obtenção da temperatura de desenovelamento ( $T_m$ ).

Fonte: Elaborada pela autora.

Foram testadas diferentes concentrações da sonda fluorescente (Sypro Orange) e proteína, em duplicata, e a figura 36 apresenta as curvas obtidas com melhor sinal, na menor concentração de sonda e 4 $\mu$ M de proteína.

Na figura 36 (a), é possível observar um aumento de fluorescência por volta de 40°C, quando o desenovelamento da proteína acontece, seguido de uma queda, em uma curva com um formato muito semelhante ao da proteína selvagem, o que permite inferir que a U5-15K truncada possui uma estrutura terciária globular bem definida. A partir das curvas da figura 36

(b) foram calculadas as temperaturas de desenovelamento ( $T_m$ ) que resultou em  $45\pm 1^\circ\text{C}$  para a U5-15K nativa e  $41\pm 1^\circ\text{C}$  para a U5-15K sem alça, mostrando que, embora a proteína mutante tenha um enovelamento globular bem definido, a estabilidade térmica é menor sem a alça flexível. Apesar disso, a diferença de  $T_m$  é inferior a  $7^\circ\text{C}$  o que, segundo a literatura<sup>85</sup>, não representa grandes mudanças na probabilidade da proteína formar cristais, no tampão em que o experimento foi realizado.

#### 4.2.4 Teste de clivagem

Também foi avaliada a influência da alça flexível na atividade de autoclivagem. O teste seguiu o mesmo procedimento citado para as demais mutantes usadas nesse trabalho (seções 3.8.1 e 4.1.3).

Na figura 37 (a), pode-se ver o aparecimento de uma banda de massa molecular reduzida por volta de 24h, e sua inibição dessa clivagem com a adição de quimostatina, um dos inibidores da atividade da proteína selvagem.

Esse experimento demonstra que a atividade de autoclivagem da proteína foi mantida, bem como sua inibição por quimostatina, sendo mais um argumento a favor da consideração de que a estrutura terciária da U5-15K sem alça é muito semelhante a da U5-15K nativa.

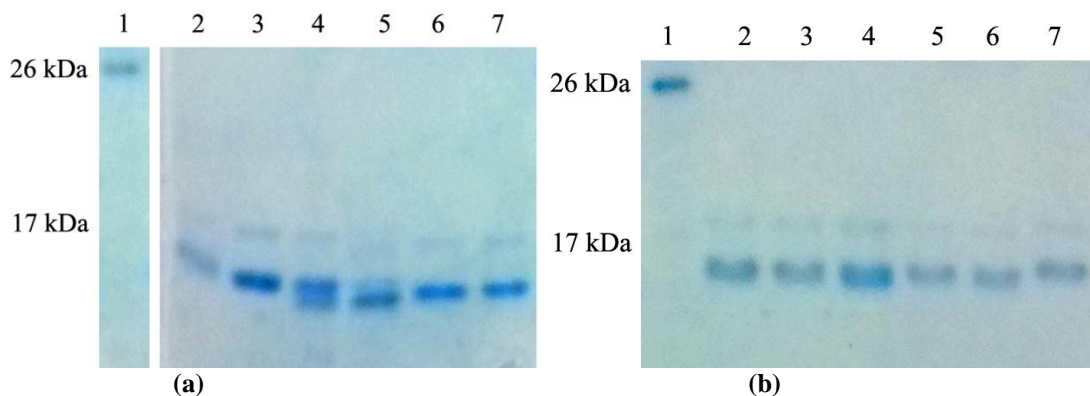


Figura 37 – Gel SDS-PAGE 20% do teste de clivagem da U5-15K sem a alça. A imagem da direita se refere ao teste controle e o da direita ao teste da U5-15K truncada na presença de  $6\ \mu\text{g mL}^{-1}$  de quimostatina. Em ambos géis a coluna 1 possui marcador de massa molecular, 2 alíquota da proteína após a purificação, 3 alíquota após 12h da purificação, 4 24h após a purificação, 5 36h após a purificação, 6 48h após a purificação, 7 60h após a purificação.

Fonte: Elaborada pela autora.



#### 4.2.5 Ensaio de cristalização

A partir dos dados obtidos nas seções anteriores (proteína mutante enovelada e com atividade de autoclivagem), prosseguiu-se com os ensaios de cristalização. Foram testados os kits comerciais Crystal Screen HT (Hampton Research), MORPHEUS HT, SG1 Screen™ e JCSG-plus HT (Molecular Dimensions) como descrito na seção 3.10.2. A concentração de proteína usada foi de  $3\text{mg.mL}^{-1}$ , uma vez que concentrações acima dessa resultavam na agregação e precipitação da proteína.

Apesar de se esperar que a deleção de uma região flexível fosse aumentar a probabilidade de formação de cristais, estes não foram observados, o que indica que a rede cristalina da U5-15K de *T. brucei* não segue exatamente a obtida nos cristais da U5-15K de humano.

Outra alternativa para a obtenção de cristais seria a realização de ensaios de cristalização com uma mutante cuja atividade de autoclivagem estivesse inativada, uma vez que, como foi dito no início desta seção, a mudança da população de proteína (clivada e não clivada) na solução pode estar prejudicando a formação dos cristais.

Partindo desse ponto, e tendo em mãos a U5-15K S126A, que como foi visto, não possui atividade de autoclivagem, foram realizados novos ensaios de cristalização com os kits comerciais Index HT (Hampton Research), MORPHEUS HT, SG1 Screen™, BCS Screen e JCSG-plus HT (Molecular Dimensions), como descrito na seção 3.10.2. Para esses ensaios foi usada uma concentração de aproximadamente  $4\text{mg.mL}^{-1}$ .

Depois de algumas semanas alguns cristais apareceram em algumas condições dos kits comerciais MORPHEUS HT, JCSG-plus HT e BCS Screen (figura 38), sendo que a solução de cristalização está descrita na tabela abaixo.

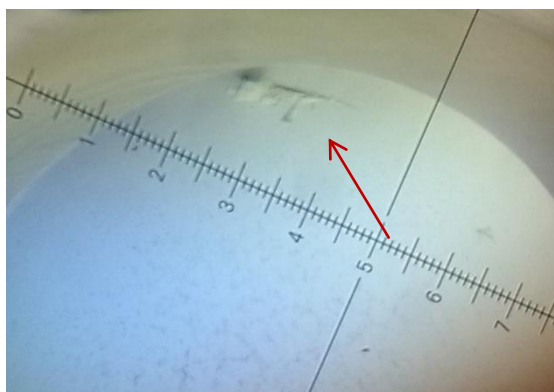


Figura 38 – Cristal da proteína U5-15K S126A na condição 1 da tabela 10, indicado pela seta vermelha.  
Fonte: Elaborada pela autora.



Tabela 10 – Soluções de cristalização dos kits comerciais utilizados, que resultaram em cristais promissores.

Número	Kit comercial	Solução de cristalização
1	MORPHEUS HT	0,09M NPS (0,3M Nitrato de Sódio, 0,3M Fosfato de sódio dibásico, 0,3M Sulfato de amônio) 0,1M Tampão 3 (0,1M Tris (base)), pH8,5 37,5% v/v mistura de precipitação 4 (25% v/v MPD; 25% PEG 1000; 25% w/v PEG 3350)
2	JCSG-plus HT	0,005M cloreto de cobalto (II) hexahidratado 0,005M cloreto de cádmio pentahidratado 0,005M cloreto de magnésio hexahidratado 0,005M Cloreto de níquel hexahidrato 0,1M HEPES pH7,5  12% w/v PEG 3350
3	BCS Screen	0,2M Acetato de amônio 0,1M acetato de sódio, pH 4,6 28% v/v PEG 5% Etileno glicol
4	BCS Screen	0,15M Nitrato de amônio 0,1M MÉS pH 6,0 20% PEG 5% v/v Etileno glicol

Fonte: Elaborada pela autora.

Testes adicionais estão em andamento para a obtenção de cristais maiores e mais simétricos.

### 4.3 Estudo do subcomplexo formado pelas proteínas U5-15K e U5-102K

#### 4.3.1 Purificação dos parceiros de interação

A purificação por afinidade em tandem (*tandem affinity purification*, TAP) é usada no isolamento de complexos proteicos para posterior identificação das proteínas pertencentes a eles. Esse sistema foi inicialmente desenvolvido para *S. cerevisiae* e tem por princípio a fusão de uma cauda (cauda PTP) no N- ou C- terminal da proteína de interesse.<sup>86</sup>

Em *T. brucei* a cauda PTP é formada por um conjunto de epítomos, sendo o primeiro deles derivado da proteína A de *Staphylococcus aureus*, com capacidade de interagir com Imunoglobulina G (IgG). O segundo epítopo é proveniente da Proteína C de humanos, que pode ser capturado por anticorpos direcionados para ele, sendo esta interação dependente da presença de íons  $Ca^{2+}$ . Os epítomos estão separados por um sítio de clivagem reconhecido por TEV protease (Figura 39).<sup>86</sup> A purificação da proteína de interesse e seus parceiros de interação é feita através de duas cromatografias de afinidade, como descrito na seção 3.11.2.

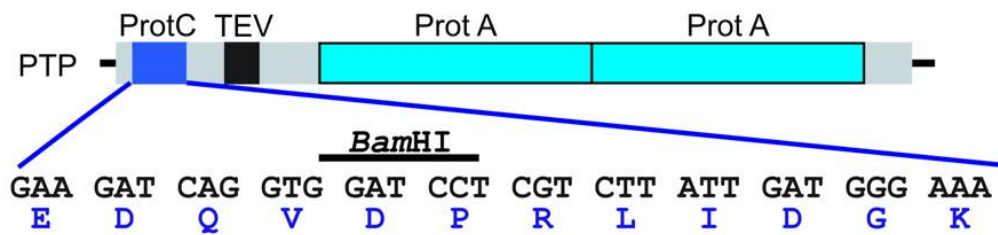


Figura 39 – Desenho esquemático da cauda PTP.  
 Fonte: SCHIMANSKI<sup>86</sup>

Em estudos prévios,<sup>79</sup> a cauda PTP foi fusionada ao C-terminal da proteína U5-15K de *T. brucei*, o que resultou na identificação da proteína U5-102K. No presente trabalho, a cauda PTP foi fusionada a proteína U5-102K usando o vetor pC-PTP-NEO, com o objetivo de observar se outras proteínas que participam da partícula U5 snRNP, da partícula U4.U6/U5 tri snRNP e complexo B poderiam ser purificadas.



Figura 40 – Posição da Prp6 (em vermelho) de levedura no U4/U6.U5 tri snRNP, com o N-terminal cruzando o complexo e interagindo com as proteínas Prp8, Prp4, Dib1 e Prp31.  
 Fonte: NGUYEN<sup>67</sup>

Essa proteína foi escolhida por ser maior do que a U5-15K (120kDa em contraste com uma proteína de 17kDa) e por se supor que interage com um número maior de proteínas e complexos, como foi visto nos complexos de humano e levedura.

Na figura 40, é possível observar como o N-terminal da Prp6 de levedura permeia a região a esquerda do tri snRNP U4/U6. U5, e o C-terminal interagem com a região a direita. Além disso, a U5-15K (DibI) deixa o *spliceosomo* durante a ativação do complexo B, o que não ocorre com a U5-102K (Prp6).<sup>67-68</sup>

Para isso, os oligonucleotídeos foram desenhados para a amplificação de uma porção de 617 pares de bases do C-terminal da U5-102K. Também foi desenhado um oligonucleotídeo denominado 102k 5'out (tabela 2), que, como o próprio nome indica, está posicionada fora dessa região de 617pb. Esse oligonucleotídeo foi usado para se certificar de que a cauda PTP foi corretamente inserida no genoma do *T. brucei* após a transfecção (figura 41).

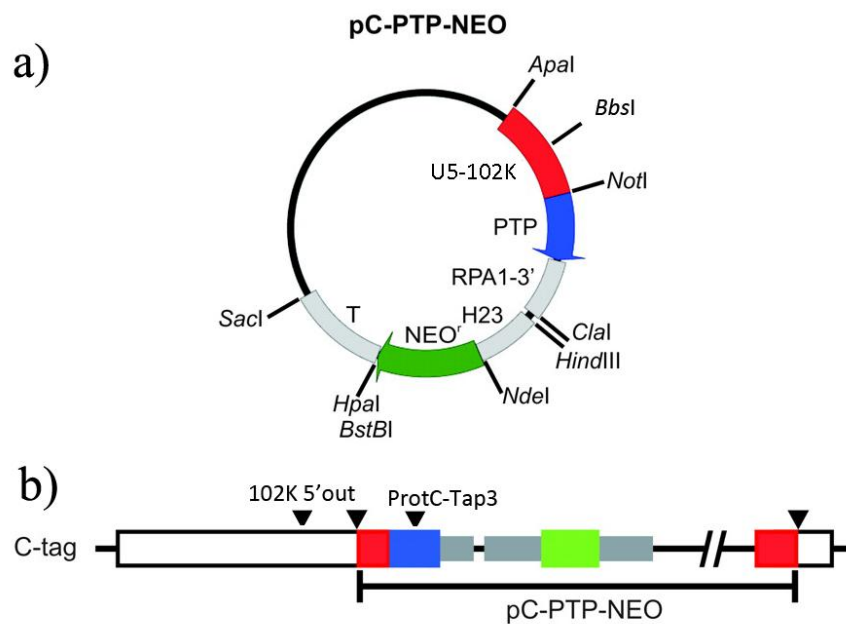


Figura 41 – a) Mapa do vetor pC-PTP-NEO contendo o C terminal da U5-102K em vermelho e b) inserção da cauda PTP no DNA genômico do *T. brucei*.

Fonte: Adaptada de SCHIMANSKI<sup>86</sup>

Essa porção C-terminal da U5-102K foi clonada em pC-PTP-NEO e transformadas em *E. coli* DH10b, como descrito nas seções 3.3.2 e 3.5.2. O DNA plasmidial foi extraído e linearizado com a enzima de restrição *BbsI*.

O vetor linearizado foi purificado e transfectado em TbPCF 427 como descrito na seção 3.6.2. A confirmação foi feita por meio de *Western Blot* (seção 3.7) com anticorpo

primário contra proteína A e uma PCR usando os oligonucleotídeos ProtC-Tap3 (complementar a proteína C da cauda PTP) e 102K 5'out (tabela 2). Esses resultados podem ser vistos na figura 42(a), onde a seta vermelha indica a banda referente à expressão da U5-102K fusionada com a cauda PTP, e na figura 42(c), onde a seta vermelha indica a banda de amplificação do C-terminal da U5-102K entre a posição onde o oligonucleotídeo 102K 5'out foi desenhado e a proteína C da cauda PTP.

