

**ANÁLISE ESTRUTURAL DO COMPONENTE
TELOMERASE TRANSCRIPTASE REVERSA DE
Leishmania major.**

MARÍA ALEJANDRA VIVIESCAS MALDONADO

UNIVERSIDADE ESTADUAL PAULISTA
"Júlio de Mesquita Filho"
CAMPUS DE BOTUCATU
INSTITUTO DE BIOCÊNCIAS DE BOTUCATU

**ANÁLISE ESTRUTURAL DO COMPONENTE
TELOMERASE TRANSCRIPTASE REVERSA DE *Leishmania*
*major***

MARIA ALEJANDRA VIVIESCAS MALDONADO

ORIENTADORA: MARIA ISABEL NOGUEIRA CANO

CO-ORIENTADOR: CARLOS ALEXANDRE HENRIQUE FERNANDES

Tese apresentada ao Instituto de
Biotecnologia, *Campus* de Botucatu,
UNESP, para obtenção do título de
Doutor no Programa de Pós-
Graduação em Ciências Biológicas
(Genética).

DEDICATÓRIA

Aos meus pais, Victor e María Mercedes, porque seu envolvimento na área acadêmica me fez sonhar desde pequena com obter um dia o título de doutora.

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP

BIBLIOTECÁRIA RESPONSÁVEL: ROSANGELA APARECIDA LOBO-CRB 8/7500

Maldonado, Maria Alejandra Viviescas.

Análise estrutural do componente telomerase transcritase reversa de Leishmania major. / Maria Alejandra Viviescas Maldonado. - Botucatu, 2018

Tese (doutorado) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências de Botucatu

Orientador: Maria Isabel Nogueira Cano

Coorientador: Carlos Alexandre Henrique Fernandes

Capes: 20804008

1. Leishmania. 2. Telômero. 3. Telomerase. 4. Bioinformática.
5. Proteínas - Análise.

Palavras-chave: Leishmania; bioinformatica; estrutura;
telomerase; telômeros.

AGREDECIMENTOS

Aos meus pais, Victor Viviescas e María Mercedes Maldonado e minha irmã Mariana Viviescas porque sem eles, eu não seria a pessoa que sou hoje.

Ao meu parceiro, Ariel Barcellos, pela companhia e apoio nos momentos difíceis e por me emprestar seu computador com sistema operativo Linux, onde o PyMol roda muito mais rápido. Também a Vera, Otto e Davi, por terem me acolhido e acompanhado durante o doutorado.

Aos membros da minha família, especialmente minha prima Luisa Riviere, pelo seu apoio e amor durante todos os momentos da minha vida.

À Profa. Dra. Maria Isabel Nogueira Cano, por me acolher no seu grupo de pesquisa, me permitir desenvolver minha pesquisa no seu laboratório e pela sua orientação durante o doutorado.

Ao Dr. Carlos Alexandre Henrique Fernandes, pela sua co-orientação durante o doutorado e por me ensinar tudo o que eu sei sobre biologia estrutural.

Ao Prof. Dr. Marcos Roberto de Mattos Fontes, por me permitir fazer uso do seu laboratório e equipamentos.

Aos integrantes e ex-integrantes do Laboratório de Telômeros pelo apoio, companheirismo e convívio: Marcelo Santos da Silva, Marcela Segatto, Edna Gicela Ortiz Morea, Marina Roveri, Whisnayder Gentil, Alessandro Vassilievitch, Gabriel Arantes e Camila Storti. Um agradecimento especial a Cristiane de Santis pela força no meu português e aos meus estudantes de iniciação científica, Vitor Da Silva e Stephany Paiva, por me permitirem transmitir meu conhecimento, pelo apoio com alguns experimentos e por levar este projeto adiante.

Aos integrantes e ex-integrantes do Laboratório de Biologia Molecular Estrutural, por me fazerem sentir como parte do grupo, Guilherme Salvador, Fábio Cardoso, Andrea Barros, Lino de Lima, e especialmente ao Fabio Matioli pela sua ajuda na geração dos modelos tridimensionais.

O presente trabalho foi realizado com apoio do CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico, nº do processo 161947/2014-0

À FAPESP pelo apoio geral ao meu Projeto de Doutorado que esteve vinculado ao Projeto Temático 2012/50161-8

Ao programa de Pós-graduação de Ciências Biológicas Área de Concentração Genética.

Ao Instituto de Biociências de Botucatu da Universidade Estadual Paulista “Julho de Mesquita Filho” - UNESP.

“Eu faço parte dos pensam que a Ciência é belíssima.
Um cientista em um laboratório não é apenas um técnico,
ele é também uma criança diante de fenômenos
naturais que o impressionam como um conto de fadas.”

Marie Curie

SUMÁRIO

Lista de Figuras	iii
Lista de Tabelas.....	v
Resumo.....	vi
Abstract	vii
LISTA DE ABREVIATURAS E SIGLAS	viii
1. INTRODUÇÃO	1
1.1 Parasitos do gênero <i>Leishmania</i> e Leishmaniose	1
1.2. Telômeros e telomerase.....	5
1.3. Telômeros como alvo terapêutico	12
1.4 Estrutura do complexo telomerase	13
1.4.1. TERT, o componente proteico da telomerase	15
1.4.2 O componente RNA (TER) da telomerase.....	29
1.5 Telômeros de <i>Leishmania</i> sp.....	32
1.5.1 Complexo telomérico de <i>Leishmania</i> sp.	32
1.5.2 Características do complexo telomerase em <i>Leishmania</i> sp.	33
2. OBJETIVOS	36
3. MATERIAIS E MÉTODOS	37
3.1 Alinhamento de sequências e predição da estrutura secundária.....	37
3.2 Modelagem molecular e simulações de dinâmica molecular	37
3.3 <i>Docking</i> Molecular.....	38
3.4 Estudos filogenéticos.....	39
3.5 Desenho e construção dos vetores de expressão com os domínios TEN, RT e TRB da telomerase de <i>L. major</i>	40
3.6 Testes de mini-indução de proteína heteróloga em sistema bacteriano	42
3.7 Purificação do domínio TEN a partir da fração insolúvel do extrato proteico.....	43
3.8 Purificação do domínio TEN a partir da fração solúvel do extrato proteico.....	43
3.9 Re-enovelamento do domínio TEN purificado a partir da fração insolúvel.....	44
3.10 Recuperação de proteínas recombinantes contidas em corpúsculos de inclusão	45
3.11 <i>Western blot</i>	45
3.12 Formação de duplex parcial de oligonucleotídeos para ensaio de CD e fluorescência.....	46
3.13 Espectroscopia de dicroísmo circular (CD).....	46
3.14 Espectroscopia de fluorescência estática.....	47
3.15 Cromatografia analítica de exclusão molecular	47
3.16 Transformação de células competentes de <i>Pichia pastoris</i> KM71H.....	48
3.17 Teste de mini-indução em leveduras	48

4. Resultados e discussão	50
4.1. O componente TERT de <i>Leishmania</i> é o mais basal entre os eucariotos.....	50
4.2 Caracterização estrutural do domínio TEN do TERT de <i>Leishmania</i> sp.	53
4.3 Obtenção do domínio TEN recombinante e estudos estruturais com este domínio	59
4.4 Caracterização estrutural do domínio TRB do TERT de <i>Leishmania</i> sp.	75
4.4.1 Análise estrutural do motivo T em <i>L. major</i>	80
4.4.2 Análise estrutural do motivo CP em <i>L. major</i>	85
4.4.3 Análise estrutural do mecanismo de interação entre o domínio TRB e o componente TER em <i>L. major</i>	88
4.5 Expressão do domínio TRB de <i>L. major</i> em modelo heterólogo	102
4.6 Caracterização do domínio RT de <i>Leishmania</i> sp.	104
4.6.1 Análise dos motivos IFD e 3	106
4.6.2 Caracterização estrutural da fração catalítica do TERT de <i>L. major</i>	107
4.6.3 Caracterização estrutural do domínio RT de <i>L. major</i>	108
4.7 Obtenção do domínio RT recombinante	113
5. CONCLUSÕES E PERSPECTIVAS.....	117
6. Bibliografia	121
ANEXOS.....	129
Anexo 1: Fragmentos clonados nos vetores de expressão	129
Anexo 2: Sequências utilizadas para se fazer o cladograma	130
Anexo 3: Alinhamento múltiplo de sequência entre diferentes espécies de <i>Leishmania</i> sp.	135
Anexo 4: Alinhamento de sequência múltiplo entre diferentes Trypanossomatídeos.....	137
APÊNDICES.....	139
Apêndice 1: Capítulo de livro: “NUCLEAR AND KINETOPLAST DNA REPLICATION IN TRYPANOSOMATIDS”	139
Apêndice 2: Artigo publicado em BBA-General Subjects: “A calmodulin-like protein (LCALA) is a new <i>Leishmania amazonensis</i> candidate for telomere end-binding protein”.....	140
Apêndice 3: Artigo publicado em Biochemie: “Structural studies with BnSP-7 reveal an atypical oligomeric conformation compared to phospholipases α -like toxins”	141
Apêndice 4: Artigo aceito para publicação em Current Proteomics “Chaperones and Their Role in Telomerase Ribonucleoprotein Biogenesis and Telomere Maintenance”.....	141

LISTA DE FIGURAS

Figura 1. Ciclo de desenvolvimento de <i>Leishmania</i> sp.....	2
Figura 2. Conformações do DNA telomérico	7
Figura 3. O ciclo catalítico da telomerase	9
Figura 4. Representação esquemática das holoenzimas telomerasas em diferentes organismos	11
Figura 5. Esquema estrutural das telomerasas.....	18
Figura 6. Representação estrutural do domínio TEN do TERT	22
Figura 7. Representação estrutural do domínio TRB do TERT	25
Figura 8. Representação estrutural domínio RT.....	27
Figura 9. Representação estrutural do domínio CTE	29
Figura 10. Elementos estruturais conservados do componente TER de diferentes grupos de organismos	31
Figura 11. Panorama atual da composição da cromatina telomérica de <i>Leishmania</i> sp.	33
Figura 12. Predição da estrutura secundária de TER em <i>L. major</i>	35
Figura 13. Mapa do vetor de expressão contendo o domínio TEN de <i>L. major</i>	41
Figura 14. Mapa do vetor de expressão contendo o domínio RT de <i>L. major</i>	41
Figura 15. Mapa do vetor de expressão contendo o domínio TRBD de <i>L. major</i>	42
Figura 16. Cladograma dos componentes TERT de diferentes eucariotos	50
Figura 17. Delimitação dos domínios estruturais na estrutura primária de <i>L. major</i>	53
Figura 18. Alinhamento múltiplo de sequências entre os domínios TEN dos TERT de <i>Leishmania</i> e mamíferos	55
Figura 19. Predição de estrutura secundária do domínio TEN de <i>L. major</i>	58
Figura 20. Teste de mini-indução de bactérias <i>E. coli</i> BL21 DE3 códon plus transformadas com pET28a-domínio TEN.....	59
Figura 21. Purificação do domínio TEN do TERT de <i>L. major</i> a partir da fração insolúvel do extrato proteico de bactérias <i>E. coli</i> DE3 códon plus, por cromatografia de troca aniônica.....	61
Figura 22. Avaliação do estado de enovelamento do domínio TEN purificado	63
Figura 23. <i>Western blot</i> confirmando a purificação do domínio TEN recombinante	64
Figura 24. Espectros de Dicroísmo circular e fluorescência estática do domínio TEN purificado na presença de DNA telomérico.....	66
Figura 25. Cromatografia analítica de exclusão por tamanho do domínio TEN.....	68
Figura 26. Comprovação de contaminação com DNA bacteriano em amostras proteicas contendo o domínio TEN.	69
Figura 27. Tratamento do pellet bacteriano com N-Lauryl sarcosine durante a lise.....	71
Figura 28. Tentativa de purificação do domínio TEN expresso a partir da fração solúvel do extrato proteico mediante cromatografia de troca iônica	72

Figura 29. Tentativa de purificação do domínio TEN expresso a partir da fração solúvel do extrato proteico mediante cromatografia de afinidade	73
Figura 30. <i>Western blot</i> comprovando que o domínio TEN foi expresso na fração solúvel do extrato proteico.....	74
Figura 31. Alinhamento de sequências do domínio TRB e demarcação da posição dos motivos	77
Figura 32. Predição da estrutura terciária do domínio TRB de <i>L. major</i>	79
Figura 33. Modificações no motivo T causadas pelas substituições presentes em <i>L. major</i>	82
Figura 34. Detalhe da pinça hidrofóbica no motivo T	84
Figura 35. Detalhes da interação entre o motivo CP e o motivo T	86
Figura 36. Detalhe dos motivos T e CP	88
Figura 37. Resíduos do modelo tridimensional do TRBD de <i>L. major</i> escolhidos como alvo para <i>docking</i> molecular.	89
Figura 38. Detalhe da região de interação entre o componente TBE e o domínio TRB	93
Figura 39. Análise de interações entre o domínio TRB de <i>T. thermophila</i> e a fita A do elemento TBE	95
Figura 40. Análise de interações entre o domínio TRB de <i>T. thermophila</i> e a fita B do elemento TBE	96
Figura 41. Análise de interações entre o domínio TRB de <i>L. major</i> e a fita A do elemento TBE	97
Figura 42. Análise de interações entre o domínio TRB de <i>L. major</i> e a fita B do elemento TBE	98
Figura 43. Motivos CP2 e TFLY	101
Figura 44. <i>Western blot</i> da mini indução do domínio TRB de <i>L. major</i> em <i>P. pastoris</i>	103
Figura 45. Alinhamento de sequência múltiplo do domínio RT	105
Figura 46. Comparação das superfícies da telomerase de <i>T. castaneum</i> e de <i>L. major</i>	108
Figura 47. Estrutura terciária do domínio RT de <i>L. major</i>	109
Figura 48. Conservação estrutural dos motivos do domínio RT em <i>L. major</i>	110
Figura 49. Conservação estrutural do sitio ativo do TERT em <i>L. major</i>	112
Figura 50. Teste de expressão e de solubilidade do domínio RT de <i>L. major</i> usando <i>E. coli</i> BL21 Artic Express pRIL	114
Figura 51. Purificação do domínio RT mediante isolamento de corpúsculos de inclusão.....	115
Figura 52. <i>Western blot</i> do domínio RT do TERT de <i>L. major</i> após purificação por isolamento de corpúsculos de inclusão	116
Figura 53. Fluorescência estática do domínio RT de <i>L. major</i> re-enovelado	117

LISTA DE TABELAS

Tabela 1. Oligonucleotídeos utilizados neste trabalho	49
Tabela 2. Comparação entre as sequências primárias dos componentes TERT de protozoários dos gêneros <i>Leishmania</i> e <i>Trypanosoma</i>	52
Tabela 3. Homologia estrutural entre o domínio TEN de LmTERT com estruturas cristalográficas depositados em PDB	56
Tabela 4. Resíduos conservados no motivo T de diferentes organismos, o asterisco denota que o resíduo não é conservado em <i>L. major</i> quando comparado com os eucariotos derivados.....	81
Tabela 5. Valores obtidos para cada um dos alvos após o <i>Docking</i> molecular.....	91
Tabela 6. Resumos das interações hidrostáticas e ligações de hidrogênio entre o componente TBE e o domínio TBD em <i>T. thermophila</i> e <i>L. major</i>	99
Tabela 7. Resíduos do motivo 3 importantes para a processividade do TERT humanos, o asterisco denota um resíduo não conservado	106

RESUMO

Os parasitos do gênero *Leishmania* são protozoários primitivos entre os quais estão os causadores de um espectro de doenças conhecidas como leishmaniose, as quais afetam milhões de pessoas no mundo inteiro, sendo o Brasil um dos países que apresenta maior número de casos por ano. Os tratamentos e vacinas disponíveis para a leishmaniose apresentam problemas como toxicidade, alto custo e baixa eficiência, tornando importante a busca por novos alvos terapêuticos. Dada sua importância para a estabilidade genômica e proliferação celular, os telômeros têm sido considerados como alvos terapêuticos potenciais no tratamento da leishmaniose. Os telômeros são as extremidades físicas dos cromossomos lineares compostos por complexos ribonucleoproteicos que envolvem a interação entre proteínas, DNA e RNA. A maioria dos eucariotos mantém o tamanho dos telômeros pela ação do complexo telomerase, que é minimamente composto por uma proteína (TERT, *Telomerase Reverse Transcriptase*) e um RNA longo não codificador (TER, *Telomerase RNA*). Os dois componentes do complexo telomerase em *Leishmania* sp. foram identificados, porém não há informação disponível sobre suas estruturas. O objetivo principal deste estudo foi a caracterização estrutural do componente TERT de *Leishmania major*. Utilizando alinhamentos múltiplos de sequências, foi possível verificar que os quatro domínios estruturais das telomerases canônicas estão presentes no componente TERT em *Leishmania* sp. Três destes domínios foram estudados independentemente utilizando várias ferramentas de bioinformática. O domínio TEN (*Telomerase essential N-terminal*) é o mais amino-terminal e é importante na interação com o DNA telomérico e proteínas teloméricas. Em *L. major*, este domínio mostra uma baixa conservação de estrutura primária quando comparado com outros eucariotos, mas preserva o único motivo estrutural. Este domínio está composto principalmente por elementos desordenados e hélices- α e mostra uma baixa conservação de estrutura primária quando comparado com outros eucariotos, mas preserva seu motivo funcional característico. O domínio TRB (*Telomerase RNA Binding*) é importante na interação entre o componente TERT e o componente TER. Em *L. major*, o domínio TRB apresenta uma conservação estrutural parcial, mas possui substituições em resíduos de aminoácidos próprias do gênero *Leishmania* que poderiam alterar o mecanismo de interação TERT/TER. O domínio RT é o centro catalítico do TERT. Este domínio é altamente conservado em *L. major* e apresenta conservação de todos os motivos funcionais, assim como do centro catalítico. Para corroborar os resultados *in silico* com dados experimentais, os três domínios foram individualmente expressos em sistema heterólogo e protocolos de purificação das proteínas recombinantes são apresentados. O objetivo é utilizar estas proteínas em estudos de interação proteínas-ácidos nucleicos e proteína-proteína para verificar o grau de conservação destas interações, uma vez que nesses parasitos o TERT compartilha características estruturais com outros TERT, embora também apresente peculiaridades do gênero. Os resultados gerados neste trabalho podem fornecer conhecimento sobre a evolução da enzima telomerase e indicar se a telomerase de *L. major* poderá ser considerada um alvo terapêutico parasita-específico.

Palavras chave: *Leishmania*, telômeros, telomerase, estrutura, bioinformática.

ABSTRACT

Parasites of the *Leishmania* genus are primitive protozoa and among them are the causative agents of a spectrum of diseases called leishmaniasis, which affect millions of people in tropical countries worldwide, being Brazil one of the countries with higher number of cases each year. The drugs or vaccines available to treat leishmaniasis present problems such as toxicity, excessive cost a low efficiency, so it is important to search for new potential therapeutic targets. Due to their importance in genome stability and cell proliferation, telomeres have been considered a potential therapeutic target against leishmaniasis. Telomeres are the ends of linear chromosomes composed by ribonucleoproteic complexes that involve the interaction between proteins, DNA and RNA. Most eukaryotes maintain their telomere length by the action of the telomerase complex, minimally composed by a protein (TERT, Telomerase Reverse Transcriptase) and a long non-coding RNA (TER, Telomerase RNA). Both components of the *Leishmania* sp. telomerase complex have already been identified, however, little is known about their structure. This study had the aim to characterize the structure of the *Leishmania major* TERT component. Using multiple sequence alignments, we were able to verify that the four structural domains of canonical telomerases are present in *Leishmania* sp. Using different bioinformatic approaches three of these domains were independently studied. The TEN (Telomerase essential N-terminal) domain is the most amino-terminal and is involved in the interaction between TERT, telomeric DNA and accessory proteins. In *L. major* this domain is mainly composed by unordered regions and α helixes showing little primary structure conservation when compared to other eukaryotes, although it preserves its unique structural motif. The TRB (Telomerase RNA binding) domain is important for the interaction between TERT and TER. In *L. major* the TRB domain shows partial structural conservation but presents characteristics that are specific to the genus and that could alter the mechanism of TERT/TER interaction. The RT (Reverse transcriptase) domain is the catalytic center of TERT. It is structurally conserved in *L. major*, presenting all the functional motifs and conservation of the catalytic core. To corroborate our *in-silico* results with experimental data, we independently expressed all three domains using heterologous system and purification protocols are presented. Our goal is to perform protein-protein and protein-nucleic acid interaction assays to verify how conserved are these interactions. In these protozoa, TERT shares conserved structural features with other TERT but also present its own peculiarities. The results gathered from these studies will provide insights on the evolution of the telomerase enzyme and will help to determine if *L. major* telomerase could be considered a parasite-specific therapeutic target.

Keywords: *Leishmania*, telomeres, telomerase, structure, bioinformatic

LISTA DE ABREVIATURAS E SIGLAS

°C	grau centígrado
Å	Angstrom
<i>g</i>	Força gravitacional
K	Graus Kelvin
kDa	KiloDaltons
M	Molar
mg	Miligramas
ml	Mililitro
pH	Potencial de hidrogênios
ps	Picosegundos
rpm	Rotações por minuto
μ	micro

CD	Dicroísmo circular
CTE	<i>C-terminal extension</i>
DNA	Ácido desoxirribonucleico
et al	E colaboradores
FISH	<i>Fluorescence in situ hybridization</i>
HIV	Vírus da Imunodeficiência Humana
IPTG	Isopropyl-B-D-Galactoside
LaTERT	TERT de <i>Leishmania amazonensis</i>
LCTAS	<i>Leishmania conserved telomere associated sequences</i>
NMR	Ressonância magnética nuclear
PBD	<i>Protein Data Bank</i>
POT1	<i>Protection of Telomeres 1</i>
RAP1	<i>Repressor-Activator protein 1</i>
RNA	Ácido ribonucleico
RT	<i>Reverse transcriptase</i>
TBE	<i>template boundary element</i>
TEN	<i>Telomerase essential N-terminal</i>
TER	<i>Telomerase RNA</i>
TERT	<i>Telomerase reverse transcriptase</i>

TIN2	<i>TRF1-interacting protein 2</i>
TPP1	<i>TIN2 e POT1 Interacting Factor</i>
TRB	<i>Telomerase RNA binding</i>
TRF1	<i>Telomeric Repeat Factor 1</i>
TRF2	<i>Telomeric Repeat Factor 2</i>

1. INTRODUÇÃO

1.1 Parasitos do gênero *Leishmania* e Leishmaniose

O gênero *Leishmania* pertence à ordem Kinetoplastida, família Trypanosomatidae, classificados entre os organismos eucariotos mais primitivos do planeta, tendo divergido de outros eucariotos cerca de 200-500 milhões de anos atrás (SIMPSON et al., 2004; SILVA, DA et al., 2017). O gênero *Leishmania* é composto por parasitos digenéticos com ciclo de vida heteroxeno, apresentando duas formas de vida principais presentes em dois hospedeiros (WORLD HEALTH ORGANIZATION, 2012; ALVAR et al., 2012). A primeira é denominada de promastigota e se apresenta como uma forma alongada, flagelada e móvel que habita o trato digestivo de insetos dípteros, dos gêneros *Phlebotomus* no velho mundo e *Lutzomia* no novo mundo (WORLD HEALTH ORGANIZATION, 2012). A segunda forma de vida é denominada amastigota e se apresenta arredondada, com flagelo inaparente e habita o interior das células da linhagem monocítica-fagocitária do hospedeiro vertebrado (CHAPPUIS et al., 2007).

O ciclo de desenvolvimento de *Leishmania* sp. (Figura 1) se inicia quando fêmeas infectadas do inseto vetor se alimentam com o sangue do hospedeiro mamífero, liberando promastigotas metacíclicos na sua corrente sanguínea (CHAPPUIS et al., 2007; SILVA, DA et al., 2013). Os promastigotas são fagocitados por macrófagos e células dendríticas. No interior da célula, o promastigota se transforma em amastigota passando a sobreviver dentro dos fagolisossomos. Os amastigotas se dividem por fissão binária até causar o rompimento da célula, e assim, são liberados na corrente sanguínea infectando novas células (CHAPPUIS et al., 2007). O ciclo é fechado quando um inseto vetor, não infectado, se alimenta de um hospedeiro mamífero infectado ingerindo

células contendo as formas amastigotas. No trato digestivo do inseto vetor, os amastigotas se transformam em promastigotas procíclicos, uma forma altamente proliferativa, mas não infectiva, que ao longo do trato digestivo do inseto se transformam em parasitos não proliferativos mas altamente infectivos, denominados promastigotas metacíclicos, que migram para a probóscide do inseto vetor (Da SILVA, et al., 2013).

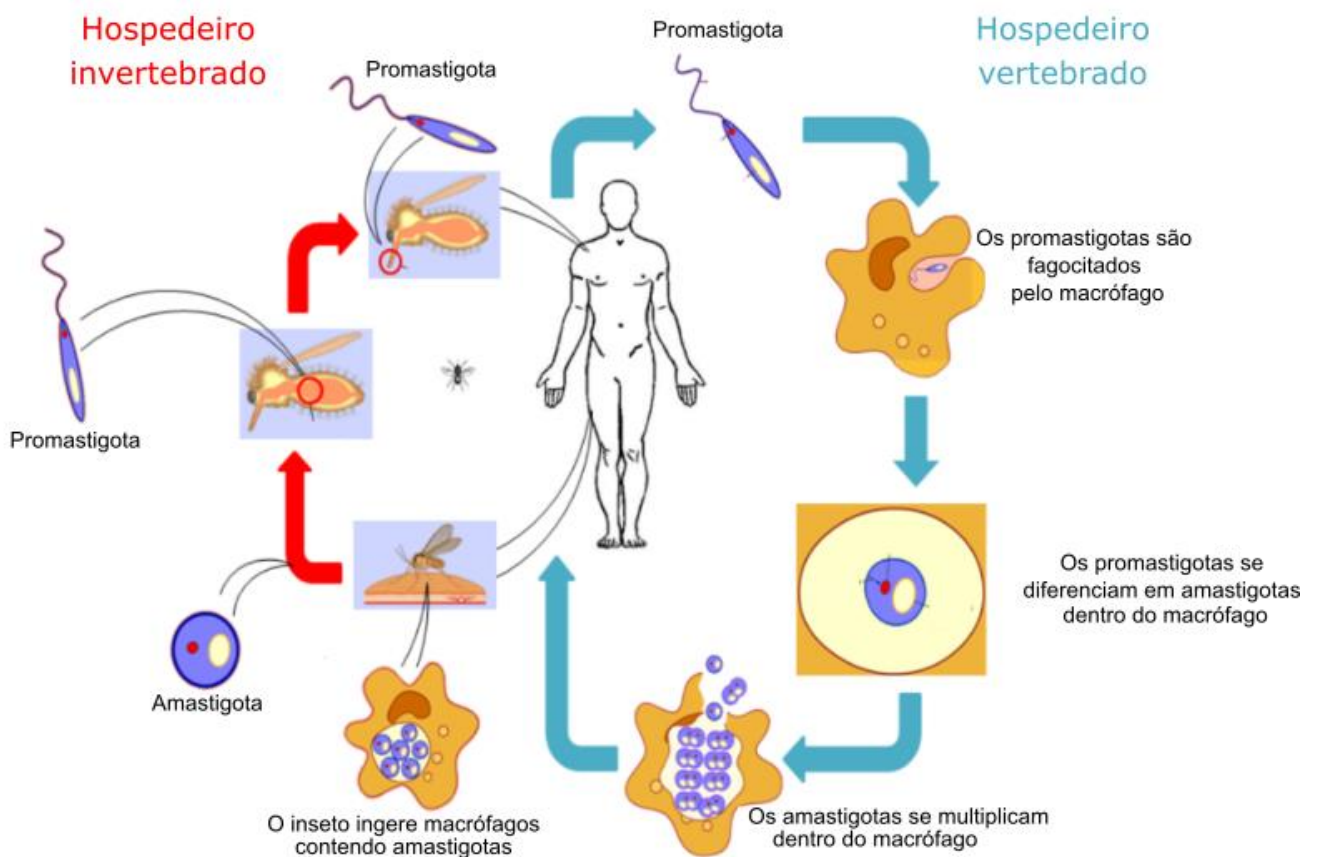


Figura 1. Ciclo de desenvolvimento de *Leishmania* sp.

Representação esquemática do ciclo de desenvolvimento dos parasitos do gênero *Leishmania* no hospedeiro vertebrado e no hospedeiro invertebrado. A figura representa os diferentes estágios de desenvolvimento do parasita durante seu passo pelo dos hospedeiros, o mamífero e o invertebrado figura traduzida de Da SILVA, et al., (2013).

O gênero *Leishmania* é composto por aproximadamente 53 espécies, das quais 31 são parasitos de mamíferos e aproximadamente 20 espécies dos subgêneros *Leishmania* e *Viannia*, infectam humanos (AKHOUNDI et al., 2016).

Os parasitos do gênero *Leishmania* são causadores de um conjunto de doenças conhecidas como leishmanioses que apresentam diversas formas clínicas (WORLD HEALTH ORGANIZATION, 2012). A leishmaniose cutânea (LC) consiste em lesões na pele que variam em severidade e duração, podendo ser auto curativas. A leishmaniose mucocutânea (LMC) é uma doença desfigurante e mutilante que causa úlceras nas mucosas do nariz, boca, laringe e faringe. Estas úlceras não curam espontaneamente. Finalmente, a forma mais severa de leishmaniose é a leishmaniose visceral (LV) ou *kala-azar*, onde os parasitos infectam órgãos internos como o fígado e o pâncreas (CHAPPUIS et al., 2007). A forma clínica e a intensidade da doença estão associadas à espécie infectante e ao estado imunológico do hospedeiro mamífero (CHAPPUIS et al., 2007; WORLD HEALTH ORGANIZATION, 2012).