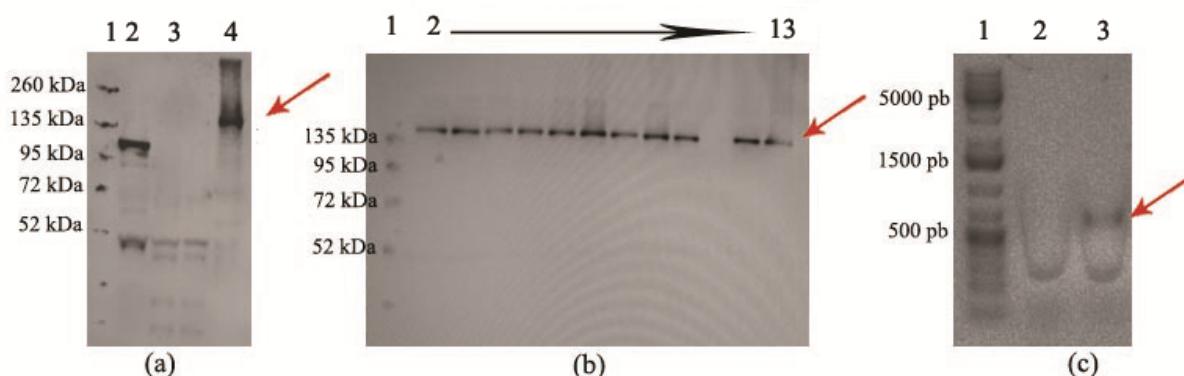


Figura 42 – a) Confirmação por *western blot* da expressão da U5-102K com a cauda PTP, onde 1 é o marcador de massa molecular, 2 controle positivo do *western blot*, 3 forma procíclica do *T. brucei* 427 não transfectada e 4 3 forma procíclica do *T. brucei* 427 expressando a U5-102K b) *Western blot* da seleção das colônias positivas, onde 1 é o marcador de massa molecular e 2 a 13 são as 12 colônias testadas. c) Eletroforese em gel de agarose da inserção da cauda PTP no gene da U5-102K, onde 1 é o marcador, 2 3 forma procíclica do *T. brucei* 427 não transfectada e 3 3 forma procíclica do *T. brucei* 427 expressando a U5-102K. As bandas relativas à proteína U5-102K (a e b) e ao C-terminal do gene da U5-102K (c) estão indicadas pelas setas vermelhas.

Fonte: Elaborada pela autora.

Confirmada a transfecção, foi feita a seleção dos clones como descrito na seção 3.7.2. Da placa de 96 poços, foram escolhidas 12 colônias positivas, inoculadas em 10ml de meio SDM-79. Essas células foram centrifugadas, suspensas com tampão de amostra, fervidas por 5 minutos e aplicadas em SDS-PAGE 10%. Foi então feito o *western blot*, como descrito na seção 3.7.

Na figura 42(b), a seta vermelha indica a banda referente a expressão da U5-102K nas 12 colônias escolhidas, sendo que 11 dessa colônias eram positivas. Desses 11 clones, um deles foi escolhido para a expressão em 4L de SDM-79, como descrito na seção 3.5.1.

Prosseguiu-se com a lise e purificação como descritos nas seções 3.11.1 e 3.11.2 e alíquotas de cada etapa de purificação foram aplicadas em SDS-PAGE. As bandas das proteínas presentes nessas alíquotas podem ser vistas na figura 43. As bandas referentes às proteínas eluídas da coluna com anti-Proteína C foram recortadas e prosseguiu-se como descrito na seção 3.11.4.

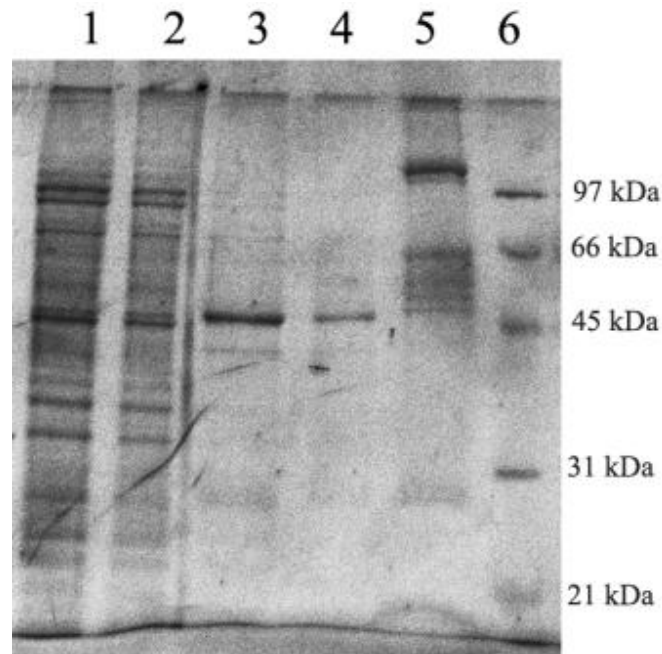


Figura 43 – SDS-PAGE 12% da purificação dos parceiros de interação da U5-102K, onde as coluna 1 e 2 possuem uma alíquota das proteínas que não se ligaram à coluna, 3 uma alíquota das proteínas eluídas da coluna contendo IgG (primeira etapa da purificação), 4 alíquota das proteínas que não se ligaram a segunda coluna (anti proteína C), 5 alíquota das proteínas eluídas da coluna contendo anti-proteína C (segunda etapa de purificação) e 6 marcador de massa molecular.

Figura: Elaborada pela autora.

A identificação das proteínas copurificadas foi feita por espectrometria de massas no Instituto Butantã, em parceria com a Dra Júlia Cunha, e as identificações estão representadas na tabela abaixo.

Tabela 11 – Proteínas identificadas por espectrometria de massas.

Código	TritypDB	Score	Cobertura	Peptídeos	Resíduos	Massa(kDa)
Tb427_110175400	U5-102K	342	13%	12	1009	111,8
Tb427.01.2330	Beta Tubulina	84	9%	3	442	50,3
Tb427_110023800	TCP-1 subunidade beta	68	4%	2	529	58,4
Tb427.10.2100	EF-1-alfa	41	1%	1	449	49,5

Fonte: Elaborada pela autora.

Das proteínas identificadas, foram escolhidas apenas as que apresentavam score superior a 38. O score das proteínas identificadas é derivado da pontuação dos íons, essa pontuação por sua vez é  $-10\log(P)$ , onde P é a probabilidade de a correspondência observada seja um evento aleatório.

As proteínas com maior score, e massa molecular equivalentes às das bandas no SDS-PAGE 12% estão indicadas na figura a seguir.

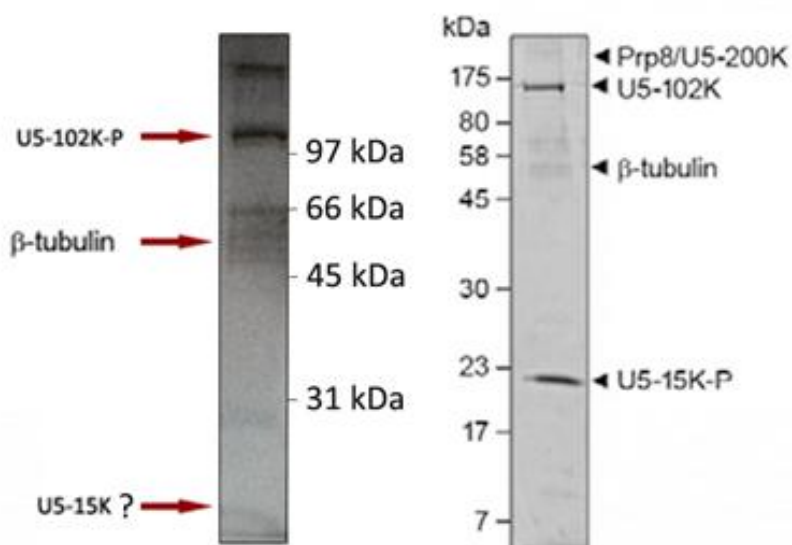


Figura 44 – A coluna da esquerda se refere à identificação das proteínas purificadas pela técnica de PTP-tag na U5-102K e a coluna da direita se refere às proteínas identificadas na purificação pela técnica de PTP-tag no C-terminal da U5-15K<sup>79</sup>.

Fonte: Elaborada pela autora.

A espectrometria de massa não identificou nenhuma proteína diferente das identificadas com a cauda no C-terminal da U5-15K e não foram identificadas proteínas com massas moleculares abaixo de 31 kDa, sendo que na figura 44 é possível observar duas bandas tênues nessa região. Novas purificações seriam necessárias, partindo de volumes maiores de cultura, visando a correta identificação das bandas presentes no gel.

Alternativamente, a inserção da cauda PTP no N-terminal da U5-102K pode levar a identificação de outras proteínas, pois se observarmos a posição do C-terminal da Prp6 no tri-snRNP de levedura (figura 45) a proteína parece se dobrar seu redor do C-terminal, de forma que a cauda PTP adicionada nessa região pode ficar escondida pelo restante de estrutura, ou mesmo atrapalhar o enovelamento dessa região, impedindo que as proteínas com a cauda PTP participem dos complexos. Essas duas situações dificultariam a purificação do complexo.



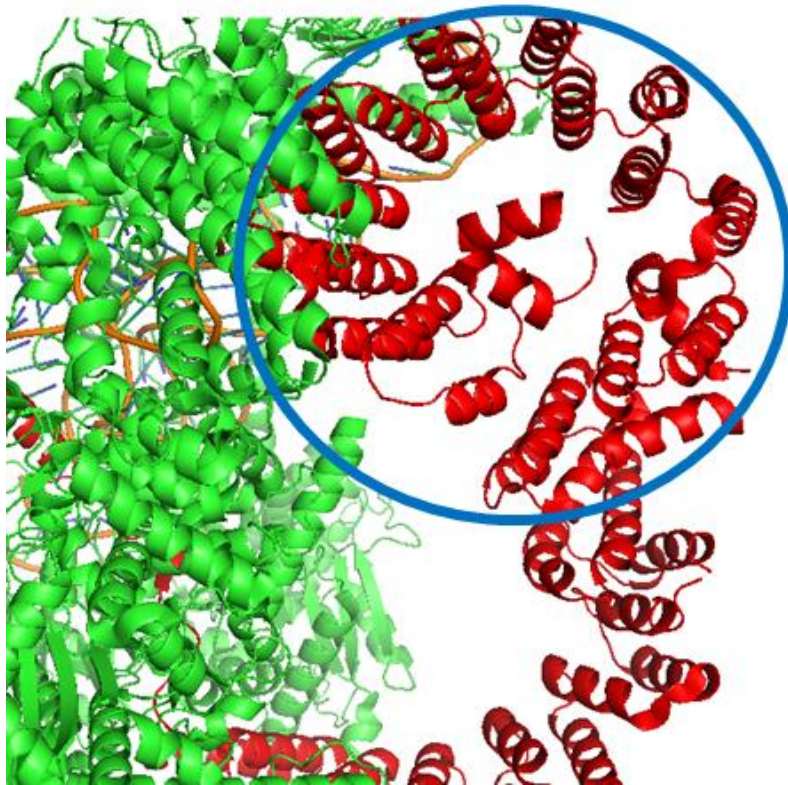


Figura 45 – Estrutura do U4/U6.U5 tri snRNP de levedura com o C-terminal da Prp6 em evidência (círculo azul).

Fonte: Adaptada NGUYEN <sup>67</sup>

#### 4.3.2 Obtenção da proteína U5-102K recombinante

Para estudar a interação entre U5-15K e U5-102K *in vitro*, foi necessário obter a U5-102K recombinante. A abordagem inicial foi utilizar a construção pET28\_102K cedida pelo Dr. Marco Tulio A. da Silva. Esse vetor foi transformado em *E. coli* BL21(DE3) e a proteína foi expressa como descrito nas seções 3.4.2 e 3.6.1.

O SDS-PAGE do teste de expressão é mostrado na figura 46, nela é possível observar o aparecimento de banda equivalente a massa molecular superior a 97kDa, indicada pela seta vermelha. Essa banda aparece apenas após a indução com IPTG e é referente a expressão da U5-102K.

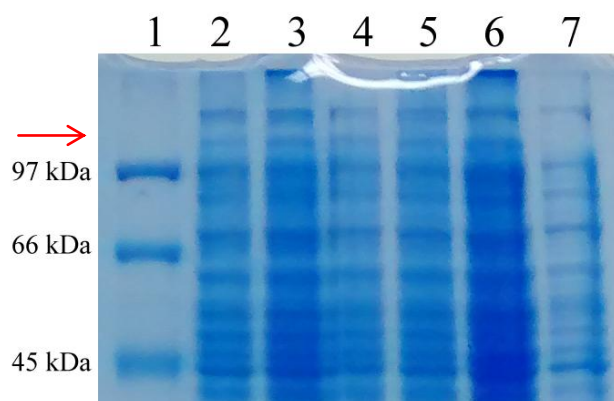


Figura 46 – SDS-PAGE 10% do teste de expressão da U5-102K, clonada em pET28a. A coluna 1 representa o marcador de massa molecular, a coluna 2 uma alíquota 1h após a indução com 0,5mM de IPTG, 3 2h após a indução, 4 3h após a indução, 5 4h após a indução, 6 16h após a indução e 7 uma alíquota antes da indução. A banda de expressão da U5-102K está indicada com a seta vermelha.

Fonte: Elaborada pela autora.

Em seguida foi feita a purificação por cromatografia de afinidade, como descrito na seção 3.6.3. Foram separadas alíquotas das proteínas solúveis totais, das proteínas que passaram pela coluna sem ficarem retidas na resina e alíquotas do pico de eluição do gradiente de imidazol. Essas alíquotas foram aplicadas em SDS-PAGE 10% e o resultado é mostrado na figura 47.