A organização mundial da saúde considera as leishmanioses doenças de importância em saúde pública, sendo classificadas entre as doenças tropicais negligenciadas, para as quais não existem métodos profiláticos eficazes nem tratamento adequado, além de baixo investimento por parte das indústrias farmacêuticas, apesar do grande número de vítimas acometidas pela doença. As leishmanioses ameaçam 350 milhões de pessoas em 98 países, e estima-se que existem atualmente 12 milhões de pessoas infectadas (WORLD HEALTH ORGANIZATION, 2012; AKHOUNDI et al., 2016). Estima-se que todos os anos ocorrem entre 700 mil e um milhão de novos casos de leishmaniose cutânea e, entre 200 e 400 mil novos casos de leishmaniose visceral (AKHOUNDI et al., 2016), dos quais entre 20 e 30 mil levam à morte do paciente (VIJAYAKUMAR; DAS, 2018)(<http://www.who.int/mediacentre/factsheets/fs375/en/>). Aproximadamente 90% dos casos ocorrem na Índia, em Bangladesh, no Brasil, no Nepal e no Sudão (WORLD HEALTH ORGANIZATION, 2012).

A primeira opção de tratamento das leishmanioses na maioria de países são os antimoniais pentavalentes (BERMAN, 2005; CROFT et al., 2006; CHAPPUIS et al., 2007; WORLD HEALTH ORGANIZATION, 2012), mesmo quando estes fármacos apresentam problemas de toxicidade e induzem resistência nos parasitos (CROFT et al., 2006; CHAPPUIS et al., 2007; WORLD HEALTH ORGANIZATION, 2012). A segunda linha de tratamento é a Anfotericina B, o fármaco de escolha de muitos países graças à sua alta eficácia e poucos efeitos secundários. No entanto, seu alto custo e a falta de cobertura dos planos de saúde não permitem sua utilização por países em desenvolvimento (CHAPPUIS et al., 2007; WORLD HEALTH ORGANIZATION, 2012). A miltefosina é o único fármaco anti-leishmaniose oral disponível atualmente, mas possui maior eficácia contra a forma visceral, além de ser teratogênico, o que impede sua aplicação em toda a população (CHAPPUIS et al., 2007). Os protocolos de tratamentos anti-leishmaniose também podem não ser eficazes porque os pacientes não terminam o tratamento e, também, porque seu estado imunológico afeta sua resposta aos diferentes fármacos, somando um grau de complexidade ao tratamento (CROFT et al., 2006).

Com o intuito de se estabelecerem novos e mais efetivos protocolos no tratamento da leishmaniose, componentes de várias rotas metabólicas, que sejam de importância na forma amastigota do parasita e não sejam compartilhadas com o humano têm sido avaliadas como possíveis alvos terapêuticos (VIJAYAKUMAR; DAS, 2018). Do mesmo jeito, esforços têm sido encaminhados ao desenvolvimento de uma vacina que permita prevenir novos casos de leishmaniose. No entanto, todos os tratamentos disponíveis até a data têm, implicações negativas, por isso, mais pesquisas em buscas de novos alvos terapêuticos são necessárias (DIDWANIA et al., 2017; GHORBANI; FARHOUDI, 2018; VIJAYAKUMAR; DAS, 2018). Também é importante o

fortalecimento das medidas de prevenção da doença como o controle de insetos vetores, mediante inseticidas ou barreiras, e o controle de reservatórios mamíferos, como cães e gado. Entretanto, estes programas são de difícil manutenção devido à irregularidade de investimentos em profilaxia nos países em desenvolvimento (CHAPPUIS et al., 2007; WORLD HEALTH ORGANIZATION, 2012), ainda que estudos recentes ressaltaram que o investimento na prevenção e controle em doenças negligenciadas têm o potencial de gerar um grande retorno econômico, graças ao maior bem-estar da população (LENK et al., 2018).

Deste modo, a falta de fármacos e de estratégias de prevenção eficazes evidenciam a necessidade de novos protocolos terapêuticos que sejam de baixo custo e seguros para os pacientes. Recentemente, a disponibilidade de novas informações sobre a fisiologia dos parasitos e a disponibilidade das sequências dos genomas de diferentes espécies de *Leishmania* podem possibilitar a identificação de novos alvos terapêuticos.

Neste contexto, os telômeros têm sido considerados como alvos terapêuticos potenciais dada sua relação direta com a estabilidade do genoma e a proliferação celular (SEKARAN et al., 2014).

1.2. Telômeros e telomerase

Os telômeros são as extremidades dos cromossomos lineares dos eucariotos. São estruturas ribonucleoproteicas cuja principal função é a de proteger os terminais cromossômicos, evitando que sejam reconhecidos como quebras de DNA na forma de dupla fita e de fusões terminais (BLACKBURN, 1994). Eles permitem a replicação completa do DNA e evitam a perda de material genético codificante durante as divisões celulares. Por isso, a homeostase telomérica tem sido implicada em processos como ciclo celular, envelhecimento e câncer (GIARDINI et al., 2014).

O DNA telomérico é composto por sequências repetitivas curtas, que são similares mesmo entre espécies altamente divergentes (BLACKBURN; GALL, 1978; GREIDER; BLACKBURN, 1985; CANO, 2001), podendo variar em tamanho dependendo do organismo (WU et al., 2017). As repetições teloméricas são formadas por uma fita rica em guaninas e a outra rica em citosinas. Na extremidade 3' da fita rica em guaninas (G) uma protrusão fita simples denominada "*3' G-Overhang*" é formada (Figura 2) (GIARDINI et al., 2014). O DNA telomérico também pode assumir estruturas secundárias, como é o caso do *T-loop*, uma conformação na qual o *3' G-Overhang* invade a região dupla fita formando um *loop* (LANGE, DE, 2004). Esta conformação é de grande importância para evitar que os telômeros sejam reconhecidos como quebras no DNA. Uma segunda conformação comum do DNA telomérico é o *3' G-Overhang*, um arranjo que é possível dado o alto conteúdo de guaninas presente no DNA telomérico (GRIFFITH et al., 1999). A presença destas conformações no DNA genômico depende da interação com proteínas específicas. Estas conformações na cromatina telomérica evitam que os telômeros sejam reconhecidos pela maquinaria de reparo celular como quebras no DNA, garantindo assim, a estabilidade genômica (GIARDINI et al., 2014).

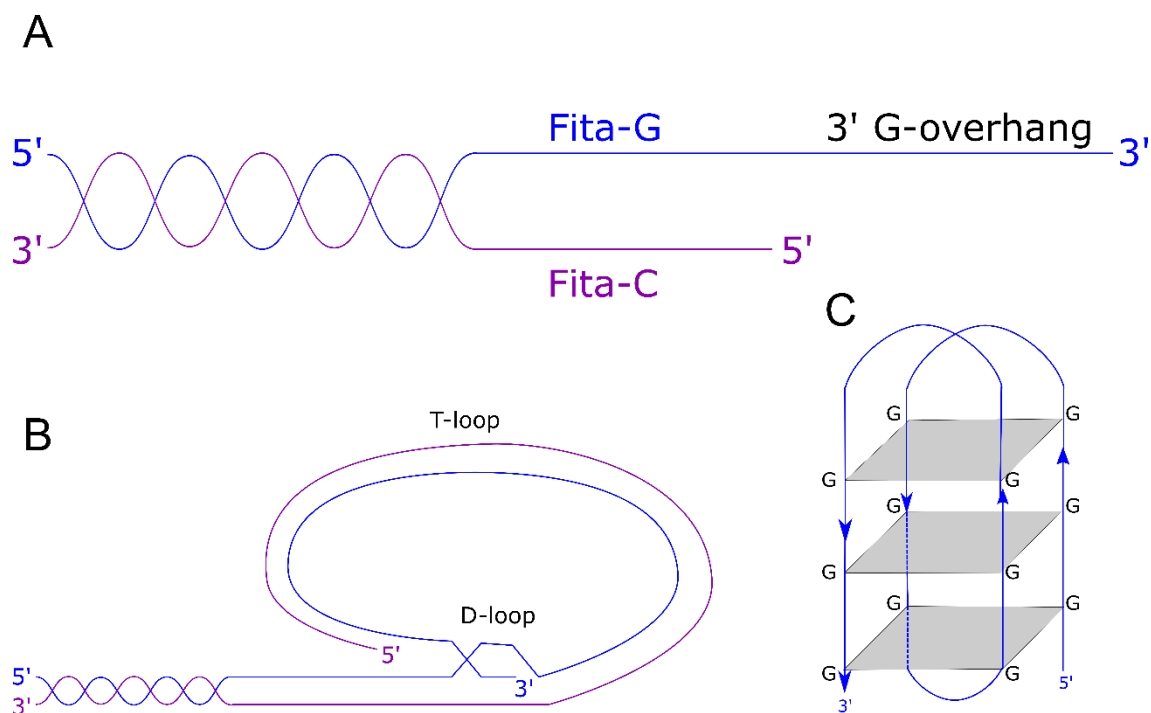


Figura 2. Conformações do DNA telomérico

(A) Estrutura primária do DNA telomérico mostrando a região dupla fita e a protrusão da fita simples rica em G ou 3' *G-Overhang*. (B) *T-loop*, conformação especial na qual o 3' *G-Overhang* invade a região dupla fita formando um *loop*. Esta conformação é de grande importância para a proteção telomérica e para evitar que os telômeros sejam reconhecidos como quebras de DNA pela maquinaria de reparo (C) O alto conteúdo de guaninas do DNA telomérico permite que o 3' *G-Overhang* forme arranjos na forma de estruturas conhecidas como *G-quadruplex*. Figura traduzida de VIVIESCAS et al., (2018).

Os conjuntos de proteínas encarregados de proteger o DNA telomérico recebem o nome de complexo de *capping*. Nos seres humanos, este complexo é denominado *shelterin*, sendo composto por 6 proteínas: TRF1 e TRF2 (*Telomeric Repeat Factor 1 e 2*), TIN2 (*TRF1-interacting protein 2*), RAP1 (*Repressor-Activator protein 1*), POT1 (*Protection of Telomeres 1*), e TPP1 (*TIN2 e POT1 Interacting Factor*) (COHEN; BLACKBURN, 1998; DMITRIEV et al., 2003; LANGE, DE, 2005; SCHMUTZ; LANGE, DE, 2016; WU et al., 2017). Os complexos de *capping* variam em complexidade e composição dependendo do organismo, mas todos possuem a função de manutenção da homeostase dos telômeros.

Devido ao fato da replicação do DNA ser semiconservativa, os cromossomos lineares eucariotos perdem uma fração de DNA genômico a cada ciclo de replicação,

caracterizando um fenômeno conhecido como “*the end replication problem*” (WATSON, 1972). A presença de regiões repetitivas não codificantes nas extremidades cromossômicas evita a perda de material genético codificante durante a replicação celular (LANGE, DE, 2009).

Além disso, os organismos contam com mecanismos que permitem a elongação dos telômeros. Na maioria dos casos, os telômeros são elongados pela ação do complexo enzimático telomerase. A telomerase possui dois componentes principais, um componente proteico e um componente RNA (*Telomerase RNA*, TER). O componente proteico é denominado *telomerase reverse transcriptase* (TERT), também chamado de subunidade catalítica, que funciona como uma enzima especializada, que utiliza uma sequência molde contida no componente RNA para adicionar repetições ao terminal 3'-OH dos telômeros (Figura 3). A telomerase é capaz de adicionar diversas repetições teloméricas a um mesmo terminal cromossômico. Este fenômeno recebe o nome de processividade na adição de repetições (RAP).

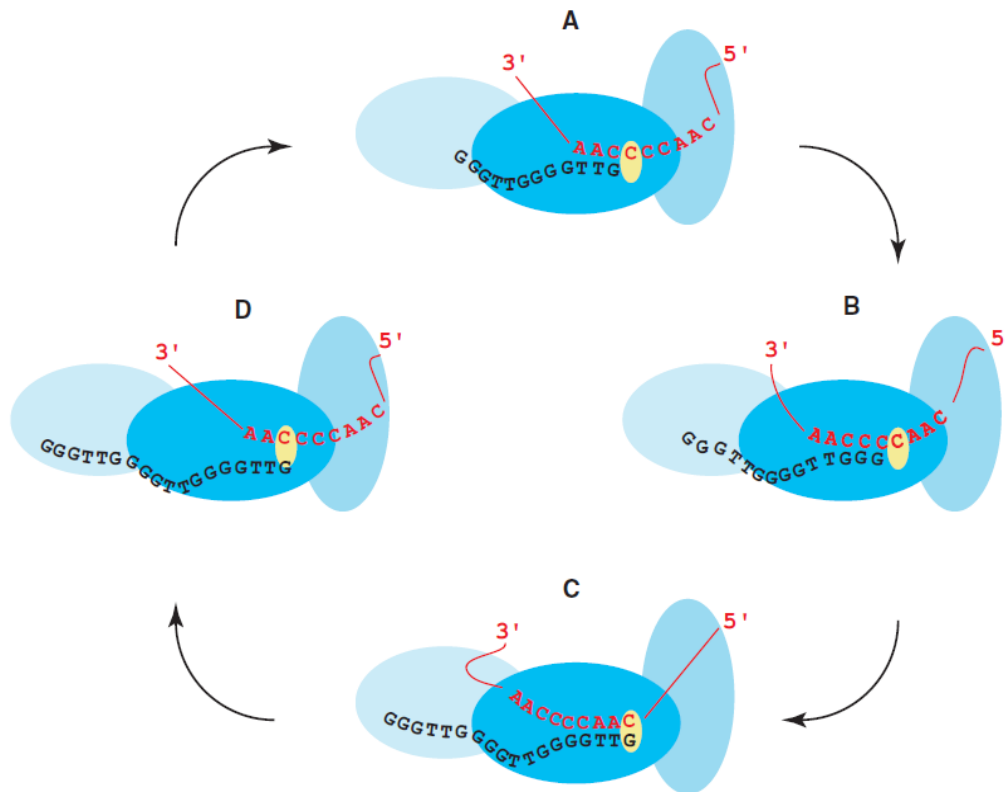


Figura 3. O ciclo catalítico da telomerase

(A). O molde do componente RNA (vermelho) formado por uma região correspondente ao reverso complemento de uma repetição telomérica e meia, e o DNA genômico (preto) são posicionados no sítio ativo da enzima (amarelo). (B). Parte da sequência contida no molde, correspondente ao reverso complemento de uma repetição telomérica, é retrotranscrita e adicionada ao DNA telomérico. (C). A adição de nucleotídeos ao terminal 3'-OH dos telômeros pela telomerase é conhecida como adição processiva de nucleotídeos (NAP). (D). Ao término da adição de uma repetição, o sítio ativo do TERT se dissocia e se transloca, reposicionando o molde do componente RNA na repetição telomérica recém-formada, reiniciando assim, a adição de uma nova repetição telomérica. Figura retirada de BLACKBURN; COLLINS, (2011).

A co-expressão dos componentes TERT e TER é suficiente para reconstituir a atividade de telomerase *in vitro* (BEATTIE et al., 1998; COLLINS; GANDHI, 1998; GIARDINI et al., 2014). No entanto, a regulação da telomerase *in vivo* requer que TERT e TER sejam montados na presença de uma grande quantidade de proteínas formando grandes complexos ribonucleoproteicos que compõem a holoenzima ativa; a holoenzima deve então conseguir interagir com as repetições teloméricas no final dos cromossomos, processo este que requer a remodelação e o auxílio dos complexos

capping (HOLT et al., 1999; COLLINS, 2006; WOJTYLA et al., 2011; SCHMUTZ; LANGE, DE, 2016; VIVIESCAS et al., 2018).

Assim, as holoenzimas telomerasas dos diferentes organismos são grandes complexos ribonucleoproteicos que interagem com uma grande variedade de proteínas de forma permanente ou temporária. No entanto, o mecanismo de ação da telomerase é conservado evolutivamente, evidenciado pelo fato de que as proteínas envolvidas com a homeostase telomérica em diferentes organismos podem ser classificadas em 4 grupos principais: as proteínas que interagem com DNA, as proteínas que servem de ponte entre a telomerase e o DNA, as proteínas que estabilizam o TER e as proteínas acessórias (COLLINS, 2006; VIVIESCAS et al., 2018) (Figura 4).

As proteínas que interagem diretamente com o DNA telomérico são encarregadas de remodelar o telômero para permitir a entrada da holoenzima. Muitas destas fazem parte dos complexos de *capping* e têm-se como exemplos deste grupo as proteínas Pot1 em humanos (SCHMIDT; CECH, 2015), Cdc13 em leveduras (MALYAVKO et al., 2014) e Teb1 em *Tetrahymena thermophila* (JIANG et al., 2015).

As proteínas que formam uma ponte entre as proteínas que interagem com o DNA telomérico e o centro catalítico da holoenzima interagem diretamente com os componentes dos complexos de *capping*, assim como com o TERT e são responsáveis pela localização da holoenzima nos telômeros. As proteínas p50 em *T. thermophila*, (JIANG et al., 2015), Est1 em levedura (TAGGART et al., 2002; CHAN et al., 2008) e Tpp1 em humanos (COHEN; BLACKBURN, 1998; LANGE, DE, 2005) fazem parte deste grupo.

As proteínas que estabilizam o TER são de grande importância já que conferem a esta subunidade sua estrutura secundária, que é essencial para a montagem do

complexo, assim como evitam a degradação do TER (COLLINS, 2006). O complexo de proteínas H/ACA, conformado por Dyskerina, NOP10, NHP2 e GAR1 faz a função de estabilizar o TER em humanos (POGACIĆ et al., 2000), mutações em proteínas deste grupo estão associadas à doenças degenerativas (MACNEIL et al., 2016).

As proteínas acessórias, entre as quais cabe destacar um grande número de chaperonas dão estabilidade ao complexo durante sua biogênese, são exemplos destas chaperonas a HSP90 e sua co-chaperona p23 (VIVIESCAS et al., 2018).

A Figura 4 apresenta os complexos ribonucleoproteicos das telomerases mais estudadas de 3 grupos de organismos: ciliados (*T. thermophila*), leveduras de brotamento (*Saccharomyces cerevisiae*) e humanos.

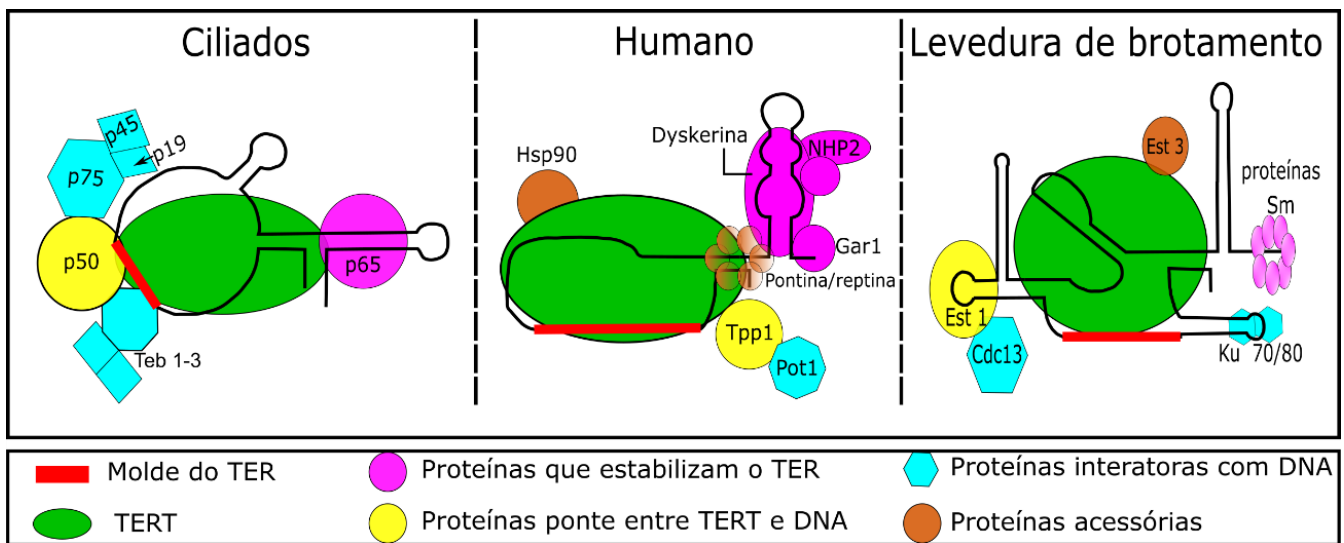


Figura 4. Representação esquemática das holoenzimas telomerasas em diferentes organismos

Ilustração dos complexos ribonucleoprotéicos telomerase ativos de três organismos modelo, o ciliado *Tetrahymena thermophila*, O humano e a levedura de brotamento *Saccharomyces cerevisiae*. Cada um destes organismos apresenta proteínas associadas ao telômero próprias, mas cabe destacar que há uma conservação geral nos grupos de proteínas que estão presentes nos diferentes complexos e por tanto são conservados ao longo da evolução dos eucariotos. Estes grupos de proteínas são: As proteínas que estabilizam o TER, as proteínas que interagem com DNA telomérico, as proteínas que servem de ponte entre o DNA telomérico e o TERT e as proteínas acessórias. Figura traduzida de (VIVIESCAS et al., 2018).

1.3. Telômeros como alvo terapêutico

A homeostase telomérica é vital para o bom funcionamento celular em diferentes organismos (COLLINS, 2006) e é mantida graças à complexa interação de proteínas com ácidos nucleicos e de proteínas entre si, formando complexos de alta ordem nos terminais dos cromossomos. A pesquisa de telômeros é primordial em áreas como desenvolvimento, senescência celular, envelhecimento dos organismos e câncer (ZHANG, 2007).

Na maioria das células somáticas humanas, a expressão da telomerase se encontra desativada, resultando nos processos naturais de senescência celular e envelhecimento (BLASCO, 2005). Entretanto, em 80-90% de células tumorais a expressão da enzima é reativada, o que implica em uma relação direta entre a atividade de telomerase e a ocorrência de câncer, tornando os componentes do complexo telomerase alvos terapêuticos potenciais para tratar este tipo de doença (MERGNY et al., 2002; PHATAK; BURGER, 2009; SEKARAN et al., 2014; BERARDINELLI et al., 2017).

A atividade de telomerase é também fundamental nos organismos unicelulares, os quais dependem desta para a proliferação indefinida da população celular. Isto é especialmente importante no caso de parasitos protozoários que dependem de uma alta atividade proliferativa para poder infectar seus hospedeiros (SILVA, DA et al., 2013; DEY; CHAKRABARTI, 2018). Este fato sugere que, indiretamente, a ação da telomerase pode ser importante nos processos infecciosos. É sabido que diversos genes implicados na evasão do sistema imune em alguns protozoários (ex. *Plasmodium* sp. e *Trypanosoma brucei*) estão localizados na região subtelomérica e que a inativação da telomerase pode levar à perda progressiva dos telômeros (DEY; CHAKRABARTI,

2018). Além disso, em células de *T. brucei*, que não apresentam atividade de telomerase, também ocorre a perda de minicromossomos e rearranjos entre os cromossomos de tamanho intermediário, principalmente entre aqueles que apresentam sítios de expressão subteloamérica transcripcionalmente silenciados (FIGUEIREDO et al., 2002; BARRY et al., 2003; DREESEN; CROSS, 2008). Estas duas observações indicam que os componentes da biologia telomérica de parasitos protozoários podem ser apontados como possíveis alvos terapêuticos no tratamento das doenças ocasionadas por estes organismos.

1.4 Estrutura do complexo telomerase

Com poucas exceções, os eucariotos mantêm o comprimento dos telômeros mediante ação da telomerase, o que sugere que o aparecimento desta enzima ocorreu precocemente durante a evolução do grupo, provavelmente nos ancestrais dos protozoários (AUTEXIER; LUE, 2006; GIARDINI et al., 2014). Entender a estrutura dos complexos telomerase e as interações entre seus constituintes é de grande importância para se compreender o mecanismo de ação da telomerase, tanto em termos gerais, quanto as particularidades de cada organismo.

As primeiras estruturas dos complexos telomerase foram desvendadas em 2013 e correspondem a do ciliado *T. thermophila* e a humana. Estas estruturas foram obtidas a partir da técnica de microscopia eletrônica de partícula única e, possuem baixa resolução (~25 Å) (SAUERWALD et al., 2013; JIANG et al., 2013; SANDIN; RHODES, 2014). A estrutura da holoenzima de *T. thermophila* mostra uma molécula de TERT interagindo com uma molécula de TER na forma de um monômero, que também interage com 6 proteínas acessórias (JIANG et al., 2013). Por sua vez, a estrutura da holoenzima humana parecia estar composta por duas moléculas do componente proteico

TERT, interagindo com duas moléculas do TER (SAUERWALD et al., 2013), sugerindo que o centro catalítico da telomerase em humanos seria um dímero (SAUERWALD et al., 2013; JIANG et al., 2013; SANDIN; RHODES, 2014). No entanto, dados recentes obtidos por criomicroscopia eletrônica a uma resolução de ~8 Å, acerca da estrutura da telomerase humana ligada ao substrato, mostraram um complexo formado por dois lóbulos, o primeiro correspondente aos componentes TERT e TER que interagem com o DNA telomérico e o segundo correspondente a duas cópias do complexo de proteínas H/ACA (dyskerina, NOP10, NHP1 e GAR10), encarregadas de estabilizar o componente TER e uma cópia da proteína de localização TCAB1 (*Telomere Cajal Body protein 1*) encarregada da localização do complexo telomerase nos corpos de cajal em humanos (NGUYEN et al., 2018). Dado que só foi encontrada uma molécula de cada componente (TERT e TER) no centro ativo da telomerase, os autores concluíram que a telomerase humana se comporta como um monômero e não como um dímero. Esta evidência é corroborada por ensaios bioquímicos que mostraram que a formação de dímeros na telomerase humana depende do método de purificação e se dá por um motivo na região amino-terminal do TERT, que quando removido não afeta a atividade enzimática (WU et al., 2015; CHAN et al., 2017; WANG; FEIGON, 2017).

Estudos bioquímicos em *S. cerevisiae*, utilizando as técnicas de imunoprecipitação e FISH (*Fluorescence in situ hybridization*), mostram que o centro catalítico nestes organismos também se comporta como monômero (BAJON et al., 2015). Estes resultados suportam a hipótese atualmente aceita de que a telomerase atua como um monômero, e que este mecanismo é conservado de ciliados até humanos (BAJON et al., 2015; WU et al., 2015; WANG; FEIGON, 2017; NGUYEN et al., 2018).

O pequeno tamanho das proteínas teloméricas de *T. thermophila* comparado aos outros organismos, sua alta atividade de telomerase (GREIDER; BLACKBURN, 1985) e o fato que o complexo telomerase, junto com todas as proteínas acessórias, é montado de modo constitutivo (JIANG et al., 2013), tem facilitado o desenvolvimento de uma grande quantidade de estudos estruturais dos componentes da telomerase neste organismo (JIANG et al., 2013, 2015; CHAN et al., 2017; WANG, FEIGON, 2017). Por essa razão, a telomerase de *T. thermophila* é a mais conhecida atualmente, tanto em termos estruturais, quanto de interações entre os componentes do complexo. A estrutura da holoenzima telomerase de *T. thermophila*, juntamente com as outras nove proteínas que formam parte do complexo e o componente RNA (TER) foi resolvida em 2015, utilizando-se a técnica de criomicroscopia eletrônica, com uma resolução de 9 Å (JIANG et al., 2015). Esta estrutura mostra a interação do componente proteico TERT com o TER e, com o DNA telomérico, assim como a interação com complexos proteicos acessórios homólogos aos reportados em humanos e levedura, sugerindo um mecanismo de ação comum entre o complexo telomerase de diferentes organismos (Figura 4).

1.4.1. TERT, o componente proteico da telomerase

As telomerasas se assemelham às transcriptases reversas clássicas na sua capacidade de sintetizar DNA a partir de RNA, contudo, a presença de um componente RNA intrínseco dentro do complexo, que contém o molde que gera as repetições teloméricas, é uma característica exclusiva das telomerasas (AUTEXIER; LUE, 2006; LEWIS; WUTTKE, 2012). Consequentemente, elas apresentam características estruturais e funcionais conservadas com as transcriptases reversas clássicas, mas também características estruturais e funcionais próprias (AUTEXIER; LUE, 2006).

Em termos estruturais, os TERT canônicos contêm quatro domínios principais (Figura 5A). Os domínios da região amino-terminal, TEN (*Telomerase Essential N-terminal*) e TRB (*Telomerase RNA-Binding*) são regiões exclusivas das telomerasas (JACOBS et al., 2006; ROUDA; SKORDALAKES, 2007) e funcionalmente são fundamentais na interação com ácidos nucleicos (DNA genômico/TER) e com outras proteínas do complexo telomerase (XIA et al., 2000; ARMBRUSTER et al., 2001; HUANG et al., 2014; JANSSON et al., 2015; JIANG et al., 2015). Os domínios da região carboxi-terminal, RT (*Reverse Transcriptase*) e CTE (*C-Terminal Extention*), são estruturalmente similares aos das transcriptases reversas canônicas e funcionalmente estão associados com a catálise enzimática. O domínio RT compartilha uma alta conservação estrutural com as transcriptases reversas clássicas e forma as regiões estruturais da proteína conhecidas como os “dedos” (*finger*) e a “palma da mão” (*palm*) (LINGNER; CECH, 1996; LINGNER et al., 1997; PENG et al., 2001; BANIK et al., 2002; GILLIS et al., 2008); enquanto o domínio CTE, que representa o “polegar” (*thumb*), é uma região menos conservada, que também tem importância na atividade enzimática (Figura 5B) (GIARDINI et al., 2006; HOFFMAN et al., 2017). Algumas exceções à estruturas canônicas são encontradas nos TERT de alguns insetos, os quais não possuem o domínio TEN, e nos de alguns nematoídes, cujos TERTs são muito pequenos e não parecem possuir o domínio TEN, nem o domínio CTE (WYATT et al., 2010).