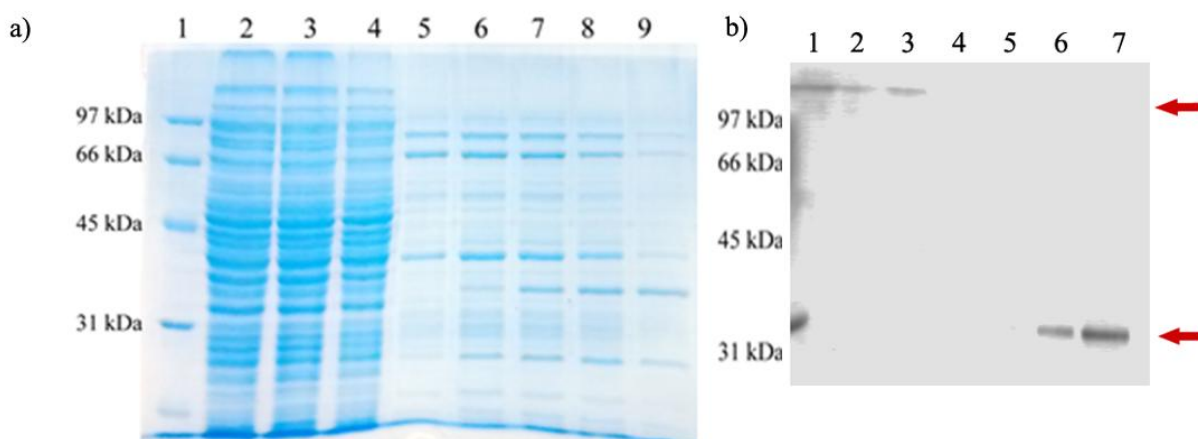


Figura 47 – a) Gel SDS- PAGE da purificação da U5-102K, onde 1 representa o marcador de massa molecular, 2 as proteínas solúveis totais na amostra, 3 as proteínas que não interagiram com a resina, 4 as proteínas que foram eluídas com tampão sem imidazol e os poços 5, 6, 7, 8 e 9 representam as frações de coleta do ÄKTA correspondentes ao pico de eluição no gradiente de concentração de imidazol. b) Western blot da purificação da U5-102K onde 1 representa a porção de proteínas que não interagiram com a coluna, 2 as proteínas eluídas em tampão sem imidazol e os poços 3, 4, 5, 6 e 7 são as frações de coleta do ÄKTA correspondentes ao pico de eluição do gradiente de concentração de imidazol.

Fonte: Elaborada pela autora.

A quantidade de bandas de tamanhos distintos que pode ser vista no gel da purificação, principalmente nos poços de 5 a 9 (figura 47 (a)) indica que muitas proteínas



contaminantes têm a mesma afinidade pela coluna que a U5-102K. O *western blot* (com anticorpo conta a cauda de histidina), feito a partir das mesmas amostras do gel da figura 47 (a), sugere uma degradação da proteína durante o processo de purificação, que pode ser visto pelo aparecimento de uma banda de aproximadamente 35kDa na figura 47 (b), demonstrando instabilidade da U5-102K em solução.

Como estes dois membros spliceossomais foram co-purificados a partir de células de *T. brucei*, se levantou a hipótese que a co-expressão destas proteínas poderia estabilizar U5-102K. Para tal foi empregando o vetor pET-Duet, cuja principal característica é a presença de dois cassetes de clonagem, como mostrado na figura 48. Essas clonagens foram feitas como descrito na seção 3.2.2.

Para separar a U5-102k do vetor de propagação pJET1-2, foi necessária a digestão do vetor com a enzima de restrição Eam1105I, uma vez que, por terem aproximadamente o mesmo tamanho, não é possível separar a banda do gene da banda do vetor por eletroforese em gel de agarose 1%.

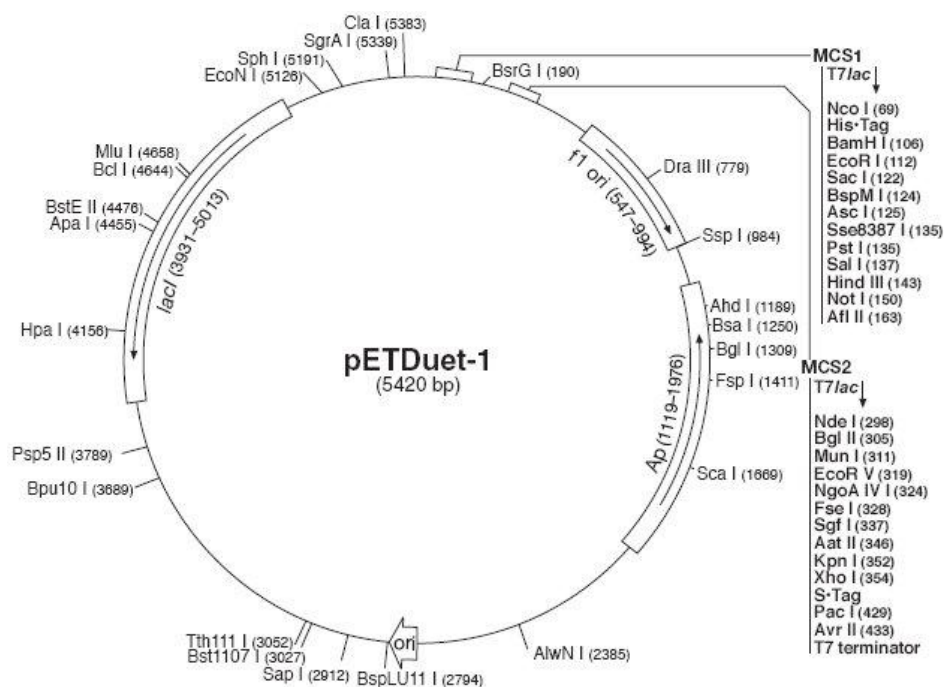


Figura 48 – Mapa de vetor pETDuet-1  
Fonte: MANUAL pETDuet<sup>87</sup>

O vetor de expressão pETDuet-1 é um vetor que permite a expressão de duas proteínas simultaneamente, sendo uma fusionada a cauda de histidina ao N-terminal e a outra com S-Tag, peptídeo derivado da ribonuclease A (essa cauda não será usada nesse trabalho).<sup>87</sup> Uma

vez que a co-purificação das proteínas seria realizada por cromatografia de afinidade por íons  $\text{Co}^{2+}$ , decidiu-se por adicionar a cauda da histidia à proteína com menor taxa de expressão. Essa medida visa diminuir a quantidade de proteína purificada sem a parceira de interação.

Baseado nos SDS-PAGE da expressão de U5-15K clonada em pET28a, obtida em trabalhos anteriores<sup>79</sup>, e da U5-102K (figura 46), e nos tamanhos das proteínas, 120kDa para a U5-102K e 17kDa para a U5-15k, decidiu-se pela clonagem da U5-102K fusionada a cauda de histidinas.

Com as clonagens confirmadas por gel de agarose 1% e sequenciamento dos plasmídeos, a construção pETDUET\_102K\_15K foi transformada inicialmente em *E. coli* BL21(DE)3, como descrito na seção 3.5.2 e foram feitos testes de expressão com diversas colônias.

Como se trata da co-expressão de duas proteínas de tamanhos muito distintos, foi necessário aplicar as amostras do teste de expressão em dois SDS-PAGE, um 10% para a observação da banda referente à U5-102K e um 15% para a observação da U5-15K (figura 49).

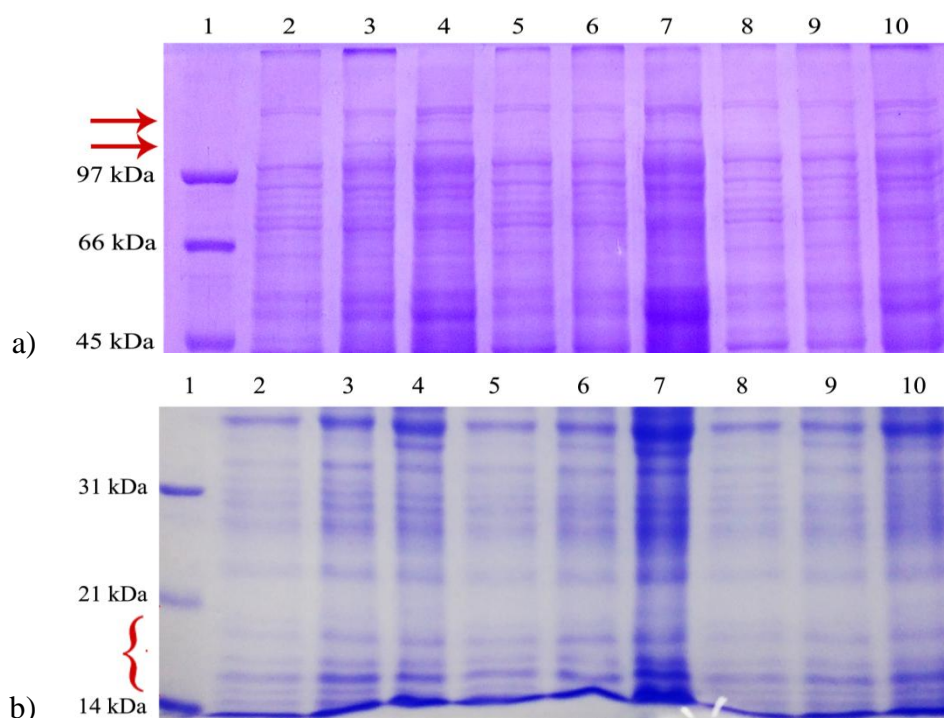


Figura 49 – SDS-PAGE do teste de co-expressão das proteínas U5-15K e U5-102K em pET DUET a) 10% para a visualização da U5-102K, onde as setas vermelhas apontam as bandas que podem corresponder à expressão da proteína. b) 15% para a visualização da U5-15K, onde a chave vermelha indica a região do gel onde a banda de expressão apareceria. Em ambos géis a coluna 1 contém marcador de massa molecular, 2 uma alíquota da colônia 1 antes da indução com IPTG, 2 2h após a indução e 3 16h após a indução, 5 alíquota da colônia 2 antes da indução, 6 2h após a indução e 7 16h após a indução, 8 alíquota da colônia 3 antes da indução com IPTG, 9 2h após a indução e 10 16h após a indução.

Fonte: Elaborada pela autora.

Na figura 49 (a), podemos ver o aparecimento de duas bandas acima de 97 kDa. Como o marcador da massa molecular usado não possui proteínas com massa superior a 97 kDa, não é possível afirmar qual das duas bandas realmente corresponderia à U5-102K. Já na figura 49 (b), não observamos o aparecimento de nenhuma banda de super expressão.

Foram feitos novos testes de expressão com outras cepas (seção 3.4.2) como *E. coli* Rosetta Gami B(DE3) e *E. coli* BL21(DE3) *codon plus*. Também foi feito o *western blot* dessas amostras, com anticorpo contra a cauda de histidina, para a identificação da U5-102K, mas não foram observadas bandas de expressão nesses experimentos. Dados os resultados negativos, foi definida uma nova estratégia, que será descrita na seção seguinte.

### 4.3.3 Obtenção do N-terminal da proteína U5-102K

A resolução do tri snRNP de levedura e do complexo B do spliceossomo de humanos permitiu visualizar os pontos de interação entre as proteínas homólogas à U-515K e U5-102K (figura 50).

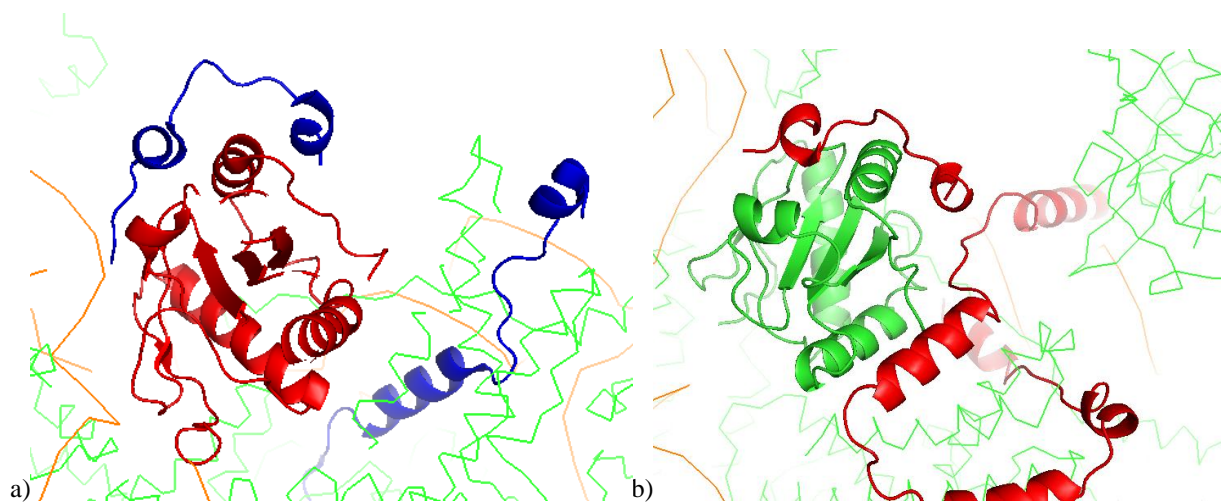


Figura 50 – a) Interação entre as proteínas U5-15K (vermelha) e a porção N-terminal da proteína U5-102K (azul) no tri snRNP de levedura. b) Interação entre as proteínas U5-15K (verde) e a porção N-terminal da proteína U5-102K (vermelha) no complexo B do spliceossomo humano.

Fonte: Adaptada de NGUYEN<sup>68</sup>; BERTRAM<sup>88</sup>

Como pode ser visto na figura 50 (a), a interação entre as duas proteínas no tri snRNP de levedura ocorre apenas em uma pequena região do N-terminal da Prp6 formada por duas hélices alfa compreendidas entre os resíduos 4 e 26. Já no complexo B humano, parece haver uma aproximação da hélice alfa entre os resíduos 115 e 135, além da região dos resíduos de 6 a 26 do N-terminal da U5-102K.<sup>68,88</sup>

Com essas informações em mãos, foi feito o alinhamento dessas sequências com a sequência da U5-102K de *T. brucei* (figura 51) e a predição de suas estruturas secundária usando o servidor PSIPRED.

```

Tbrucei      MSAAAPVLSTKVDKYEQVLAKYDWRSPFPPRGYQYGIGRGAKAFITSAELSATTGVSAPL
Scerevisiae -----MERPSFLDQEPAGYVFGIGRGATGFSTKEKQVSNDDKG--
Hsapiens    -----MNKKKPFPGMPAPLGYVFGLRGATGFTTRSDIGPARDANDPV
              :  :  .  *  **  *:*:*:*..*  *  .  :  .  .

Tbrucei      LA-----LGGEQSAM
Scerevisiae --RRIPKR-----YRENLNHLQSQPKDDEDEAANV
Hsapiens    DDRHAPPGKRTVGDQMKKNQAADDDDELDNDTNYDEFNGYAGSLFSSGPYEKDEEADAI
              .  *  :

Tbrucei      LDALDR-LEERRHKRRREDVKGGATSREGKKPKLLLSMED-----IATVG
Scerevisiae  FKTLELKLAKKKKRRAN---EKDDDNSVDSSNVKRQFADLKESLAAVT ESEWM DIPDAT
Hsapiens    YAALDKFMDERRKERREQREKEEIEKYRMRERPKIQQFSDLKRKLAEVTEEEWLSIPEVG
              :*  :  :  :  :  :  *  .  .  .  :  :  *  *

Tbrucei      SGASNQSLVKKRKARTEDMDVVDGYAFDDEAVVTASDLLRGRA-----
Scerevisiae DFTR---RNKRNRIOEQQLNR-KTYAAPDSL---IPG-----
Hsapiens    DARN---KRQENPRYEKLT---PVPDSE---FAKHLQTGENHTSV DPRQTQFGGLNTP
              :  :  *  :  :  *

```

Figura 51 – Porção N-terminal do alinhamento das sequências das proteínas U5-102K de *T. brucei*, U5-102K de *H. sapiens* e Prp6 de *S. cerevisiae* feito no ClustalW. As regiões de hélice alfa estão grifadas em rosa. As regiões de interação com a U5-15K e Dim1 em *H. sapiens* e *S. cerevisiae* respectivamente, então marcadas em vermelho.