Grande parte do conhecimento sobre a estrutura dos TERT já descritos se deve a duas estruturas cristalográficas obtidas do TERT de *Tribolium castaneum*, o besouro da farinha. Uma dessas estruturas foi obtida na ausência de ácidos nucleicos (GILLIS et al., 2008) (PDB ID 3DU5, resolução 2.71 Å) (Figura 5B), e mostra um complexo em forma de anel contendo os domínios TRB, RT e CTE. Esta conformação anelar parece ser

mantida por interações não covalentes entre os domínios TRB e CTE. No interior do anel, se forma uma cavidade capaz de acomodar um híbrido RNA/DNA dupla fita de aproximadamente 8 nucleotídeos de comprimento, que corresponde à região do molde no TER em *T. castaneum* e ao 3' *G-overhang* da fita telomérica rica em guaninas (GILLIS et al., 2008; MITCHELL et al., 2010; WYATT et al., 2010). A outra estrutura cristalográfica foi obtida a partir do complexo da proteína e um híbrido DNA/RNA sintético (MITCHELL et al., 2010) (PDB ID 3KYL, resolução 2.7-Å) e mostra que o híbrido RNA/DNA interage com resíduos dos três domínios que formam o anel (GILLIS et al., 2008; MITCHELL et al., 2010) (Figura 5C).

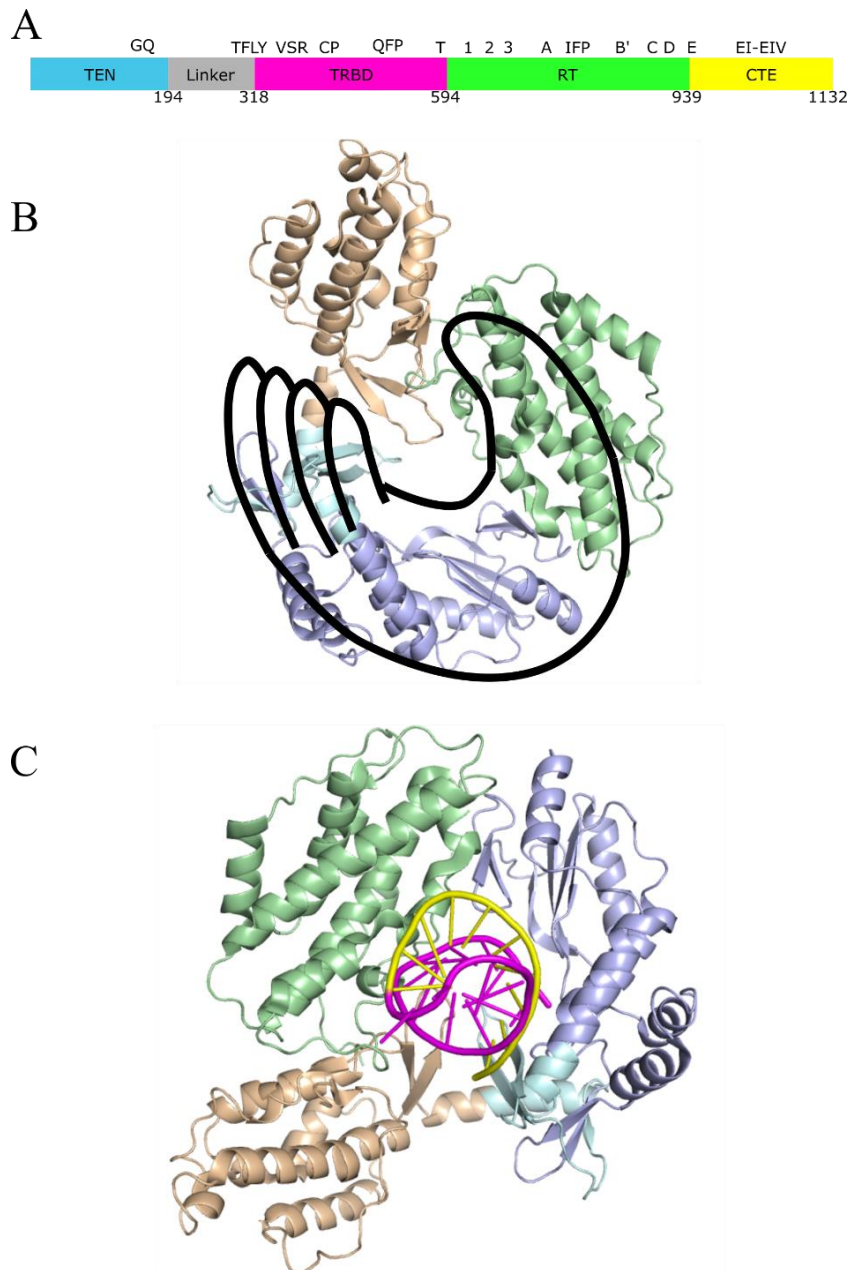


Figura 5. Esquema estrutural das telomerases

(A) Estrutura primária do TERT humano, utilizada como exemplo para mostrar os 4 domínios estruturais presentes nos TERTs canônicos, TEN, em azul; *Linker*, em cinza; TRB, em rosa; RT em verde e CTE em amarelo e os motivos funcionais de cada domínio, denotados acima de cada um deles. Figura adaptada de CHAN et al, 2017 (B) Destaque para o arranjo dos motivos do domínio RT como parte de uma mão. Esta organização é característica de transcriptases reversas canônicas e componentes TERT de diferentes organismos. (C) Representação em *cartoon* da estrutura cristalográfica da telomerase de *T. castaneum*, mostrando a configuração em anel, formada pelos três domínios estruturais da proteína. Os domínios TRB em marrom, RT em azul, destacando os subdomínios que formam os dedos (azul claro) e a palma (azul escuro), e o CTE em verde. A configuração em anel é devida a interações não covalentes entre os domínios TRB e CTE. Na cavidade do anel, se acomoda um híbrido de ácidos nucleicos compostos pela região do molde no TER (roxo) e o DNA telomérico simples fita (amarelo). As representações da proteína em *cartoon* foram geradas com o programa PyMol a partir da estrutura cristalográfica da telomerase de *T. castaneum* (PDB: 3DU5 e 3KLY) (GILLIS et al., 2008; MITCHELL et al., 2010).

O TERT de *T. castaneum* não apresenta o domínio TEN, por isso, até 2015, a localização deste domínio em relação aos três domínios que formam o anel era desconhecida. Esta pergunta foi respondida graças à resolução da estrutura da holoenzima de *T. thermophila* por microscopia crio-eletrônica (JIANG et al., 2015), que permitiu mapear o domínio TEN na estrutura do TERT, mostrando que o domínio TEN se localiza próximo ao domínio CTE e, interage tanto com o DNA telomérico, quanto com os complexos proteicos que se associam ao 3'*G-overhang* (JIANG et al., 2015) (Figura 6A).

1.4.1.1 Domínios da região amino-terminal do TERT (TEN e TRB)

A importância de regiões conservadas na região amino-terminal do TERT foi primeiramente mostrada em leveduras de brotamento pela indução de mutações em aminoácidos conservados nestas regiões. Foi observado que algumas mutações incapacitavam a levedura de manter o tamanho dos telômeros, porém nem sempre aboliam a atividade enzimática da telomerase ou sua processividade (FRIEDMAN; CECH, 1999; XIA et al., 2000; BOSOY et al., 2003). Também foi mostrado que a redução no tamanho dos telômeros estava, na maioria dos casos, relacionada com uma incapacidade de associação entre os componentes TERT e TER, sugerindo que estes mutantes eram defeituosos na montagem do complexo ribonucleoproteico e não na atividade catalítica (BOSOY et al., 2003).

O domínio TEN é o domínio menos conservado tanto em sequência de aminoácidos quanto em estrutura terciária. Porém, sua presença é indispensável para a atividade de telomerase em humanos, *S. cerevisiae* e *T. thermophila* (FRIEDMAN; CECH, 1999; BRYAN et al., 2000; ARMBRUSTER et al., 2001; LAI et al., 2001), pois está envolvido com o reconhecimento do DNA telomérico como substrato e na processividade da adição das repetições teloméricas (MORIARTY et al., 2004;

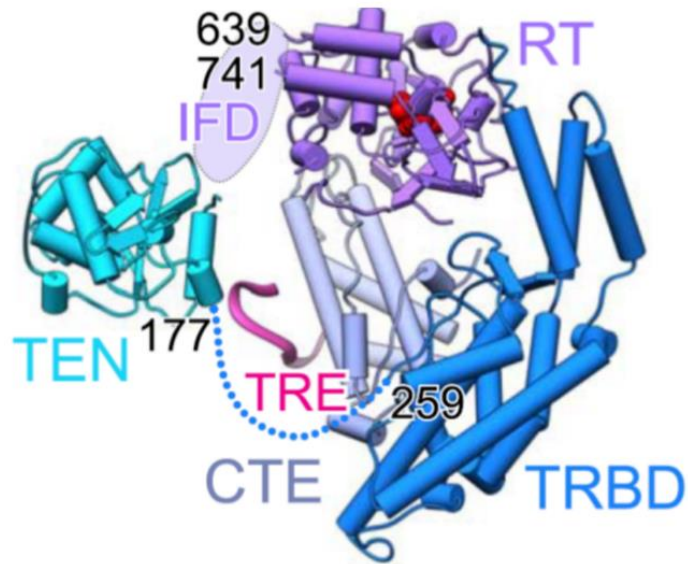
WYATT et al., 2010). O domínio TEN parece servir como uma âncora que auxilia na estabilização do DNA telomérico fora da estrutura anelar e permite sua acomodação no sítio catalítico (WYATT et al., 2010; JIANG et al., 2015). Isto aumenta a processividade da enzima, já que evita que a associação entre o DNA telomérico e o TERT seja perdida durante a translocação (Figura 3). O domínio TEN *in vivo* também apresenta a função de interação com proteínas que servem de ponte com o 3' *G-overhang*. Esta interação é fundamental para o recrutamento da telomerase para os telômeros e na biogênese do complexo ribonucleoproteico (JIANG et al., 2015; CHAN et al., 2017).

De modo interessante, a telomerase de *T. castaneum* é funcional e processiva, apesar de carecer do domínio TEN (MITCHELL et al., 2010). Atualmente, existem duas hipóteses para explicar este fato: a primeira é a de que neste organismo outras partes da enzima sejam suficientes para manter a interação com o DNA telomérico (MITCHELL et al., 2010); a segunda é a de que existe uma outra proteína ainda desconhecida e que não está ligada covalentemente ao complexo, que exerce a função atribuída ao domínio TEN em outros organismos (MASON et al., 2011). Esta segunda hipótese é suportada por experimentos que mostram que a ação de telomerase em humanos pode ser reconstituída pela expressão conjunta de dois fragmentos independentes do TERT humano, um correspondente a uma versão truncada do TERT, carente de domínio TEN, e o segundo correspondendo ao domínio TEN isolado (ROBART; COLLINS, 2011). Estes resultados mostram que, embora todos os domínios sejam necessários para se ter atividade de telomerase, a ligação covalente entre o domínio TEN e os domínios que conformam o anel não é indispensável.

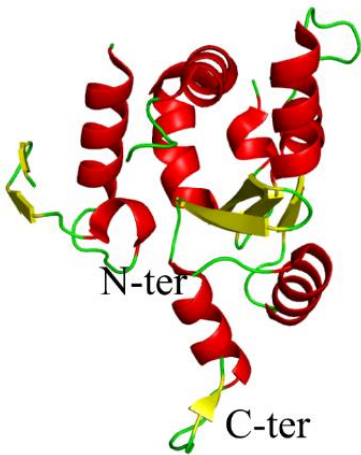
Atualmente, existem duas estruturas tridimensionais disponíveis do domínio TEN, a do ciliado *T. thermophila* (JACOBS et al., 2006) (PDB ID: 2B2A, resolução

2.22Å) e a da levedura *Ogataea polymorpha* (POLSHAKOV et al., 2016) (PDB ID: 5LGF) (Figura 6B-C). A estrutura do domínio TEN de *T. thermophila* foi resolvida por cristalografia de raios x e apresenta uma organização estrutural com alternância de hélices- α e folhas- β (JACOBS et al., 2006) (Figura 6B). Já a estrutura do domínio TEN de *O. polymorpha* foi resolvida por ressonância magnética nuclear (NMR) (PDB ID: 5LGF) (POLSHAKOV et al., 2016) e apresenta uma grande quantidade de hélices- α e de regiões desenoveladas, não apresentando similaridade estrutural com a estrutura de *T. thermophila*. Estes dados evidenciam a baixa conservação estrutural deste domínio (Figura 6C).

A



B



C

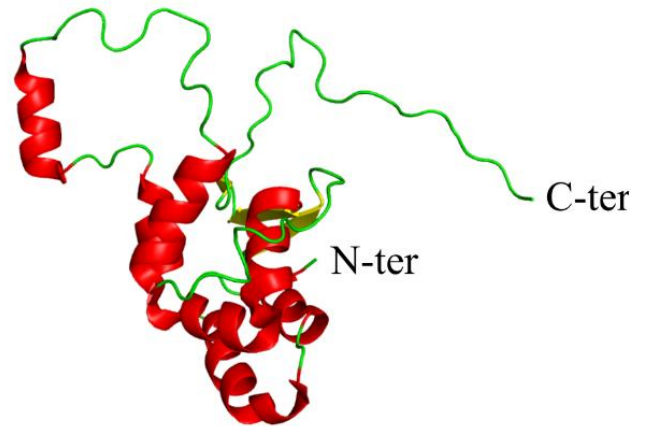


Figura 6. Representação estrutural do domínio TEN do TERT

(A) Representação em *cartoon* do TERT de *T. thermophila* mostrando a localização do domínio TEN em relação aos três domínios que constituem a conformação anelar. Figura retirada de JIANG et al. 2015. (B) Estrutura cristalográfica do domínio TEN de *T. thermophila* (JACOBS et al., 2006) PDB: 2B2A. As hélices- α são mostradas em vermelho, as folhas- β em amarelo e os *loops* e elementos desordenados em verde. Este domínio está principalmente conformado por hélices- α e folhas- β . (C) Estrutura representativa do domínio TEN de *O. polymorpha* resolvida por NMR PDB:5LGF (POLSHAKOV et al., 2016). O código de cores é o mesmo do que para *T. thermophila*. esta estrutura está formada principalmente por hélices- α e apresenta uma grande quantidade de regiões desordenadas. As representações em *cartoon* foram geradas com o programa PyMol.