Fonte: Elaborada pela autora.

Do alinhamento, podemos observar resíduos conservados na região N-terminal marcados em vermelho, apesar de a predição de estrutura secundária não conservar as hélices alfa na proteína de *T. brucei*. Dessa forma, escolhemos a porção entre a prolina 6 e fenilalanina 140 para clonagem em pET SUMO e expressão em *E. coli* (figura 52), como citado nas seções 3.2.2 e 3.6.1. Essa região resultaria na expressão de uma proteína um pouco menor que a U5-15K, de aproximadamente 15kDa, mas não tão pequena que sua observação em SDS-PAGE se tornasse particularmente complicada.

```

MSAAAPVLSTKVDKYEQVLAKYDWRSPFPPRGYQYGIGRGAKAFITSAELSATTGVSAPLLALGGEQSAML
LDALDRLEERRHKRRREDVKGGATSREGKKPKLLLSMEDIATVGGSGASNQSLVKKRKARTEDMDVVDGYAFDDE
TAVVTASDLLRGRALAAANQTLENILNMGSSSEQTTWITHSRALREMGMTKKAQQTIVEGSRITGSKG
PLIWKEQLQLHNDPAAQRQLLEQAVAAACKTCEELWFLLLQYQPPHEQLHLWLQKAVMACPASESLWLR
ILGYVTAPRDQRKIIRKALEVTPTLPSLWAMLARLEDYERGAIFNAAAAEHPSMKIIIEGAKFEFHLR
GQVGVGSEKEGGNEQLHNAKIHSVLRQAAQRFLMDEEKSRTEWLDAARVAVSERYILTSAYMILLHFV
CDTGAATLTIPVTWLEDLAALVPDMWRAHDVLCATAWATMLIGRDAAVKGTNEAVTEKDVQIIGFAL
RSAPFTSLETAVSAVTRYISSTTVSYTSRVEDGMKKEGIKEEEEEELGRPVATIQAAAPTCTSAATA
SQCLPLPLVSVLLYTLFRAPAFNTIDVMLSVAKVFYDHHFYDGAHRVLMLAITRYPVNALLFTAACK
TKMAMGCRKGAEEILLKGT AIDPSSDIAWVKLAVHRRSNREDIVPLLDEALVHFPTSERLWLMRLEAEGMKA
QQLLEDAMARGSSAPTINGLRRVYQKALSSEHCRLSPTIWCYAAYKLESNLFSDAGAARALLLEGLV
VCEKGCQHLHHNGSTKRAEVRATFGLARCHVELRHAGSETALETVKEVLQQLPKVDGCFVSVGELV
ALSIDLENPATRGRAAAQAVQHWVRDPLVFSSVAKLYHAAGRHEKALEQAMKAVQMSKGRCGDAVAL
WLKLSMFTYRKIVMEKMGLTVKDESNEEETALQNGVLLAWLWKQLANGSSGGTSGSEAMEDMKGSTES
ATHSEPPKPNSGPLWIYVSKSRDPSNVTLLGFRDSVEVMLTQVEMIKL

```

Figura 52 – Sequência de aminoácidos da proteína U5-102K de *T. brucei*, a região em vermelho indica a sequência do N-terminal referida nesse trabalho como U5-102K N-term.

Fonte: Elaborada pela autora.

O resultado do teste de expressão pode ser visto na figura 53, é possível observar o surgimento de uma banda entre 31 e 21 kDa após a indução com IPTG, para as 4 colônias testadas. Essas bandas são referentes a expressão do N-terminal da U5-102K e estão apontadas pelas setas vermelhas.

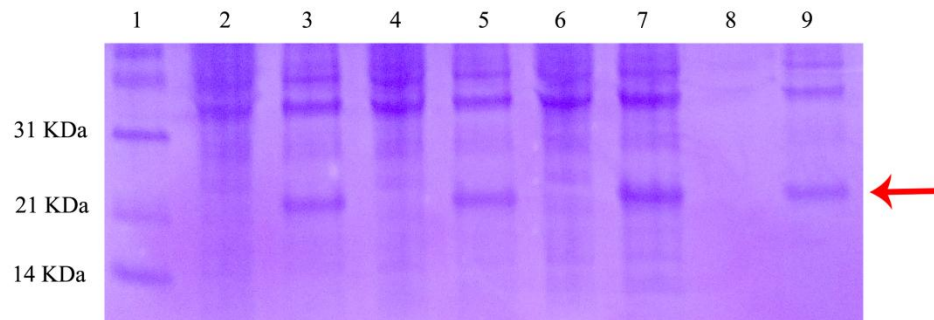


Figura 53- Gel SDS-PAGE 15 % do teste de expressão em *E. coli* BL21(DE3) do N-terminal da proteína U5-102k clonado em pET SUMO, onde a coluna 1 contém marcador de massa molecular, 2 alíquota da colônia 1 antes da indução com IPTG, 3 alíquota da colônia 1 16h após a indução, 4 alíquota da colônia 2 antes da indução com IPTG, 5 alíquota da colônia 2 16h após a indução, 6 alíquota da colônia 3 antes da indução com IPTG, 7 alíquota da colônia 3 16h após a indução, 8 alíquota da colônia 4 antes da indução com IPTG, 9 alíquota da colônia 4 16h após a indução e a seta vermelha indica a banda de expressão da proteína.

Fonte: Elaborada pela autora.

Em seguida foi testada a solubilidade da proteína, como descrito na (seção 3.6.2). O resultado pode ser visto na figura 54 (a), onde o gel da esquerda possui o extrato das proteínas insolúveis (coluna 1) e solúveis (coluna 2), foi feito o *western blot* dessas amostras (gel da direita, sendo que uma banda é visível na fração solúvel (apontada pela seta vermelha).



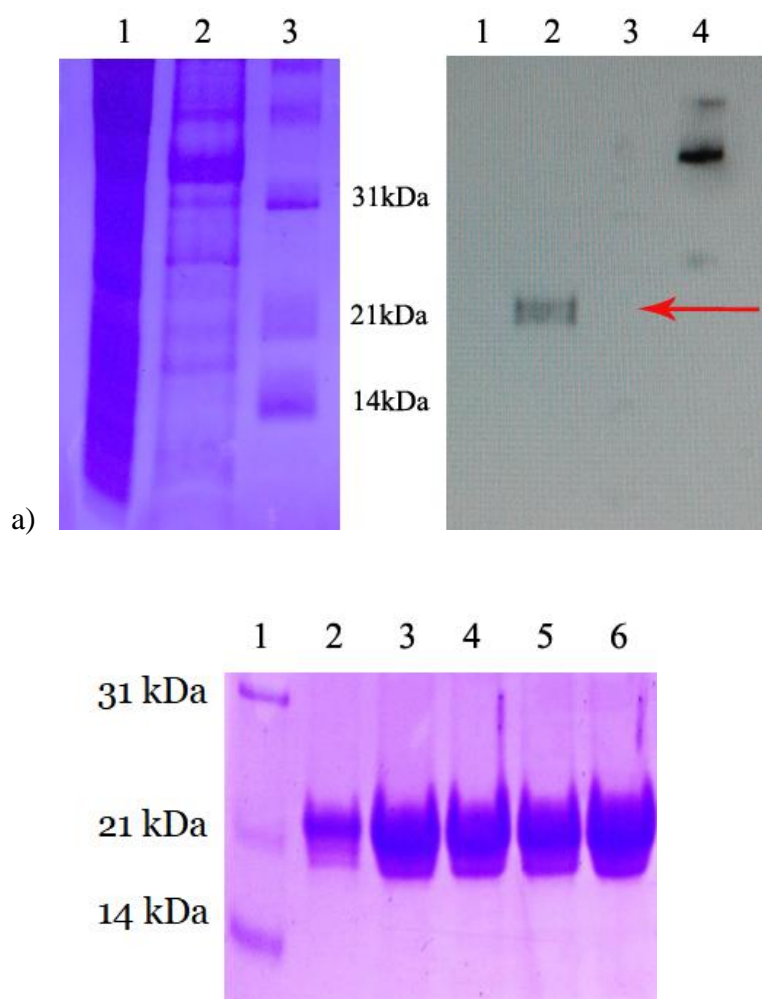


Figura 54 – a) SDS-PAGE 15% do teste de solubilidade do N-terminal da U5-102K e *western blot* das mesmas amostras, onde 1 é a alíquota das proteínas insolúveis, 2 a alíquota das proteínas solúveis, 3 o marcador de massa molecular e 4 o controle positivo do *western blot* feito com anticorpo contra a cauda de histidinas, b) SDS-PAGE 15% das frações do ÄKTA correspondentes ao pico da purificação por afinidade do N-terminal da U5-102K, onde 1 é o marcador de massa molecular e 2 a 6 são alíquotas das frações do pico de purificação.

Fonte: Elaborada pela autora.

Com a proteína solúvel, foi feita a purificação por cromatografia de afinidade, como descrito na seção 3.6.3. A figura 54 (b) mostra as frações equivalentes ao pico de eluição da proteína. A U5-102K N-term foi então concentrada por ultra-filtração em membrana Milipore sob rotação de  $2500 \times g$  a  $4^{\circ}C$  até aproximadamente  $1mgml^{-1}$  e usada nos testes de interação com a U5-15K.

#### 4.3.4 Ensaios de interação

Para os ensaios de interação, as duas proteínas foram concentradas separadamente até aproximadamente  $1mgml^{-1}$ . Uma alíquota de cada foi retirada para o cálculo do raio

hidrodinâmico a partir da técnica de DLS (seção 3.9.2). Em seguida, alíquotas das duas proteínas foram transferidas para um mesmo microtubo na proporção de 1:1 para interação e o raio hidrodinâmico foi novamente calculado. A comparação das três curvas está representada na figura a baixo.

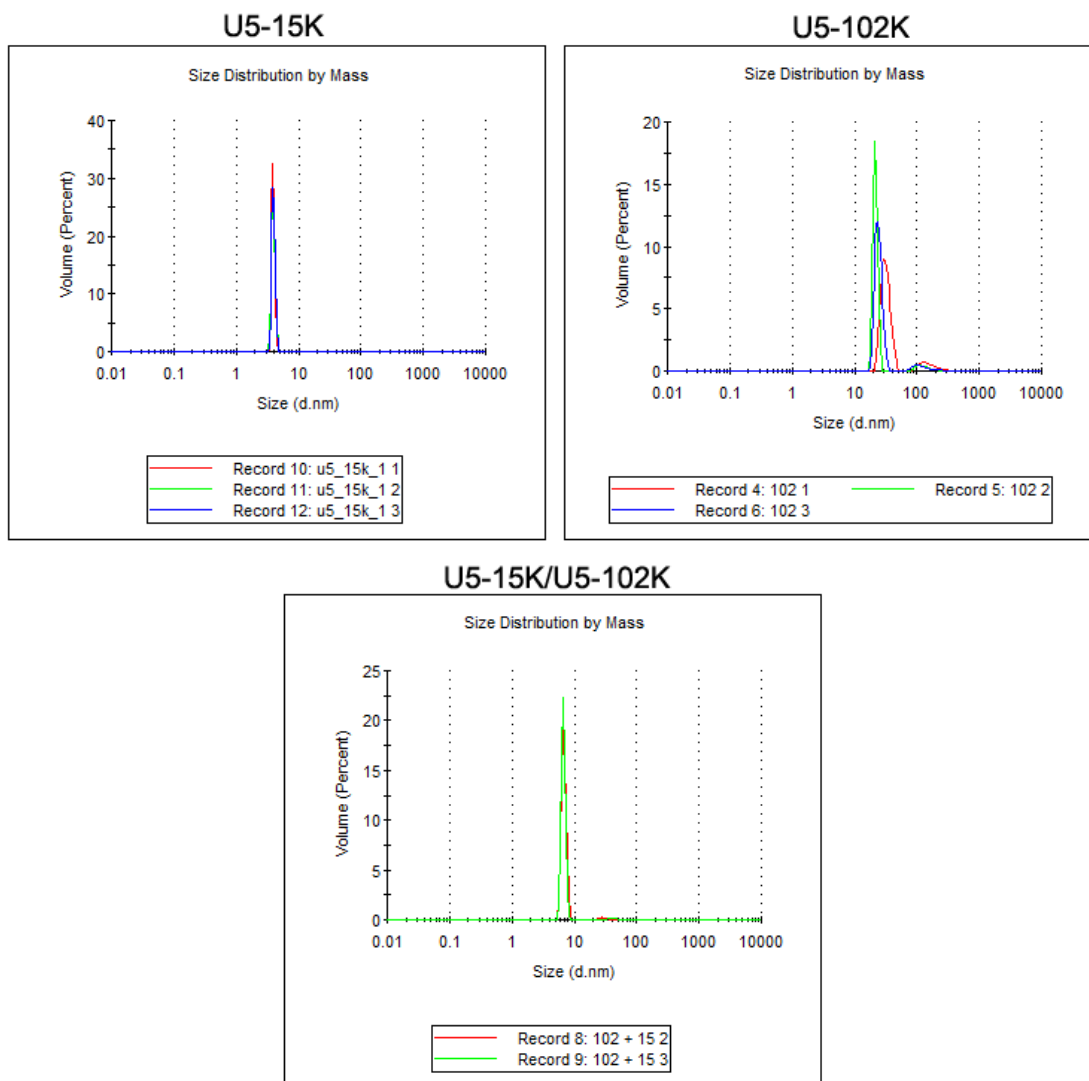


Figura 55 – Distribuição do diâmetro em nanômetros da U5-15K, da U5-102K e da mistura de 1:1 das duas proteínas, como assinalado no topo de cada quadro.

Fonte: Elaborada pela autora

O raio hidrodinâmico ( $R_H$ ) calculado para a U5-15K foi de  $19 \pm 2$  Å e massa molecular de  $16 \pm 3$  kDa,  $127 \pm 4$  Å para o N-terminal da U5-102K e  $33 \pm 4$  Å para a mistura das duas proteínas e massa molecular de  $50 \pm 20$  kDa. A massa e o  $R_H$  calculados para a U5-15K são bem próximos dos valores teóricos e calculados anteriormente<sup>64</sup>, já o valores para o N-terminal da U5-102K sugerem que a proteína esteja agregada.

Além disso, a técnica usada para o cálculo do  $R_H$  aproxima as partículas a esferas, e se levarmos em consideração que o N-terminal da U5-102K humana e Prp6 de levedura é pouco estruturado e flexível, é de se esperar que o cálculo do  $R_H$  para essa proteína seja menos preciso.

Observando os valores calculados para a mistura da U5-15K e N-terminal da U5-102K, pode-se notar que o  $R_H$  dobrou, enquanto a massa molecular teórica do heterodímero formado pelas duas proteínas (aproximadamente 33 kDa) está dentro do intervalo da massa molecular calculada pela técnica de DLS.

Em paralelo, foi feito um gel de poliacrilamida não desnaturante, para a separação dos diferentes complexos proteicos sendo formados. Foram aplicadas amostras da U5-15K, U5-102K e da mistura, na proporção de 1:1, das duas proteínas. O resultado pode ser observado na figura abaixo.

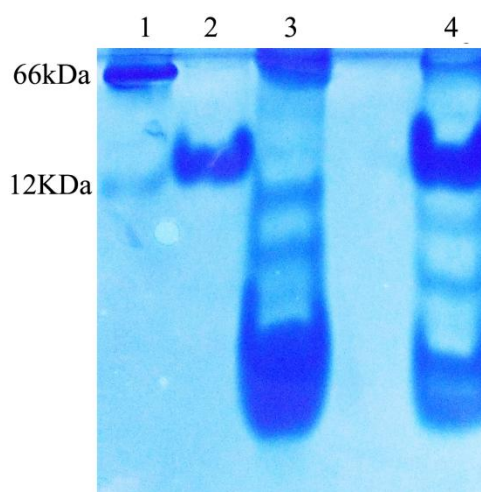


Figura 56 – Gel nativo com 20% de poliacrilamida do teste de interação entre U5-15K e o N-terminal da U5-102K, sendo que a coluna 1 contém o marcador de massa molecular, 2 uma alíquota da U5-15K purificada, 3 uma alíquota do N-terminal da U5-102K purificado e 4 uma alíquota da mistura de 1:1 das duas proteínas.

Fonte: Elaborada pela autora.

Na coluna onde foi aplicada a amostra da mistura das proteínas, podemos ver uma soma das bandas das outras colunas, o que indica que não houve interação entre as duas proteínas. Isso pode ter ocorrido porque a U5-102K se encontra em sua maioria agregada. Ou a interação não é forte o bastante para não ser quebrada durante a eletroforese.

Além disso, é possível que a U5-102K N-term, assuma um enovelamento que não é compatível com a interação com a U5-15K, dessa forma seria necessário expressar as duas proteínas na mesma célula.



#### 4.4 Identificação dos parceiros de interação da proteína Prp43

Para a identificação dos parceiros de interação da proteína Prp43, 699 pares de base do seu C-terminal foram clonados no pC-PTP-NEO, como descrito nas seções 3.2.2 e 3.4.2. Dentro dessa região, existe um sítio de restrição para a enzima *EagI*, usada para a clonagem em pC-PTP-NEO. Por esse motivo foi necessário realizar uma mutação nesse sítio de forma a trocar a sequência de bases do códon da Ala 702 (seção 3.3).

Na figura 57 (a) é possível ver o aparecimento de uma banda abaixo de 500pb na digestão feita antes da mutação, indicada pela seta vermelha. Na figura 57 (b), essa banda não aparece, indicando que a construção (pJET\_Prp43) possui apenas um sítio de restrição para a enzima *EagI*.

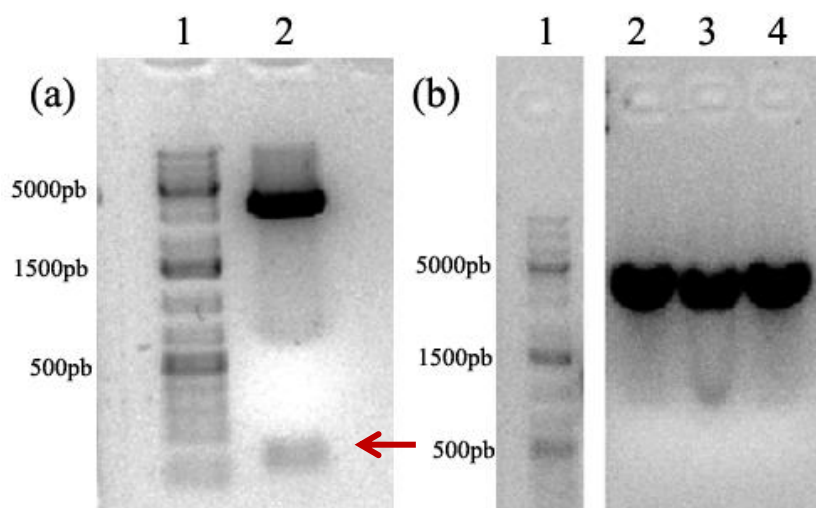


Figura 57 - (a) Digestão do pJET\_Prp43 com *EagI* antes da mutação, a seta vermelha indica a porção que codifica o C-terminal digerida. (b) Digestão do pJET\_Prp43, proveniente de três colônias, com *EagI* após a mutação sítio dirigida.

Fonte: Elaborada pela autora.

Em seguida a sequência do C-terminal foi clonada em pC-PTP-NEO, o vetor foi linearizado com a enzima de restrição *BspEI* e transfectado em TbPCF427, como descrito na seção 3.5.2. A transfecção foi confirmada por *western blot* (seção 3.7), como é possível observar na figura 58 com o aparecimento de uma banda entre 140kDa e 100kDa, que não é vista na cultura de *T brucei* não transfectada.

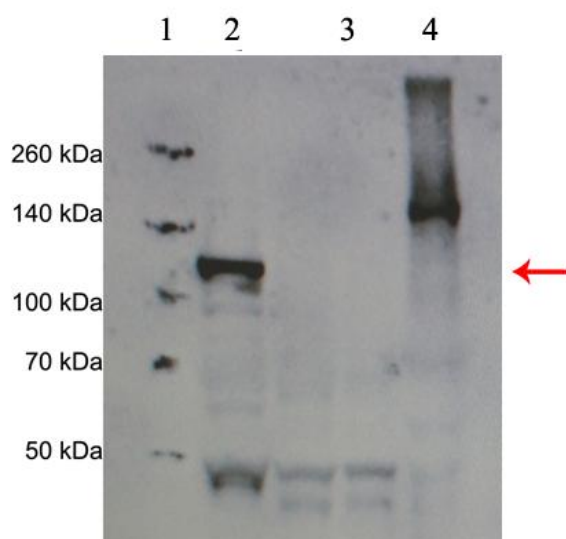


Figura 58 – *Western blot* da confirmação da transfecção da cepa 427 da forma procíclica de *T. brucei* com pC-PTP-NEO\_Pr43, onde a coluna 1 contém o marcador de massa molecular, 2 uma alíquota de *T. brucei* PCF 427 transfectado, 3 a cepa selvagem de *T. brucei* PCF 427 e 4 o controle positivo para o *western blot*. A seta vermelha indica a altura da banda da Prp43 com a cauda PTP.

Fonte: Elaborada pela autora.

A seleção dos clones foi feita em uma placa de 96 poços e a expressão da Prp43 com a cauda PTP foi novamente analisada por *western blot*. Na figura é possível observar a banda de expressão da proteína nos 5 clones testados.

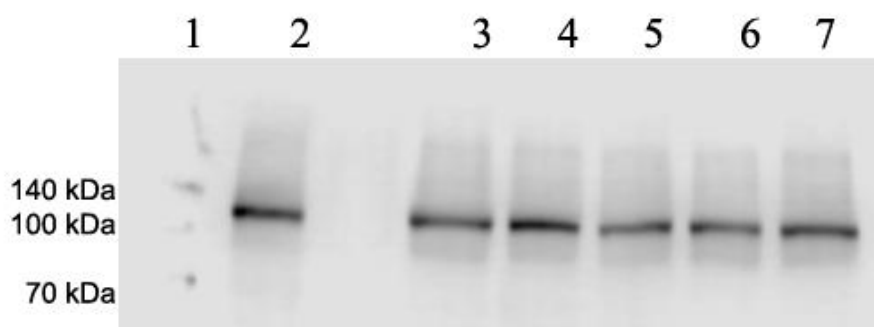


Figura 59 – *Western blot* da seleção dos clones da cepa 427 da forma procíclica de *T. brucei* transfectado com pC-PTP-NEO\_Pr43. A coluna 1 contém marcador de massa molecular, a 2 o controle positivo do *Western blot*, feito do extrato proteico da cultura de *T. brucei* antes da seleção na placa de 96 poços e as colunas de 3 a 7 são colônias selecionadas na placa de 96 poços.

Fonte: Elaborada pela autora.

Dos clones positivos, foi escolhido um e prosseguiu-se com a cultura de *T. brucei* em 4L de meio SDM-79. As células foram lizadas por prensa French e foi feita a purificação dos complexos. As alíquotas retiradas durante esse processo foram aplicadas em SDS-PAGE 12% e o resultado de cada etapa pode ser analisado na figura a seguir.

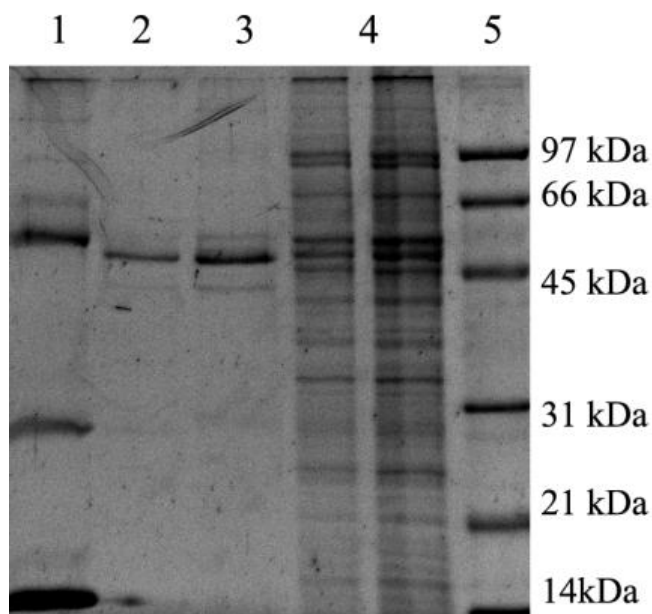


Figura 60 – SDS-PAGE 12% da purificação dos parceiros de interação da proteína Prp43, onde 1 são as proteína purificadas da coluna com anti-proteína C, 2 as proteínas que não interagiram com com a matriz contendo anti-proteína C, 3 as proteínas purificadas da coluna contendo IgG, 4 as proteínas que não interagiram com a matriz contendo IgG e 5 o marcador de massa molecular.

Fonte: Elaborada pela autora.

As bandas da última etapa de purificação (coluna 1 da figura 60) foram recortadas e armazenadas a  $-20^{\circ}\text{C}$ . O cultivo de *T. brucei* em 4L de SDM-79 e subsequente purificação das proteínas foi realizado mais uma vez, como já descrito. Nessa etapa é importante tomar o cuidado de colocar as bandas de mesma massa molecular juntas.

As amostras foram tripsinizadas e dessalinizadas como descrito na seção 3.11.4 e foi feita a identificação das proteínas por espectrometria de massas, no Instituto Butantã, em parceria com a Dra. Júlia Cunha.

As proteínas identificadas com score acima de 50, excluindo-se as comumente encontradas como contaminantes desse tipo de experimento, como chaperonas, elementos de transposição, proteínas ribossomais, fatores de alongação, etc<sup>88</sup>, estão listadas na tabela 12 e figura 61. Também foram excluídas da análise proteínas relacionadas com estresse celular por escassez de alimento.

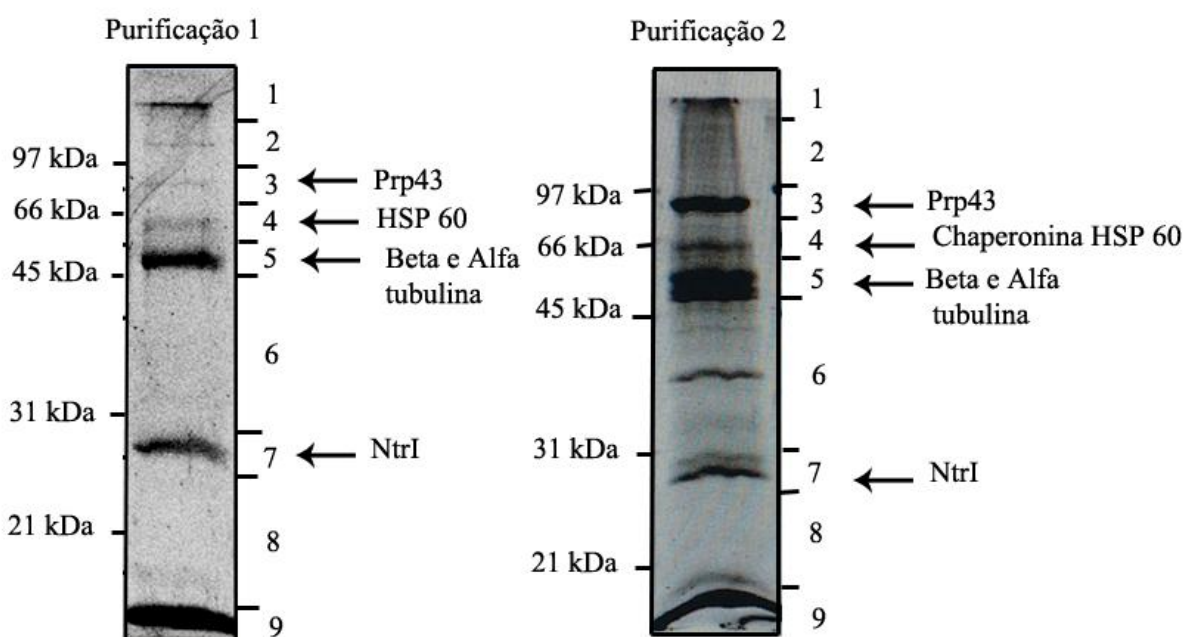


Figura 61 – Identificação por espectrometria de massas das proteínas co-purificadas com a Prp43.  
Fonte: Elaborada pela autora.