Entre o domínio TEN e o domínio TRB há uma região chamada “*linker*” (Figura 5A). Esta região é intrinsecamente desenovelada, apresenta uma conservação muito baixa de estrutura primária entre os eucariotos e seu tamanho varia entre 20 e 500 aminoácidos. A sua função parece ser a de conferir flexibilidade ao domínio TEN em

relação ao restante da proteína (XIA et al., 2000; MORIARTY et al., 2004; AUTEXIER; LUE, 2006; JIANG et al., 2015). Esta flexibilidade parece ser relevante para a ligação e estabilização de ácidos nucleicos, e também, para permitir a correta localização do domínio TEN dentro do complexo enzimático (LEWIS; WUTTKE, 2012; JIANG et al., 2015). Em ciliados, próximo à região do *linker* encontra-se o motivo CP2 (JANSSON et al., 2015), cujo homólogo em mamíferos é o motivo TFLY (HARKISHEIMER et al., 2013). Ambos têm importância funcional e estão envolvidos na interação com o TER e a montagem do complexo telomerase (HARKISHEIMER et al., 2013; JANSSON et al., 2015).

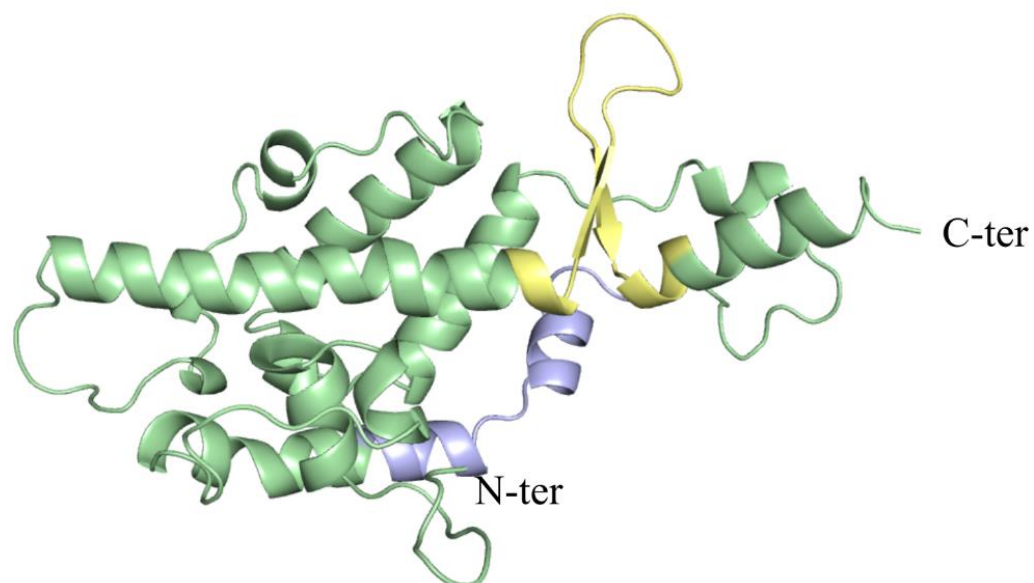
O domínio TRB dos TERTs é estruturalmente conservado nos eucariotos (ROUDA; SKORDALAKES, 2007; RICE; SKORDALAKES, 2016). Ele apresenta três motivos, CP, QFP e T, que interagem com regiões conservadas dentro do componente TER, envolvidas com o início da replicação telomérica e a delimitação do limite do molde (XIA et al., 2000; AUTEXIER; LUE, 2006; WYATT et al., 2010; JANSSON et al., 2015).

Existem diversas estruturas cristalográficas do domínio TRB de diferentes organismos (ROUDA; SKORDALAKES, 2007; GILLIS et al., 2008; HARKISHEIMER et al., 2013; JANSSON et al., 2015) (PDB IDs: 2R4G, resolução 1,71 Å; 4LMO, resolução 2,37 Å; 4O26, resolução 3,0 Å; 5C9H, resolução 3,0 Å). Todas mostram uma boa conservação estrutural geral, com um grande conteúdo de hélices- α (aproximadamente 10) e 2 folhas- β curtas (ROUDA; SKORDALAKES, 2007) (Figura 7).

O motivo T tem a conformação de um β -*hairpin* e é altamente conservado. O motivo CP, que é menos conservado, é constituído por duas hélices- α . Juntos, estes

motivos formam duas cavidades com aminoácidos conservados expostos ao solvente (ROUDA; SKORDALAKES, 2007). A região TBE (*template boundary element*) do componente RNA interage com estes dois motivos, posicionando-se para o início da replicação telomérica (JANSSON et al., 2015). Acredita-se que esta interação limita a quantidade de nucleotídeos do TER que têm acesso ao sítio ativo do TERT, garantindo assim, que só a região do molde seja copiada durante a síntese do DNA telomérico (CHEN; GREIDER, 2003; JANSSON et al., 2015).

A



B

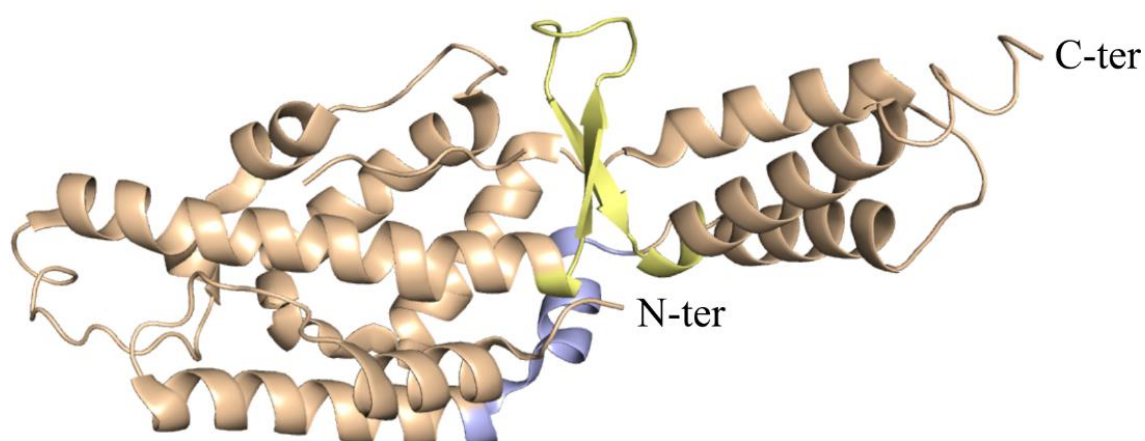


Figura 7. Representação estrutural do domínio TRB do TERT

(A) Representação em *cartoon* da estrutura cristalográfica do domínio TRB de *Takifugu rubripes* (HARKISHEIMER et al., 2013) (PDB: 4LMO) os dois motivos principais, estão destacados, T em amarelo e CP em azul. (B) representação em *cartoon* da estrutura cristalográfica do domínio TRB de *T. thermophila* (ROUDA; SKORDALAKES, 2007) (PDB: 2R4G) os dois motivos principais, estão destacados, T em amarelo e CP em azul. O domínio TRB nos dois organismos mostra uma organização estrutural similar, com uma estrutura rica em hélices- α e o β -*hairpin* característico que forma o motivo T. As representações das proteínas foram geradas com o programa PyMol.

1.4.1.2 Domínios da região carboxi-terminal do TERT (RT e CTE)

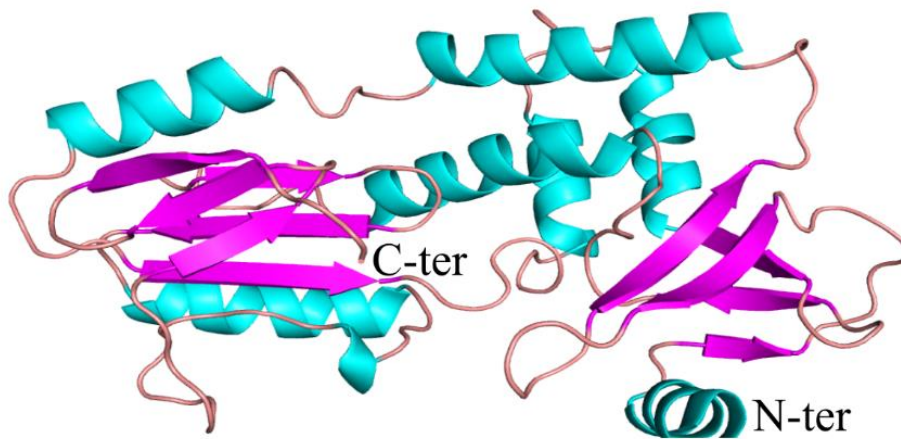
Os domínios catalíticos da região carboxi-terminal e mais especificamente, o domínio RT das telomerasas de *S. cerevisiae* e *Euplotes aediculatus* (LINGNER; CECH, 1996; LINGNER et al., 1997), foram os primeiros a serem descritos, devido à conservação de motivos compartilhados com transcriptases reversas não teloméricas (PENG et al., 2001; WYATT et al., 2010). Estudos bioquímicos e utilizando mutagênese sítio-dirigida mostram que neste domínio existem motivos de grande importância para a atividade catalítica da telomerase (WEINRICH et al., 1997; LINGNER et al., 1997).

O domínio RT apresenta sete motivos conservados, presentes também em transcriptases reversas canônicas. São eles os motivos 1, 2, A, B', C, D e (Figura 5A). Como mencionado anteriormente, esses motivos formam as regiões estruturais conhecidas como palma (*palm*) e os dedos da mão (*fingers*) (Figura 5B) (PENG et al., 2001). Mais recentemente, foram descritos dois motivos adicionais nos TERT, o motivo 3, presente entre os motivos 2 e A, que é importante para a velocidade de adição de nucleotídeos e a processividade da enzima (XIE et al., 2010; LEWIS; WUTTKE, 2012) e uma inserção entre os motivos A e B' que formam o motivo IFD, que tem importância na interação com proteínas que estabilizam o complexo (WYATT et al., 2010; DEY; CHAKRABARTI, 2018). O motivo IFD do TERT de *T. thermophila* parece interagir com a proteína acessória p50 e o domínio TEN cumprindo algum papel no recrutamento do TERT ao DNA telomérico (CHAN et al., 2017).

A única estrutura cristalográfica disponível do domínio RT é a do TERT de *Tribolium castaneum* (GILLIS et al., 2008) (PDB ID 3DU5, resolução 2.71 Å) (Figura 8A), composto por uma mistura de hélices- α e folhas- β , formando os dois subdomínios estruturais, dedos e palma. A estrutura de *T. castaneum* compartilha alta conservação

estrutural com outras transcriptases reversas, como a do HIV (Vírus da Imunodeficiência Humana) (Figura 8B) (GILLIS et al., 2008; WANG; FEIGON, 2017). A estrutura cristalográfica do TERT de *T. castaneum* associada a um híbrido DNA/RNA, mostra que os mecanismos de associação do RT com ácidos nucleicos e o sítio catalítico do TERT são muito similares à transcriptase reversa do vírus HIV (MITCHELL et al., 2010).

A



B

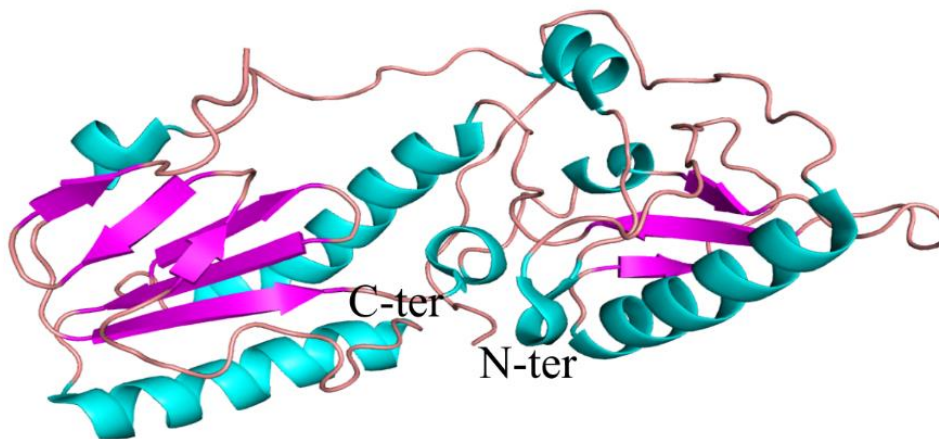


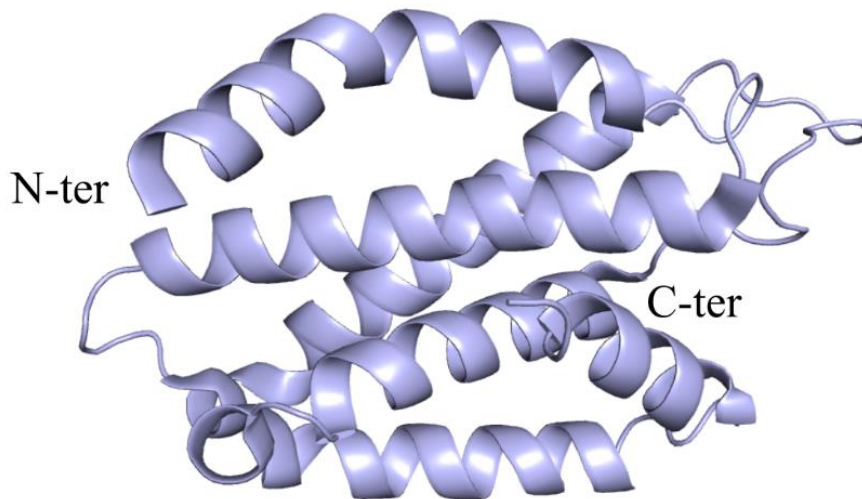
Figura 8. Representação estrutural domínio RT

(A) Representação em *cartoon* do domínio RT do TERT de *T. castaneum* (GILLIS et al., 2008) PDB: 3DU5. As hélices- α estão mostradas em azul, as folhas- β em magenta e os *loops* e elementos desordenados em rosa. (B). Representação em *cartoon* do domínio RT da transcriptase reversa do vírus HIV (SARAFIANOS et al., 2001) PDB: 1HYS. O código de cores é o mesmo do usado para *T. castaneum*. Há uma conservação estrutural notória entre as duas estruturas, evidenciando que o domínio catalítico do TERT é muito similar às transcriptases reversas canônicas. As representações das proteínas foram geradas com o programa PyMol.