Tabela 12 - Proteínas identificadas por espectrometria de massas.

Código TritypDB	Score	Cobertura	Peptídeos	Resíduos	Massa (kDa)
Tb427.05.1150	733	23%	64	735	85,8
Tb427.07.5640	536	30%	9	271	29,5
Tb427.01.2330	451	19%	11	442	52,3
Tb427.01.2340	298	19%	9	451	53,1
Tb427.10.5920	83	2%	2	791	89,9
Tb427.03.1140	64	4%	1	390	44,0

Fonte: Elaborada pela autora.

Das proteínas listadas, foram identificadas duas proteínas hipotéticas associadas ao *splicing*, a Tb427.07.5640 e a Tb427.03.1140. A proteína Tb427.03.1140 possui um domínio SURP, sugerido como domínio de ligação com RNA, que está presente em algumas proteínas relacionadas ao U2 snRNA. A busca por proteínas homólogas com auxílio do BlastP mostrou que pode se tratar de um fator de *splicing* (FS3a). A interação da Prp43 com proteínas do U2 snRNP já havia sido mostrada em humano.<sup>71</sup>

Essa proteína aparece no banco de dados TriTrypDB como sendo uma proteína putativa da família SURP, e foi identificada como sendo o fator de *splicing* SF3a120 em estudos que usaram a proteína SF3a60 fusionada com a cauda PTP.<sup>97</sup>

A análise bioinformática da sequência Tb427.07.5640 no servidor ProtParam resultou em uma proteína que possui uma massa molecular de 29168 Da e ponto isoelétrico calculado de 6,98. Sua composição de aminoácidos se encontra na tabela 13.

Tabela 13 – Composição de aminoácidos na sequência da U5-15K em porcentagem.

Resíduo	Porcentagem (%)	Resíduo	Porcentagem (%)	Resíduo	Porcentagem (%)
<b>Ala (A)</b>	8,1	<b>Gly (G)</b>	11,1	<b>Pro (P)</b>	3,7
<b>Arg (R)</b>	7,4	<b>His (H)</b>	1,8	<b>Sec (U)</b>	0,0
<b>Asn (N)</b>	3,0	<b>Ile (I)</b>	2,6	<b>Ser (S)</b>	10,3
<b>Asp (D)</b>	9,2	<b>Leu (L)</b>	4,4	<b>Thr (T)</b>	5,9
<b>Cys (C)</b>	0,7	<b>Lys (K)</b>	9,2	<b>Trp (W)</b>	1,1
<b>Gln (Q)</b>	2,6	<b>Met (M)</b>	1,5	<b>Tyr (Y)</b>	1,5
<b>Glu (E)</b>	7,4	<b>Phe (F)</b>	1,5	<b>Val (V)</b>	7,0

Fonte: Elaborada pelo autor.

A análise da sua sequência no servidor SMART indicou a presença do domínio conservado G-Patch entre os resíduos 12 e 58, com e-value de  $1,77e^{-09}$ . Esse domínio rico em Glicinas é encontrado em proteínas que se ligam em RNA.

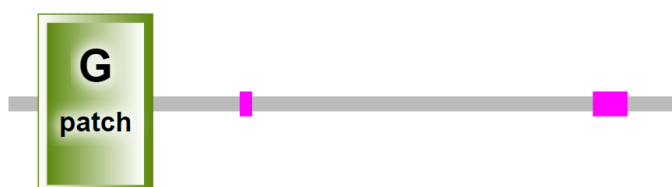


Figura 62 – Localização de domínio G-Patch pelo servidor SMART.

Fonte: Elaborada pela autora

Como já foi mencionado, a helicase Prp43 de levedura interage com o complexo NTR, mais especificamente uma proteína denominada NtrI, durante os ciclos de *splicing*. Proteína essa que possui o domínio G-patch<sup>72</sup>. Alinhando a sequência de *T. brucei* com a de levedura, com auxílio do ClustalW (figura 63) podemos ver que existe conservação na região do domínio G-patch.

Essa análise também mostra que a Ntr1 de *T. brucei* é uma proteína menor que a de levedura, sendo que não foram identificadas sequências em *T. brucei* homólogas entre a região não alinhada.

Tbrucei	-----MSEDPNGTRWLKS	13
Scerevisiae	MEDSDSNTDKKFFFKRRIDSYNYSDEEDNNSMNSDMTYNDALKTSSGNAPTISKLTK	60
	* : *..	
Tbrucei	ENNVGRSLLLRRGGWTEGSLGKEKDGVTTHVKVRQKDG-TMGIGYAGTV-HETWSTQSV	71
Scerevisiae	TYGIGAKLLSSMGYVAGKGLGKDGSGITTPPIETQSRPMHNAGLGMFSNTNNSNYHSENE	120
	.* ** *.. *..***: .*:** ::::: . :* * .. .: :..	
Tbrucei	FA----DVLEKIKAAVSVNDDSDAEGGSSSQAKNEERSDAGGCGKAAKPSGGRHVGM	127
Scerevisiae	YLSSEDEVVEGIEQ-VKFNKTSTEVLGEA-----LLNDSGDMTIVRTLRELRAGVQL	172
	: :*: * : *..* * : : : : * : * . . . * : *	
Tbrucei	KRHALKT-EFL---RAKDG---KYSEEILGSASF-SRKQGRVKDN--TVTADGDDV--	174
Scerevisiae	PESILKELDPLNAVPKPKDVVEILQELLGIEKSLAIRQRTSPLLEVQVKEYYQERLL	232
	. ** : * : *.. : :*:** . . : *.. * . * : *	
Tbrucei	DNAESKAD-SESSRAKRKRRPAQDEKTEEIASTLRSVAVLQRLMVRTP----KHEPRQVS	229
Scerevisiae	SELEVTLRDESKHVSLYDKIGAILKLSDDDELIDRLTSCLLRKELLIEFDLDHLEKPN	292
	.. * . * * : * . . : * : * * : * : : : : * : * : *	
Tbrucei	KGKAE---DGDGSDVDDSD--DSGDGNIVRIVKPYPRPK-----	263
Scerevisiae	DELQIIELLAYRMDTTSKFLNRTQTTFKVV--IYPKLLKFWEGFDMTKSIDSAITLL	350
	. : : : * * : : : * : : * : * * : *	
Tbrucei	-----CTDTPFIE-----	271
Scerevisiae	DFQQVLSFIGCKEHIMEEFVYPKLLQELDNWELHDEVDHVSPIRWLDFMVLIDDKIKDT	410
	*. : : *	
Tbrucei	-----	271
Scerevisiae	IVDKIEAKFFAYCKNWHRESFCITNSDIIFIKELICERRYKILCKEFLPKFLDELWER	470
Tbrucei	-----	271
Scerevisiae	HNDPIYELEDWKEKQEKDSEGGFFFMKKLRSYTHYFHPKQYELMMRGTFNNINKILYQ	530
Tbrucei	-----	271
Scerevisiae	WHLYSTVEDLHKSQWLNWLMNTVFEHSLPTEIELSEIRKSYNIFAMSHRYHLDKSTLDE	590
Tbrucei	-----	271
Scerevisiae	DFDLRQGLRNLMTQVDDISQSEQEPTYTVQNIPLGKVVSSFKDVEDVYCLEKGYLSK	650
Tbrucei	-----	271
Scerevisiae	IPNRYTQLPYGRDQDCIVPLFEIRNGKKKMEVALKHDILWVEDSSGTFKPIYLWALDL	708

Figura 63 – Alinhamento entre as sequências da suposta proteína Ntr1 de *T. brucei* e da Ntr1 de levedura.  
Fonte: Elaborada pela autora

Já o alinhamento entre as possíveis Ntr1 dos trypanosomatídeos mostra que a proteína presente e conservada entre esses organismos, sendo que todas elas não menores em relação à proteína de levedura (figura 64).

Essa identificação sugere que os trypanosomatídeos, assim como a levedura, possuem um complexo NTR, e não apenas a Prp43, como ocorre em humano. Apesar disso, apenas com a nova identificação não é possível afirmar se a Ntr1 de *T. brucei* tem papel relevante na liberação do Y-intron.







## 5 Conclusões

Nesse trabalho foram apresentados os resultados da participação da cisteína 66 e das quatro serinas mais próximas ao C-terminal da U5-15K na atividade de autoclivagem, da caracterização estrutural da U5-15K sem alça, da caracterização dos parceiros de interação da U5-102K, da obtenção do subcomplexo formado pelas proteínas U5-15K e U5-102K e da caracterização dos parceiros de interação da helicase Prp43.

O testes de clivagem das proteínas resultantes da mutação individual da cisteína 66, das serinas 120, 126, 137 e 144 e de 5 aminoácidos próximos ao ponto de clivagem calculado anteriormente mostram que a serina 126 está diretamente ligada à atividade de autoclivagem, pode ser o resíduo catalítico, uma vez que a sua substituição por uma alanina inibe a atividade.

As mutantes U5-15K pc, U5-15K S137A e U5-15K S144A também ajudaram a definir o ponto de clivagem da U5-15K. No teste de clivagem dessas mutantes, há o surgimento de mais de uma banda de clivagem, mostrando que a adição de alaninas no C-terminal insere novos pontos de clivagem. Isso sugere que a autoclivagem ocorre em uma alanina do C-terminal, mais especificamente a alanina142.

O resíduo catalítico da autoclivagem ser a serina 126, ao invés da cisteína 66, como tinha sido proposto em trabalhos anteriores<sup>64</sup>, o que permite que a autoclivagem ocorra dentro do complexo B, contribuem para a ideia da autoclivagem estar ligada à sinalização para a saída da U5-15K e ativação do complexo B.

A modelagem por homologia da U5-15K permitiu a visualização de uma alça flexível que poderia dificultar a cristalização da proteína. Baseado nisso foi traçada a estratégia de deleção de 11 aminoácidos da alça flexível para novos ensaios de cristalização. Os experimentos de DSF e clivagem da proteína mutante U5-15K truncados mostram que mesmo com essa deleção a estrutura globular e a atividade de autoclivagem são mantidas. Apesar da proteína mutante possuir uma estrutura termicamente tão estável quanto a da proteína nativa, não foram observados cristais nos ensaios de cristalização. Já os ensaios usando a U5-15K S126A resultaram em pequenos cristais. O que indica que a autoclivagem é mais relevante para a cristalização que a presença da alça flexível e que os pontos de contato pra formação do cristal da proteína de *T. brucei* são diferentes dos da proteína de humano.

Os experimentos de purificação dos parceiros de interação da proteína U5-102K não resultaram em proteínas que já não tivessem sido caracterizadas pela mesma técnica usando a

U5-15K como sonda. Isso pode ter ocorrido porque a cauda-PTP no C-terminal da U5-102K atrapalha seu enovelamento e entrada na partícula U5/U4.U6 tri-snRNP e complexo B, ou pela quantidade de amostra usada para a identificação por espectrometria de massas ter sido insuficiente, mais experimentos são necessários.

Baseado na interação entre Dib1 e Prp6 em levedura, e U5-15k e U5-102k em humano, foram desenhados oligonucleotídeos para a expressão de uma porção do N-terminal da U5-102K de aproximadamente 15kDa. As tentativas de interação dessa proteína com a U5-15K resultaram na formação de partículas de  $50 \pm 20$  Å, calculados por DLS, que equivalem ao tamanho de um dímero formado por essas duas proteínas.

A caracterização dos parceiros de interação da helicase Prp43 mostram que essa proteína interage com supostas proteínas de *splicing* sendo elas um fator de *splicing* da partícula U2 (FS3a120) que contem o domínio SURP, proteína essa que já havia sido identificada a partir da marcação da proteína SF3a60<sup>97</sup>, e uma nova proteína que contem o domínio G-patch, homóloga à Ntr1 de levedura. Essa proteína está presente e conservada em outros tripanossomatídeos, sendo menor que a Ntr1 de levedura e homóloga apenas ao seu N-terminal.

## 6 Perspectivas futuras

Dadas as conclusões expostas no item anterior, algumas perspectivas de experimento se abrem, as principais estão listadas abaixo:

- Inibição da autoclivagem em *T. brucei* para avaliar se ela possui uma função no processo *splicing*.
- Expansão das condições de cristalização usando a U5-15K S126A para obtenção de cristais maiores.
- Nova purificação dos parceiros de interação, usando a cauda PTP no N-terminal da U5-102K.
- Análise do efeito do *knockdown* da NtrI na reação de *trans-splicing*.