Já o domínio CTE, apresenta um menor grau de conservação de sua estrutura primária, comparado aos domínios TRB e RT. O CTE está ausente no TERT de *Caenorhabditis elegans* e de outros nematóides, sugerindo que suas funções podem variar de acordo com a espécie (WYATT et al., 2010). Em termos gerais, o CTE compõe a região estrutural conhecida como polegar da mão (*thumb*), e sua presença é importante para a atividade de telomerase em *T. thermophila* e humanos, no entanto, parece ser dispensável em leveduras (LINGNER et al., 1997; BANIK et al., 2002; HUARD et al., 2003). Tem sido sugerido que ele está envolvido com a processividade e localização da enzima, mas não é indispensável para a adição de nucleotídeos (WYATT et al., 2010; MASON et al., 2011).

A estrutura cristalográfica do domínio CTE do TERT de humanos foi recentemente resolvida (HOFFMAN et al., 2017) PDB: 5UGW, resolução 2,31Å (Figura 9A), mostrando conservação estrutural quando comparado com o CTE do TERT de *T. castaneum* (Figura 9B). Este domínio contém uma estrutura rica em hélices- α com *loops* expostos ao substrato, que parecem ajudar na estabilização do híbrido DNA/RNA e, no posicionamento do DNA telomérico no sítio catalítico. (WYATT et al., 2010). O domínio CTE também tem importância farmacológica, uma vez que é nele que está contido o motivo FVYL, considerado o sítio de interação do fármaco BIBR1532, inibidor específico, não-nucleosídico, da atividade de telomerase em humanos (BRYAN et al., 2015; HOFFMAN et al., 2017).

A



B

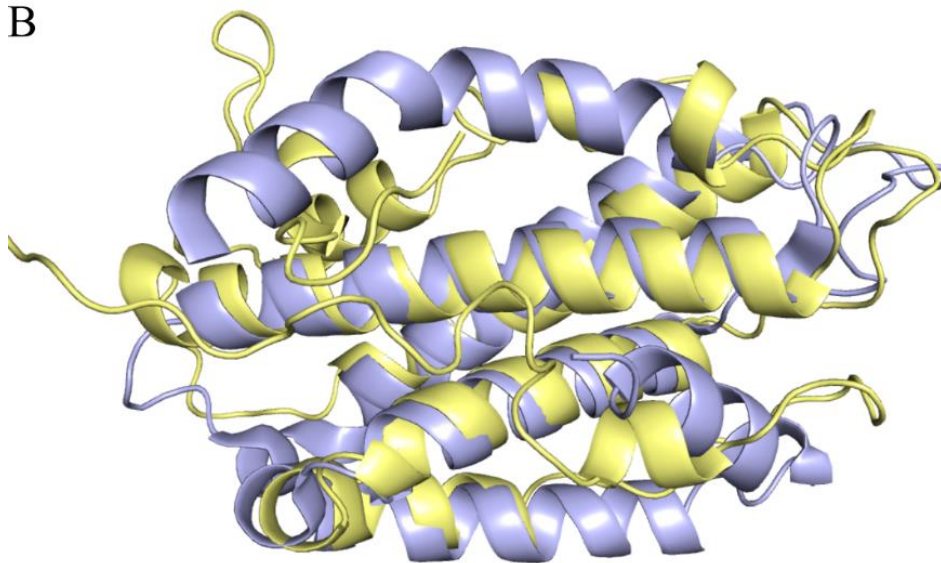


Figura 9. Representação estrutural do domínio CTE

(A) representação em *cartoon* do domínio CTE do TERT humano (HOFFMAN et al., 2017) (PDB: 5UGW). (B) superposição de carbonos α entre o domínio CTE de humano (azul) e o domínio CTE de *T. castaneum* (amarelo) (GILLIS et al., 2008) (PDB ID: 3DU5) mostrando que há uma boa superposição entre os dois domínios, mostrando conservação estrutural entre os TERTs das duas espécies. As representações da proteína em *cartoon* foram geradas com o programa PyMol.

1.4.2 O componente RNA (TER) da telomerase

O RNA da telomerase (TER) é um longo RNA não codificador (lncRNA), indispensável para o funcionamento do complexo ribonucleoproteico (LIN et al., 2004).

O componente TER é pouco conservado, apresentando variações tanto em sequência primária, quanto em tamanho, quando comparado entre os diferentes organismos. No entanto, todos os TER descritos têm elementos com estrutura secundária conservada, o que lhes permite realizar funções similares e interagir com motivos conservados entre os

TERT (MUSGROVE et al., 2018) (Figura 10). As regiões conservadas do TER interagem principalmente com resíduos conservados dentro do domínio TRB (LUE, 2004; MITCHELL et al., 2010; HUANG et al., 2014; JANSSON et al., 2015). Estudos mutacionais realizados com o TER em diferentes organismos mostraram que a substituição de nucleotídeos específicos, especialmente naqueles que alteram a estrutura secundária de segmentos conservados na molécula de RNA, são capazes de abolir a interação entre TER e TERT, resultando no encurtamento dos telômeros (LIN et al., 2004). Entre os componentes estruturais conservados do TER, está o elemento TBE, encontrado na porção 5' da molécula de RNA (LAI et al., 2001). O TBE interage com os motivos T e CP do domínio TRB da maioria dos TERT, e em grupos específicos com o domínio CP2 em ciliados e TFRY em vertebrados (HARKISHEIMER et al., 2013; JANSSON et al., 2015). Esta interação é importante para o início da síntese das repetições teloméricas e para a definição dos limites da sequência molde. A estrutura tridimensional do TBE de *T. thermophila* foi resolvida por NMR (PDB ID: 2M22). Também está disponível a estrutura cristalográfica do complexo TRBD/TBE (PDB ID: 5C9H, resolução 3,0 Å) de *T. thermophila* (RICHARDS et al., 2006; JANSSON et al., 2015).

Uma segunda região conservada e presente em todos os TERs é a região do molde (Figura 10), que consiste na sequência reversa complementar da repetição telomérica, com tamanho aproximado de uma repetição e meia (LUE, 2004). O molde se acomoda na cavidade do anel dos TERT, em contato íntimo com os motivos da palma no domínio RT (GILLIS et al., 2008; MITCHELL et al., 2010).

Já a região de *pseudoknot* (AUTEXIER; LUE, 2006; SANDIN; RHODES, 2014) embora não esteja presente em todos os TER descritos, é indispensável para atividade de telomerase *in vitro* e *in vivo* (Figura 10), dada sua importância na montagem do

complexo catalítico. As estruturas dos *pseudoknots* de diferentes organismos têm sido resolvidas por NMR e todas mostraram uma conservação estrutural geral (LEEPER et al., 2003; THEIMER et al., 2005; KIM et al., 2008; CASH; FEIGON, 2017).

Finalmente, o elemento *Stem terminus*, que em mamíferos recebe o nome de motivo CR4/5 e em *T. thermophila* é chamado de *Stem loop IV* (Figura 10), interage especificamente com domínio TRB e sua estrutura é importante para que esta interação ocorra (COLLINS et al., 2006; HUANG et al., 2014). Por conseguinte, os elementos conservados do TER são fundamentais para a atividade e processividade enzimática da telomerase (LUE, 2004; CHEN; GREIDER, 2004).

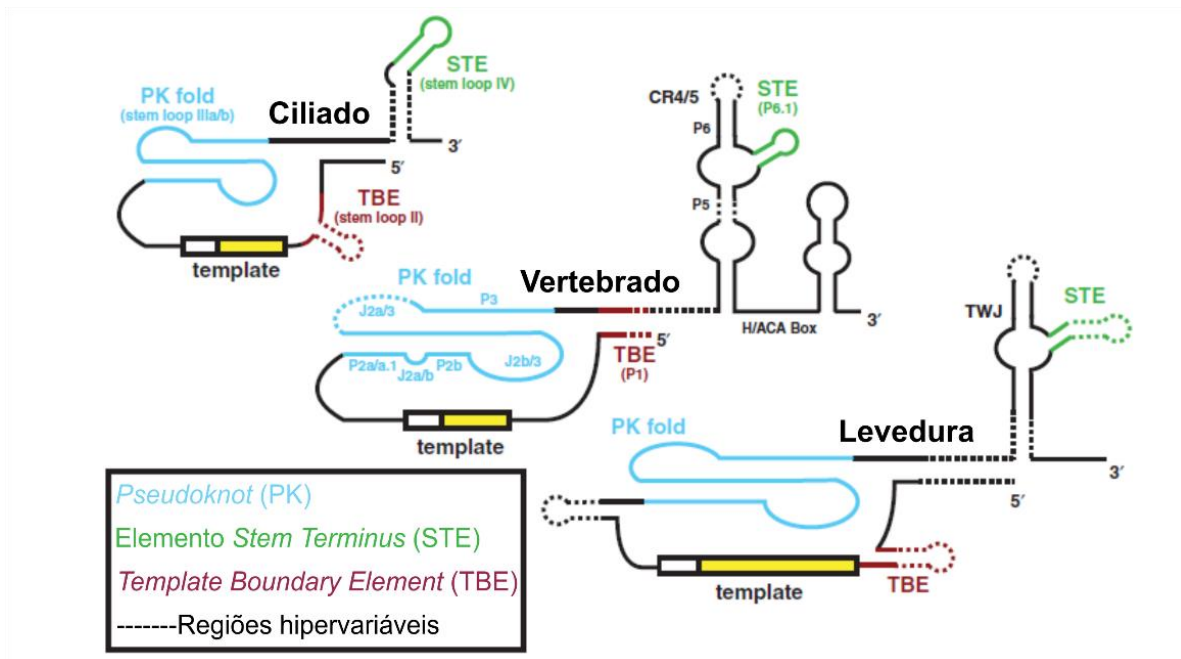


Figura 10. Elementos estruturais conservados do componente TER de diferentes grupos de organismos

Estruturas secundárias representativas dos TER de ciliados, vertebrados e leveduras. É possível se observar a diferença em tamanho e organização em cada um dos grupos, assim como a conservação de alguns elementos. As regiões conservadas estão destacadas, o molde para a sequência telomérica em amarelo, o *pseudonknot* em azul, o elemento *stem terminus* em verde e o TBE em vermelho. Figura traduzida de (MUSGROVE et al., 2018)

1.5 Telômeros de *Leishmania* sp.

A sequência telomérica repetitiva nos finais dos cromossomos dos parasitos do gênero *Leishmania* é a mesma dos telômeros de vertebrados, 5'-TTAGGG-3', e eles apresentam atividade de telomerase comprovada (CANO et al., 1999; GIARDINI et al., 2011). As regiões subteloéricas de *Leishmania* são formadas principalmente por repetições de aproximadamente 100 pb (pares de base) conhecidas como LCTAS (*Leishmania conserved telomere associated sequences*) que variam em tamanho, mesmo em cromossomos da mesma espécie (FU; BARKER, 1998; CONTE; CANO, 2005).

1.5.1 Complexo telomérico de *Leishmania* sp.

Estudos realizados em *L. amazonensis* permitiram caracterizar diversos componentes do complexo telomérico de *Leishmania*, principalmente proteínas que se associam aos telômeros. As proteínas LaRbp38, LaRPA-1 e LCALa foram identificadas interagindo com o DNA telomérico simples-fita e possivelmente com o componente TERT de *L. amazonensis* (LaTERT), uma vez que todos foram isolados a partir de extratos proteicos semi-purificados que apresentavam atividade de telomerase (FERNÁNDEZ et al., 2004; SIQUEIRA-NETO et al., 2007; MOREA et al., 2017). Estas proteínas parecem ser homólogos funcionais das proteínas CDC13 de leveduras e POT1 de outros eucariotos, já que nos genomas dos Tripanosomatídeos não existem ortólogos para as mesmas (PAVANI et al., 2014). As proteínas LaTBP1, LaRbp38 e LaTRF reconhecem regiões teloméricas dupla fita e com exceção de LaRbp38 (FERNANDES et al. em preparação), compartilham semelhanças estruturais e funcionais com proteínas teloméricas descritas em outros eucariotos (ex: RAP1, TRF1 e TRF2) (LIRA; GIARDINI; et al., 2007; LIRA; SIQUEIRA NETO; et al., 2007). Um homólogo putativo da proteína PinX1, um inibidor natural da telomerase em humanos e leveduras

(GUGLIELMI; WERNER, 2002; LIN; BLACKBURN, 2004), foi encontrado no genoma de *L. major*. Atualmente estudos estão sendo realizados com esta proteína para verificar se em *Leishmania*, de forma semelhante ao que acontece em outros organismos, PinX1 atua como um inibidor natural da atividade de telomerase. Uma visão geral da estrutura telomérica conhecida atualmente dos parasitos do gênero *Leishmania* é apresentada na Figura 11.

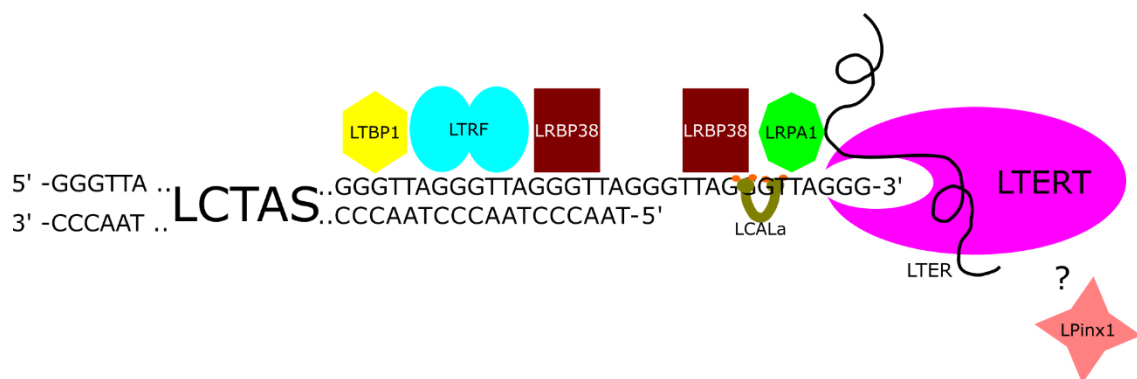


Figura 11. Panorama atual da composição da cromatina telomérica de *Leishmania sp.*

O telômero está formado por regiões subteloiméricas (LCTAS), uma região dupla fita e uma protusão simples fita 3'G-overhang. As proteínas LTBP-1, LTRF e LRbp38 interagem com o DNA dupla fita, as proteínas LRbp38, LRPA-1 e LCALa interagem com o DNA simples fita. As interações proteicas afetam a interação do complexo TERT-TER com os telômeros. Uma proteína putativa homóloga a PinX1, que pode atuar como um inibidor natural do complexo telomerase, está sendo caracterizada. Imagem atualizada a partir de (SILVA, DA et al., 2013).

1.5.2 Características do complexo telomerase em *Leishmania sp.*

A atividade de telomerase foi detectada a partir de extratos purificados de formas promastigotas de *L. tarentolae* e *L. major*. A processividade da telomerase de *Leishmania sp.* é baixa em comparação com a telomerase de ciliados e de *T. brucei*, mas preenche todos os critérios bioquímicos essenciais apresentados pelas telomerases canônicas (CANO et al., 1999).