## REFERÊNCIAS

- 1 WORLD HEALTH ORGANIZATION. **Control and surveillance of human african trypanosomiasis**: report of a WHO expert committee. Geneva, Switzerland: World Health Organization; 2013. (WHO technical report series, 984)
- 2 SIMARRO, P. P.; CECCHI, G.; FRANCO, J. R.; PAONE, M.; DIARRA, A.; RUIZ-POSTIGO, J.A.; FÈVRE, E. M.; MATTIOLI, R. C.; JANNIN, J. G. Estimating and mapping the population at risk of sleeping sickness. **PLoS Neglected Tropical Disease**, v. 6, n. 10. 2012. DOI: 10.1371/journal.pntd.0001859.
- 3 STEVERDING, D. The history of African *trypanosomiasis*. **Parasites & Vectors**, v. 12, n. 1, p.1-3. 2008. DOI: 10.1186/1756-3305-1-3.
- 4 SIMARRO, P. P.; JANNIN, J.; CATTAND, P. Eliminating human African *trypanosomiasis*: where do we stand and what comes next? **PLoS Medicine**, v. 5, n. 2. 2008. DOI: 10.1371/journal.pmed.0050055.
- 5 BRUCE, D. **Preliminary report on the tsetse fly disease or nagana in Zululand**. Durban: Bennett and Davis; 1895.
- 6 HEADRICK, D. R. Sleeping sickness epidemics and colonial responses in east and Central Africa, 1900–1940. **PLoS Neglected Tropical Diseases**, v. 8, n. 4. 2014. DOI: 10.1371/journal.pntd.0002772.
- 7 WORLD HEALTH ORGANIZATION **WHO report on global surveillance of epidemic-prone infectious diseases**. African trypanosomiasis.2000. chapter 8, p.95–106
- 8 FRANCO, J. R.; SIMARRO, P. P.; DIARRA, A.; JANNIN, J. G. Epidemiology of human African trypanosomiasis. **Clinical Epidemiology**, v. 6, p. 257-275, 2014. DOI: 10.2147/CLEP.S39728.
- 9 FRANCO, J. R.; SIMARRO, P. P.; DIARRA, A.; RUIZ-POSTIGO, J. A.; JANNIN, J. G. The journey towards elimination of gambiense human African trypanosomiasis: not far, nor easy. **Parasitology**, v. 141, n. 6, p. 748-760, 2013. DOI: 10.1017/S0031182013002102.
- 10 FRANCO, J. R.; CECCHI, G.; PRIOTTO, G.; PAONE, M.; DIARRA, A.; GROUT, L.; SIMARRO, P. P.; ZHAO, W.; ARGAW, A. Monitoring the elimination of human African trypanosomiasis: update to 2016. **PLoS Neglected Tropical Disease**, v. 12, n. 12, 2018. DOI: 10.1371/journal.pntd.0006890.
- 11 BARRETT, M. P. The elimination of human African trypanosomiasis is in sight: Report from the third WHO stakeholders meeting on elimination of gambiense human African trypanosomiasis. **PLoS Neglected Tropical Disease**, v. 12, n.12, 2018. DOI: 10.1371/journal.pntd.0006925.
- 12 MOORE, S.; SHRESTHA, S.; TOMLINSON, K. W.; VUONG, H. Predicting the effect of climate change on African trypanosomiasis: integrating epidemiology with parasite and vector biology. **Journal of the Royal Society Interface**, v. 9, n. 70, p. 817-830, 2012. DOI: 10.1098/rsif.2011.0654.

13 TONGUE, L. K.; MAVOUNGOU, J. F.; KAMKUMO, R. G.; KABA, D.; HENDJI, G. C. F.; PAMBA, R.; M'EIYI, P. M.; MBATCHI, B.; LOUIS, F. Human African trypanosomiasis in suburban and urban areas: a potential challenge in the fight against the disease. **Journal of Clinical & Experimental Pathology**, 2012. DOI: 10.4172/2161-0681.S3-002.

14 AKSOY, S. Control of tsetse flies and trypanosomes using molecular genetics. **Veterinary Pathology**, v. 115, n. 2, p.125-145, 2003.

15 KENNEDY, P. G. Clinical features, diagnosis, and treatment of human African trypanosomiasis (sleeping sickness). **Lancet Neurology**, v. 12, n. 2, p. 186-194, 2012. DOI: 10.1016/S1474-4422(12)70296-X.

16 KENNEDY P. G. Human African trypanosomiasis of the CNS: current issues and challenges. **Journal of Clinical Investigation**, v. 113, n. 4, p.496-504, 2004. DOI: 10.1172/JCI21052

17 KENNEDY, P. G. Human African trypanosomiasis neurological aspects. **Journal of Neurology**, v.253, n. 4, p. 411-416, 2006. DOI: 10.1007/s00415-006-0093-3.

18 DONELSON, J. E. Antigenic variation and the African trypanosome genome. **Acta Tropica**, v. 85, p. 391–404, 2002.

19 KENNEDY, P. G.; RODGERS, J. Clinical and neuropathogenetic aspects of human african trypanosomiasis. **Frontiers in Immunology**, v. 10, p. 39, 2019. DOI: 10.3389/fimmu.2019.00039.

20 BISSER1, S.; LUMBALA, C.; NGUERTOUM, E.; KANDE, V.; FLEVAUD, L.; VATUNGA, G.; BOELAERT, M.; BÜSCHER, P.; JOSENANDO, T.; BESSELL, P. R.; BIÉLER, S.; NDUNG'U, J. M. sensitivity and specificity of a prototype rapid diagnostic test for the detection of trypanosoma brucei gambiense infection: a multi-centric prospective study. **PLoS Neglected Tropical Diseases**, v.10, n. 4, 2016. DOI: 10.1371/journal.pntd.0004608.

21 MESU, V. K. B. K.; KALONJI, W. M.; BARDONNEAU, C.; MORDT O. V.; BLESSON S.; SIMON F.; DELHOMME, S.; BERNHARD S.; KUZIENA, W.; LUBAKI, J. F.; VUVU, S. L.; NGIMA P. N.; MBEMBO, H. M.; ILUNGA M.; BONAMA, A. K.; HERADI, J. A.; SOLOMO, J. L. L.; MANDULA, G.; BADIBABI, L. K.; DAMA, F. R.; LUKULA, P. K.; TETE, D. N.; LUMBALA, C.; SCHERRER, B.; STRUB-WOURGAFT, N.; TARRAL, A. Oral fexinidazole for late-stage African Trypanosoma brucei gambiense trypanosomiasis: a pivotal multicentre, randomised, non-inferiority trial. **Lancet**, v. 391, n. 10116, p. 144-154, 2017. DOI: 10.1016/S0140-6736(17)32758-7.

22 PÉREZ-MOLINA, J. A.; MOLINA, I. Chagas disease. **Lancet**, v. 391, n. 10115, p. 82-94, 2018. DOI: 10.1016/S0140-6736(17)31612-4.

23 SINGH, O. P.; HASKER, E.; BOELAERT, M.; SUNDAR, S. Elimination of visceral leishmaniasis on the Indian Subcontinent: critical knowledge gaps to be answered to get us

there. **Lancet Infectious Diseases**, v. 16, n. 12, p. 304-309, 2016. DOI: 10.1016/S1473-3099(16)30140-2

24 ANDRADE-NETO, V. V.; CUNHA-JUNIOR, E. F.; FAIOES, V. S.; MARTINS, T. P.; SILVA, R. L.; LEON, L. L.; TORRES-SANTOS, E. C. Leishmaniasis treatment: update of possibilities for drug repurposing. **Frontiers in Bioscience**, v. 23, p. 967-996, 2018. DOI: 10.2741/4629.

25 MARUYAMA, S. R.; DE SANTANA, A. K. M.; TAKAMIYA, N. T.; TAKAHASHI, T. Y.; ROGERIO, L. A.; OLIVEIRA, C. A. B.; MILANEZI, C. M.; TROMBELA, V. A.; CRUZ, A. K.; JESUS, A. R.; BARRETO, A. S.; DA SILVA, A. M.; ALMEIDA, R. P.; RIBEIRO, J. M.; SILVA, J. S. Non-*Leishmania* parasite in fatal visceral leishmaniasis-like disease, Brazil. **Emerging Infectious Diseases**, v. 25, n. 11, p. 2088-2092, 2019. DOI: 10.3201/eid2511.181548

26 LI, F. J.; HE, C. Y. Autophagy in protozoan parasites: *Trypanosoma brucei* as a model. **Future Microbiology**, v. 37, n.1337-1340, 2017. DOI: 10.2217/fmb-2017-0158

27 MARTINEZ-CALVILLO, S.; VIZUET-DE-RUEDA, J. C.; FLORENCIO-MARTINEZ, L. E.; MANNING-CELA, R. G.; FIGUEROA-ÂNGULO, E. E. Gene Expression in Trypanosomatid parasites. **Journal of Biomedicine and Biotechnology**, v.2010, n.525241, 2010. DOI:10.1155/2010/525241.

28 WAHL, M. C.; WILL, C. L.; LÜHRMANN, R. The spliceosome: design principles of a dynamic RNP machine. **Cell**, v. 136, p. 701–718, 2009. DOI: 10.1016/j.cell.2009.02.009.

29 WILL, C. L.; LÜHRMANN, R. Spliceosome structure and function. **Cold Spring Harbour Perspective in Biology**, v. 3, p. 1–2, 2011. DOI: 10.1101/cshperspect.a003707.

30 DE CONTI, L.; BARALLE, M.; BURATTI, E. Exon and intron definition in pre-mRNA splicing. **WIREs RNA**, v. 4, p. 49–60, 2013. DOI: 10.1002/wrna.1140.

31 PATZELT, E.; PERRY, K. L.; AGABIAN, N. Mapping of branch sites in trans-spliced pre-mRNAs of *Trypanosoma brucei*. **Molecular and Cellular Biology**, v. 9, n.10 , p. 4291-4297, 1989.

32 FICA, S. M.; TUTTLE, N.; NOVAK, T.; LI, N. S.; LU, J.; KODATHINGAL, P.; DAI, Q.; STALEY, J. P.; PICCIRILLI, J. A. RNA catalyzes nuclear pre-mRNA splicing. **Nature**, v. 503, n. 7475, p. 229–234, 2013. DOI:10.1038/nature12734.

33 MAYER, M. G.; FLOETER-WINTER, L. M. Pre-mRNA trans-splicing: from kinetoplasts to mammals, an easy language for life diversity. **Memórias do Instituto Oswaldo Cruz**, v. 100, p. 501-513, 2005.

34 MICHAELI, S. Trans-splicing in trypanosomes: machinery and its impact on the parasite transcriptome. **Future Microbiology**, v. 6, n. 4, p. 459–474, 2011.

35 LIANG, X. H.; HARITAN, A.; ULIEL, S.; MICHAELI, S. Trans and cis splicing in trypanosomatids: mechanism, factors, and regulation. **Eukaryot Cell**, v. 2, n. 5, p. 830-40, 2003.

36 HERMANN, H.; FABRIZIO, P.; RAKER, V. A.; FOULAKI, K.; HORNIG, H.; BRAHMS, H.; LÜHRMANN, R. snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein interactions. **EMBO Journal**, v. 14, n. 9, p. 2076-2088, 1995.

37 SÉRAPHIN, B. Sm and Sm-like proteins belong to a large family: identification of proteins of the U6 as well as the U1, U2, U4 and U5 snRNPs. **EMBO Journal**, v. 14, n. 9, p. 2089-2098, 1995.

38 KAMBACH, C.; WALKE, S.; YOUNG, R.; AVIS, J. M.; DE LA FORTELLE, E.; RAKER, V. A.; LÜHRMANN, R.; LI, J.; NAGAI, K. Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. **Cell**, v. 96, n. 3, p. 375-387, 1999.

39 BOESLER, C.; RIGO, N.; ANOKHINA, M. M.; TAUCHERT, M. J.; AGAFONOV, D. E.; KASTNER, B.; URLAUB, H.; FICNER, R.; WILL, C. L.; LÜHRMANN, R. A spliceosome intermediate with loosely associated tri-snRNP accumulates in the absence of Prp28 ATPase activity. **Nature**, v. 7, n. 11997, 2016. DOI:10.1038/ncomms11997.

40 VALADKHAN, S. snRNAs as the catalysts of pre-mRNA splicing. **Current Opinion on Chemical Biology**, v. 9, n. 6, p. 603-608, 2005.

41 BESSONOV, S.; ANOKHINA, M.; KRASAUSKAS, A.; GOLAS, M. M.; SANDER, B.; WILL, C. L.; URLAUB, H.; STARK, H.; LÜHRMANN, R. Characterization of purified human B(act) spliceosomal complexes reveals compositional and morphological changes during spliceosome activation and first step catalysis. **RNA**, v. 16, n. 12, p. 2384–2403, 2010. DOI: 10.1261/rna.2456210.

42 OHRT, T.; ODENWÄLDER, P.; DANNENBERG, J.; PRIOR, M.; WARKOCKI, Z.; SCHMITZOVÁ, J.; KARADUMAN, R.; GREGOR, I.; ENDERLEIN, J.; FABRIZIO, P.; LÜHRMANN, R. Molecular dissection of step 2 catalysis of yeast pre-mRNA splicing investigated in a purified system. **RNA**, v. 19, p. 902–915, 2013.

43 BERTRAM, K.; AGAFONOV, D. E.; LIU, W.T.; DYBKOV, O.; WILL, C. L.; HARTMUTH, K.; URLAUB, H.; KASTNER, B.; STARK, H.; LÜHRMANN, R. Cryo-EM structure of a human spliceosome activated for step 2 of splicing. **Nature**, v. 542, p. 318-323, 2017. DOI:10.1038/nature21079.

44 BAI, R.; YAN, C.; WAN, R.; LEI, J.; SHI, Y. Structure of the Post-catalytic Spliceosome from *Saccharomyces cerevisiae*. **Cell**, v. 171, p. 1589–1598, 2017

45 GÜNZL, A. The Pre-mRNA splicing machinery of trypanosomes: complex or simplified? **Eukaryotic Cell**, v. 9, n. 8, p. 1159–1170, 2010. DOI:10.1128/EC.00113-10.

46 DENKER, J. A.; ZUCKERMAN, D. M.; MARONEY, P. A.; NILSEN, T. W. New components of the spliced leader RNP required for nematode trans-splicing. **Nature**, v. 417, p.667–670, 2002.



47 PREUßER, C.; JAÉ, N.; BINDEREIF, A. mRNA splicing in trypanosomes. **International Journal of Medical Microbiology**, v. 302, n. 4, p. 221–224, 2012. DOI:10.1016/j.ijmm.2012.07.004.

48 HINTERBERGER, M.; PETTERSSONG, I.; STEITZ, J. A. Isolation of small nuclear ribonucleoproteins containing U1, U2, U4, U5, and U6 RNAs. **Journal of Biological Chemistry**, v. 258, n. 4, p. 2601–2613, 1983.

49 MUTO, Y.; KRUMMEL, D. P.; OUBRIDGE, C.; HERNANDEZ, H.; ROBINSON, C. V.; NEUHAUS, D.; NAGAI, K. The structure and biochemical properties of the human spliceosomal protein U1C **Journal of Molecular Biology**, v. 341, p. 185–198, 2004. DOI:10.1016/j.jmb.2004.04.078.

50 KONDO, Y.; OUBRIDGE, C.; VAN ROON, A. M. M.; NAGAI, K. Crystal structure of human U1 snRNP, a small nuclear ribonucleoprotein particle, reveals the mechanism of 5' splice site recognition. **eLife**, n.4,p.e04986, 2015. DOI: 10.7554/eLife.04986.

51 KRÄMER, A., FERFOGLIA, F., HUANG, C.-J., MULHAUPT, F., NESIC, D., TANACKOVIC, G. Structure–function analysis of the U2 snRNP-associated splicing factor SF3a. **Biochemical Society Transactions**, v. 33, n. 3, p. 439–442, 2005 DOI:10.1042/bst0330439.

52 VAN ROON, A. M. M.; OUBRIDGE, C.; OBAYASHI, E.; SPOSITO, B.; NEWMAN, A. J.; SÉRAPHIN, B.; NAGAI, K. Crystal structure of U2 snRNP SF3b components: Hsh49p in complex with Cus1p-binding domain. **RNA**, v. 23, n. 6, p. 968–981, 2017. DOI: 10.1261/rna.059378.116.

53 MARTELLY, W., FELLOWS, B., SENIOR, K., MARLOWE, T., SHARMA, S. Identification of a non-canonical RNA binding domain in the U2 snRNP protein SF3A1. **RNA**, v. 25, n. 11, p. 1509–1521, 2019. DOI:10.1261/rna.072256.119.

54 HARDIN, J. W.; WARNASOORIYA, C.; KONDO, Y.; NAGAI, N.; RUEDA, D. Assembly and dynamics of the U4/U6 di-snRNP by single-molecule FRET. **Nucleic Acids Research**, v. 43, n. 22, p. 10963–10974, 2015. DOI: 10.1093/nar/gkv1011.

55 FABRIZIO, P.; DANNENBERG, J.; DUBE, P.; KASTNER, B.; STARK, H.; URLAUB, H.; LÜHRMANN, R. The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome. **Journal of Molecular Cell Biology**, v. 36, n. 4, p. 593–608, 2009. DOI: 10.1016/j.molcel.2009.09.040.

56 LIU, S.; RAUHUT, R.; VORNLOCHER, H. P.; LÜHRMANN, R. The network of protein–protein interactions within the human U4/U6.U5 tri-snRNP. **RNA**, v. 12, p. 1418–1430. 2006.

57 LAGGERBAUER, B.; LIU, S.; MAKAROV, E.; VORNLOCHER, H. P.; MAKAROVA, O.; INGELFINGER, D.; ACHSEL, T.; LÜHRMANN, R. The human U5 snRNP 52K protein (CD2BP2) interacts with U5-102K (hPrp6), a U4/U6.U5 tri-snRNP bridging protein, but dissociates upon tri-snRNP formation. **RNA Journal**, v.11, p. 598608, 2005. DOI: 10.1261/rna.2300805.

- 58 LUZ AMBRÓSIO, D.; LEE, J. H.; PANIGRAHI, A. K.; NGUYEN, T. N.; CICARELLI, R. M.; GÜNZL, A. Spliceosomal proteomics in *Trypanosoma brucei* reveal new RNA splicing factors. **Eukaryot Cell**, v. 8, n. 7, p. 990-1000, 2009. DOI: 10.1128/EC.00075-09.
- 59 DA SILVA, M. T. A.; AMBRÓSIO, D. L.; TREVELIN, C. C.; WATANABE, T. F.; LAURE, H. J.; GREENE, L. J.; ROSA, J. C., VALENTINI, S. L., CICARELLI, R. M. B. New insights into trypanosomatid U5 small nuclear ribonucleoproteins. **Memórias do Instituto Oswaldo Cruz**, v.106, n. 2, p. 130-138, 2011.
- 60 LÜTZELBERGER, M.; BOTTNER, C. A.; SCHWELNUS, W.; ZOCK-EMMENTHAL, S.; RAZANAU, A.; KÄUFER, N. F. The N-terminus of Prp1 (Prp6/U5-102 K) is essential for spliceosome activation in vivo. **Nucleic Acids Research**, v. 38, n. 5, p. 1610-1622, 2010. DOI: 10.1093/nar/gkp1155
- 61 SU, A. I; WILTSHIRE, T.; BATALOV, S.; LAPP, H.; CHING, K. A.; BLOCK, D.; ZHANG, J.; SODEN, R.; HAYAKAWA, M.; KREIMAN, G.; COOKE, M. P.; WALKER, J. R.; HOGENESCH, J. B. A gene atlas of the mouse and human protein-encoding transcriptomes. **Proceedings of the National Academy of Science**, v. 101, p. 6062–6067, 2004.
- 62 ZHANG, Y. Z.; GOULD, K. L.; DUNBRACK, R. L. JR.; CHENG, H.; RODER, H.; GOLEMIS, E. A. The evolutionarily conserved Dim1 protein defines a novel branch of the thioredoxin fold superfamily. **Physiological Genomics**, v. 1, n. 3, p. 109–118, 1999.
- 63 REUTER, K.; NOTTROTT, S.; FABRIZIO, P.; LÜHRMANN, R.; FICNER, R. Identification, characterization and crystal structure analysis of the human spliceosomal U5 snRNP specific 15 Kd Protein. **Journal of Molecular Biology**, v. 294, n. 2, p. 515-525, 1999.
- 64 LIMA, A. L. **Caracterização estrutural da proteína spliceosomal de *Trypanosoma brucei* U5-15k**. 2015. 99p. Dissertação (Mestrado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2015.
- 65 TENCHUAN, J.; GUO, F.; WANG, Y.; ZHANG, Y. Z. Identification of Human dim1 as a Peptidase with Autocleavage Activity. **Chemical Biology Drug Design**, v. 68, p. 266–272, 2006. DOI: 10.1111/j.1747-0285.2006.00447
- 66 AN OVERVIEW of the flow of information from DNA to protein in a eukaryote. Disponível em: <https://www.nature.com/scitable/content/an-overview-of-the-flow-of-information-14711098/>. Acesso em: 06 de fevereiro de 2020.
- 67 WAN, R.; YAN, C.; BAI, R.; WANG, L.; HUANG, M.; WONG, C. C.; SHI, Y. The 3.8 Å structure of the U4/U6.U5 tri-snRNP: Insights into spliceosome assembly and catalysis. **Science**, v. 351, n. 6272, p. 466-475, 2016. DOI: 10.1126/science.aad6466
- 68 NGUYEN, T. H. D.; SCHERES, S. H. W.; NAGAI, K. Cryo-EM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 Å resolution. **Nature**, v. 530, p. 298-302, 2016. DOI:10.1038/nature16940

- 69 SCHREIB, C. C.; BOWMAN, E. K.; HERNANDEZ, C. A.; LUCAS, A. L.; POTTS, C. H. S.; MAEDER, C. Functional and biochemical characterization of Dib1's role in pre-messenger RNA splicing. **Journal of Molecular Biology**, v. 430, n. 11, p. 1640-1651, 2018. DOI: 10.1016/j.jmb.2018.04.027
- 70 HE, Y.; STALEY, J. P.; ANDERSEN, G. R.; NIELSEN, K. H. Structure of the DEAH/RHA ATPase Prp43p bound to RNA implicates a pair of hairpins and motif Va in translocation along RNA. **RNA Journal**, v. 23, n. 7, p. 1110-1124, 2017. DOI: 10.1261/rna.060954.117.
- 71 KOODATHINGAL, P.; NOVAK, T.; PICCIRILLI, J. A.; STALEY, J. P. The DEAH box ATPases Prp16 and Prp43 cooperate to proofread 5' splice site cleavage during pre-mRNA splicing. **Molecular Cell**, v. 39, n. 3, p. 385-95, 2010. DOI: 10.1016/j.molcel.2010.07.014.
- 72 CHRISTIAN, H.; HOFELE, R. V.; URLAUB, H.; FICNER, R. Insights into the activation of the helicase Prp43 by biochemical studies and structural mass spectrometry. **Nucleic Acids Research**, v. 42, n. 2, p. 1162-79, 2013. DOI: 10.1093/nar/gkt985.
- 73 GASTEIGER, E.; HOOGLAND, C.; GATTIKER, A.; DUVAUD, S.; WILKINS, M. R.; APPEL, R. D.; BAIROCH, A. Protein identification and analysis tools on the ExPASy server. **Methods in Molecular Biology**, v. 112, p. 531- 552, 1999.
- 74 BUCHAN, D. W. A.; JONES, D.T. The PSIPRED Protein analysis workbench: 20 years on. **Nucleic Acids Research**, v. 41, n. W1, p. 402-407, 2019. <https://DOI.org/10.1093/nar/gkz297>.
- 75 JONES, D. T. Protein secondary structure prediction based on position-specific scoring matrices. **Journal of Molecular Biology**, v. 292, p. 195-202, 1999.
- 76 MADEIRA, F.; PARK, Y. M.; LEE, J.; BUSO, N.; GUR, T.; MADHUSOODANAN, N.; BASUTKAR, P.; TIVEY, A. R. N.; POTTER, S. C.; FINN, R. D.; LOPEZ, R. The EMBL-EBI search and sequence analysis tools APIs in 2019 **Nucleic Acids Research**, v. 47, n. W1, p. W636–W641, 2019. <https://DOI.org/10.1093/nar/gkz268>
- 77 LIU, H.; NAISMITH, J. H. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. **BMC Biotechnol**, v. 4, n. 8, p. 91, 2008. DOI: 10.1186/1472-6750-8-91.
- 78 BRUN, R.; SCHONENBERGER, M. Cultivation in vitro cloning or procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. **Acta Tropical**, v. 36, n. 3, p. 289-292, 1979.
- 79 DA SILVA, M.T.A. **Identificacao e caracterização funcional de proteínas específicas do complexo u5 snrnp em tripanosomatídeos**. 2009. 98p. Tese (Doutorado) – Instituto de Química, Universidade Estadual Paulista, Araraquara, 2009.
- 80 WATERHOUSE, A.; BERTONI, M.; BIENERT, S.; STUDER, G.; TAURIELLO, G.; GUMIENNY, R.; HEER, F.T.; DE BEER, T.A.P.; REMPFER, C.; BORDOLI, L.; LEPORE, R.; SCHWEDE, T. SWISS-MODEL: homology modelling of protein structures and complexes. **Nucleic Acids Research**, v. 46, n. W1, p. W296-W303, 2018.

81 BIENERT, S.; WATERHOUSE, A.; DE BEER, T.A.P.; TAURIELLO, G.; STUDER, G., BORDOLI, L.; SCHWEDE, T. The SWISS-MODEL Repository - new features and functionality. **Nucleic Acids Research**, v. 45, p. D313-D319, 2017.

82 GUEX, N.; PEITSCH, M.C.; SCHWEDE, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. **Electrophoresis**, v. 30, p. S162-S173, 2009.

83 BENKERT, P.; BIASINI, M.; SCHWEDE, T. Toward the estimation of the absolute quality of individual protein structure models. **Bioinformatics**, v. 27, p. 343-350, 2011.

84 BERTONI, M.; KIEFER, F.; BIASINI, M.; BORDOLI, L.; SCHWEDE, T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. **Scientific Reports**, v. 7. 2017 DOI: 10.1038/s41598-017-09654-8

85 PANTOLIANO, M. W.; PETRELLA, E. C.; KWASNOSKI, J. D.; LOBANOV, V. S.; MYSLIK, J.; GRAF, E.; CARVER, T.; ASEL, E.; SPRINGER, B. A.; LANE, P.; SALEMME, F. R. High-density miniaturized thermal shift assays as a general strategy for drug discovery. **Journal of Biomolecular Screen**, v. 6, n. 6, p. 429–440, 2001.

86 SCHIMANSKI, B.; NGUYEN, T. N.; GÜNZL, A. Highly efficient tandem affinity purification of trypanosome protein complexes based on a novel epitope combination. **Eukaryotic Cell**, v.4, n. 1, p. 1942–1950, 2005. DOI:10.1128/EC.4.11.1942–1950.2005

87 Novagen® pETDuet-1 vector. Disponível em: [https://www.helmholtz-muenchen.de/fileadmin/PEPF/pET\\_vectors/pETDuet-1\\_map.pdf](https://www.helmholtz-muenchen.de/fileadmin/PEPF/pET_vectors/pETDuet-1_map.pdf). Acesso em: 05 de dezembro de 2019.

88 BERTRAM, K.; AGAFONOV, D. E.; DYBKOV, O.; HASELBACH, D.; LEELARAM, M. N.; WILL, C. L.; URLAUB, H.; KASTNER, B.; LÜHRMANN, R.; STARK, H. Cryo-EM Structure of a pre-catalytic human spliceosome primed for activation. **Cell**, v. 170, n. 4, p. 701-713. 2017. DOI: 10.1016/j.cell.2017.07.011.

89 XU, X.; SONG, Y.; LI, Y.; CHANG, J.; ZHANG, H.; AN, L. The tandem affinity purification method: an efficient system for protein complex purification and protein interaction identification. **Protein Expression and Purification**, v. 72, n. 2, p. 149-56, 2010. DOI: 10.1016/j.pep.2010.04.009.

90 THERMO Scientific phusion high-fidelity DNA polymerase. Disponível em: [https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0012393\\_Phusion\\_HighFidelity\\_DNAPolymerase\\_UG.pdf](https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0012393_Phusion_HighFidelity_DNAPolymerase_UG.pdf). Acesso em: 11 de dezembro de 2019.

91 THERMO Scientific T4 DNA Ligase. Disponível em: [https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0011990\\_T4\\_DNA\\_Ligase\\_5\\_Weiss\\_200\\_Weiss\\_U\\_UG.pdf](https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0011990_T4_DNA_Ligase_5_Weiss_200_Weiss_U_UG.pdf). Acesso em: 11 de dezembro de 2019.

92 THERMO Scientific Taq DNA polymerase. Disponível em: [https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0012027\\_TaqDNAPolymerase\\_recombinant\\_5\\_UuL\\_500U\\_UG.pdf](https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0012027_TaqDNAPolymerase_recombinant_5_UuL_500U_UG.pdf). Acesso em: 11 de dezembro de 2019.

93 KIM, H. S.; LI, Z.; BOOTHROYD, C.; CROSS, G. A. Strategies to construct null and conditional null *Trypanosoma brucei* mutants using Cre recombinase and loxP. **Molecular and Biochemical Parasitology**, v. 193, p.16-19, 2013.

94 HOARE, C. A.; WALLACE, F. G. Developmental stages of trypanosomatid flagellates: a new terminology. **Nature**, v. 212, p. 1385-1386, 1966.

95 BÜSCHER, P.; GILLEMANN, Q.; LEJON, V. Rapid diagnostic test for sleeping sickness. **The New England Journal of Medicine**, v.11, n. 368, p.1069-1070, 2013.

96 LUMBALA, C.; BIÉLER, S.; KAYEMBE, S.; MAKABUZA, J.; ONGARELLO, J.; NDUNG'U, J. M. Prospective evaluation of a rapid diagnostic test for *Trypanosoma brucei gambiense* infection developed using recombinant antigens. **PLoS Neglected Tropical Disease**, v.3, n. 12, p.e0006386, 2018.

97 NYAMBEGA, B.; HELBIG, C.; MASIGA, D. K.; CLAYTON, C.; LEVIN, M. J. Proteins associated with SF3a60 in *T. brucei*. **Plos One**, v.9, n. 3, p.e91956, 2014.