Posteriormente, extratos proteicos de *L. amazonensis* com atividade de telomerase foram parcialmente purificados por fracionamento em colunas

cromatográficas complementares e por afinidade ao DNA telomérico. (GIARDINI et al., 2011).

O componente TERT foi clonado e caracterizado em quatro espécies diferentes de *Leishmania* (GIARDINI et al., 2006). Os TERT de *Leishmania* e de outros Tripanosomatídeos são uns dos maiores já descritos, possuindo peso molecular de aproximadamente 156 kDa, e apresentam os mesmos domínios estruturais e funcionais compartilhados com TERT de outros organismos (Figura 5A), além de substituições em resíduos de aminoácidos espécie-específicos (GIARDINI et al., 2006; LIRA; GIARDINI; et al., 2007). Os TERT das diferentes espécies de *Leishmania* sp. mostram uma alta conservação dentro do gênero, mas suas estruturas primárias são pouco conservadas quando comparadas aos TERT de outros organismos, incluindo os do gênero *Trypanosoma* (GIARDINI et al., 2006). Apesar da importância desses estudos, eles não aprofundaram o conhecimento sobre as características estruturais do componente TERT de *Leishmania* sp.

Recentemente, os componentes TER de *L. amazonensis* e *L. major* (LeishTER) foram caracterizados como RNAs longos não codificadores de aproximadamente 2.100 nucleotídeos, possivelmente uns dos maiores componentes RNA da telomerase descritos até o momento. A predição de estrutura secundária de LeishTER mostra que a região do molde está localizada no terminal 5' do RNA, contém 12nt, e é conservada dentro do gênero *Leishmania* (Figura 12), compartilhando também alta similaridade com a sequência do molde do TER de *T. brucei* (SANDHU et al., 2013; GUPTA et al., 2013). Outras regiões próprias do TER de ciliados, como o domínio TBE e o Box C/D snoRNA (pequeno RNA nucleolar), também foram descritas no componente TER de *Leishmania* (VASCONCELOS et al., 2014), sugerindo que ao menos estruturalmente ele compartilha domínios estruturais importantes com outros protozoários.

LmTER

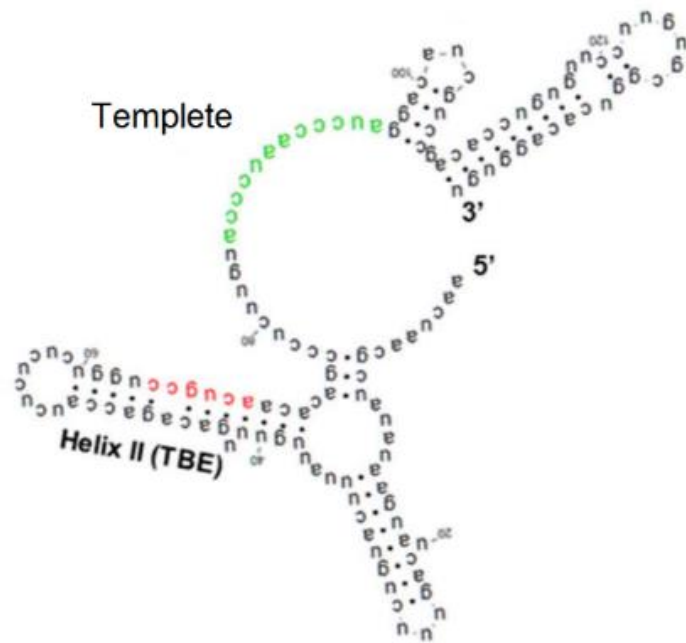


Figura 12. Predição da estrutura secundária de TER em *L. major*

Na predição parcial da estrutura secundária se destaca o elemento TBE e o motivo característico CCGUCA (vermelho) assim como a sequência do molde da repetição telomérica (verde) (Figura retirada de (VASCONCELOS et al., 2014)).

2. OBJETIVOS

Objetivo geral: Caracterizar estruturalmente o componente TERT da telomerase de *L. major*.

Objetivos específicos:

- Predizer estruturas terciárias para os diferentes domínios do componente TERT de *L. major* e realizar estudos estruturais comparativos com as estruturas cristalográficas disponíveis dos TERT de outros organismos eucariotos.
- Expressar os domínios da telomerase em sistema heterólogo e estabelecer protocolos de purificação.
- Caracterizar por métodos bioquímicos e biofísicos os domínios purificados para confirmar experimentalmente os resultados obtidos pelas técnicas de bioinformática.

3. MATERIAIS E MÉTODOS

3.1 Alinhamento de seqüências e predição da estrutura secundária

Os alinhamentos múltiplos entre as seqüências de aminoácidos dos TERT foram realizados utilizando a ferramenta Multiple Sequence Comparison by Log - Expectation (MUSCLE) da EMBL-EBI (EDGAR, 2004; MCWILLIAM et al., 2013) com o intuito de se compararem as seqüências de diferentes organismos e mapear os domínios e motivos estruturais. As predições de estrutura secundária foram realizadas utilizando-se a ferramenta GOR V (SEN et al., 2005) a partir das seqüências de aminoácidos dos TERT de diferentes espécies.

3.2 Modelagem molecular e simulações de dinâmica molecular

As modelagens moleculares foram realizadas baseadas em métodos de *threading*, que permitem analisar a estrutura de proteínas não homologas e cuja estrutura não tenha sido resolvida, mas que compartilham enovelamento e estrutura terciária similares às proteínas de estrutura conhecida. Estas análises de *threading* foram realizadas a partir das estruturas cristalográficas de domínios isolados ou do TERT de diferentes organismos, utilizando o programa Phyre2 (KELLEY et al., 2015) disponível no servidor Imperial College, Londres, Inglaterra (<http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index>).

O domínio TRB do TERT de *L. major* foi modelado a partir do domínio TRB do TERT de *Takifugu rubripes* (Score: 197,5; E value =1,4 e-30; Identidade: 26%; PDB ID 4MLO) (HARKISHEIMER et al., 2013) e do TERT completo de *Tribolium castaneum* (Score: 188,17; E value =5,7 e-29; Identidade: 15%; PDB ID 3DU5) (GILLIS et al., 2008).

A predição estrutural da região contendo os domínios TRB, RT e CTE do TERT de *L. major* foi obtida a partir da estrutura completa do TERT de *Tribolium castaneum* (Score: 725,82; E value =0; Identidade: 18%; PDB ID 3DU5) (GILLIS et al., 2008).

Os modelos gerados *in silico* foram submetidos a simulações de dinâmica molecular para minimizar sua energia livre, o utilizando GROMACS v.4.5.3 (SPOEL, VAN DER et al., 2005) (Groningen Machine for Chemical Simulation) e o campo de força GROMOS 96 53 a 6 (CHRISTEN et al., 2005), ajustando-se o pH do sistema para 7,0, a temperatura para 298 K e a pressão para 1,0 bar. A configuração inicial do sistema foi gerada após dois processos de minimização de energia utilizando um algoritmo do tipo *steepest descent*, sendo o primeiro deles realizado com todos os átomos da proteína congelados e o segundo com restrição de posicionamento dos átomos da cadeia principal. Na etapa seguinte, 200 ps de dinâmica molecular foram calculados com restrição de posicionamento dos átomos da cadeia principal da proteína (método PRMD). Para induzir a dinâmica molecular de energia dos modelos obtidos, foram utilizados mais 100-150 nanosegundos.

3.3 Docking Molecular

Foram realizados protocolos computacionais para se avaliar a interação entre o modelo *in silico* da estrutura terciária do domínio TRB e o elemento TBE extraído da estrutura cristalográfica do domínio TRB complexada ao TBE de *T. Thermophila* (PDB 5C9H). Para isto, foi utilizado o programa HADDOCK (DOMINGUEZ et al., 2003). Brevemente, este programa utiliza um protocolo de *docking* que consiste em três estágios. O primeiro estágio consiste na randomização das orientações entre as moléculas envolvidas no *docking*, aonde cada uma é randomicamente rotacionada ao redor do seu centro de massa, seguida de uma minimização de energia de corpo rígido.

Neste estágio inicial, são gerados tipicamente 1000 complexos e as 200 melhores soluções em termos de energia intramolecular passam ao estágio seguinte. O segundo estágio consiste em três refinamento de *simulated annealing*, sendo que no primeiro deles, ambas as moléculas são consideradas como corpos rígidos com refinamento de suas orientações; no segundo, as cadeias laterais da interface entre as moléculas ganham liberdade rotacional e no terceiro refinamento, elas ganham liberdade para que façam maiores rearranjos estruturais. No terceiro estágio, ocorre uma série de refinamentos de dinâmica molecular no espaço cartesiano na presença de solventes explícitos. As estruturas finais são agrupadas (“*clusterizadas*”) baseadas nos valores de desvio quadrático médio (*root mean square deviation*; r.m.s.d) entre os C α das estruturas. Um *cluster* contém estruturas que possuem um r.m.s.d. menor que 1,0 Å dos C α dos resíduos presentes na interface. Os *clusters* resultantes são ranqueados baseados na pontuação HADDOCK (*HADDOCK score*) que sumariza os diferentes valores energéticos, com maior peso para os valores de energia de interação eletrostática, de interações de van der Waals, de restrições de distância e para a superfície de área enterrada (DOMINGUEZ et al., 2003).

3.4 Estudos filogenéticos

Para os estudos filogenéticos do TERT de *L. major*, sequências homólogas a esta proteína descritas em outros eucariotos foram selecionadas através da ferramenta BLAST (ALTSCHUL et al., 1990). As sequências selecionadas foram alinhadas utilizando o programa MUSCLE (MULTiple Sequence Comparison by Log-Expectation). A partir do alinhamento final foi construída uma árvore filogenética utilizando inferência bayesiana pelo programa MrBayes v.3.1.1 (HUELSENBECK; RONQUIST, 2001). Duas cadeias de Markov-Monte Carlo de 500.000 gerações foram calculadas em quatro cadeias progressivas com trocas entre si, com um valor de fator de

temperatura de 0,2, desprezando-se as 2.500 árvores iniciais. A figura da árvore filogenética foi gerada utilizando-se o programa Mesquite v.2.72 (MADDISON, W. P. MADDISON, 2011).

3.5 Desenho e construção dos vetores de expressão com os domínios TEN, RT e TRB da telomerase de *L. major*

Os domínios do TERT de *L. major* foram delimitados por alinhamentos de sequência múltiplo e predição de estruturas secundária e terciária. Tendo os domínios delimitados, foram desenhados três clones contendo um domínio estrutural cada um, TEN (Figura 13), RT (Figura 14) ou TRB (Figura 15). Os domínios TEN e RT foram clonados no vetor pET28a+. Os plasmídeos desenhados foram construídos mediante clonagem a partir de DNA genômico pela empresa Genone (<http://www.genone.com.br/>). A sequência do domínio TRBD do TERT de *L. major* foi analisada e seus códons foram otimizados para expressão em levedura *Pichia pastoris* pela empresa GeneArt Gene Synthesis (Thermo Fisher Scientific). A sequência com os códons otimizados foi sintetizada e sub-clonada no vetor pPICZ α A pela empresa Genone (<http://www.genone.com.br/>). As sequências clonadas nos vetores encontram-se no apêndice 1.

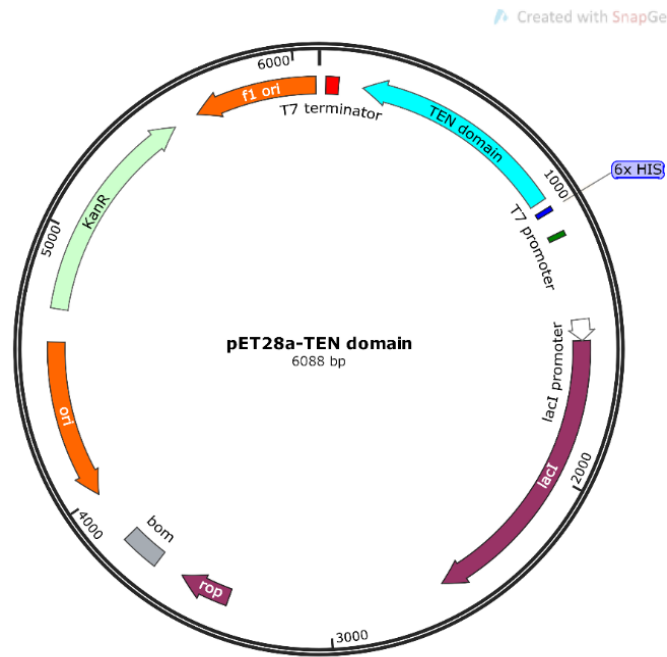


Figura 13. Mapa do vetor de expressão contendo o domínio TEN de *L. major*
 O domínio TEN de *L. major* (azul claro) foi clonado no vetor pET 28a⁺ em fusão com uma cauda de 6x Histidina no extremo amino-terminal. Imagem gerada com o programa Snap Gene Viewer 3.3.4.

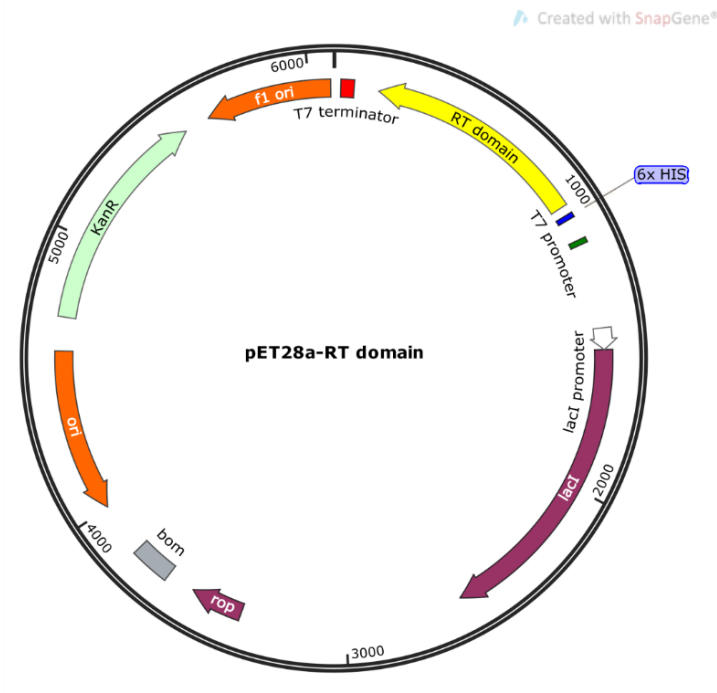


Figura 14. Mapa do vetor de expressão contendo o domínio RT de *L. major*
 O domínio RT de *L. major* (amarelo) foi clonado no vetor pET 28a⁺ em fusão com uma cauda de poli-histidina (azul) no extremo amino-terminal. Imagem gerada com o programa Snap Gene Viewer 3.3.4.

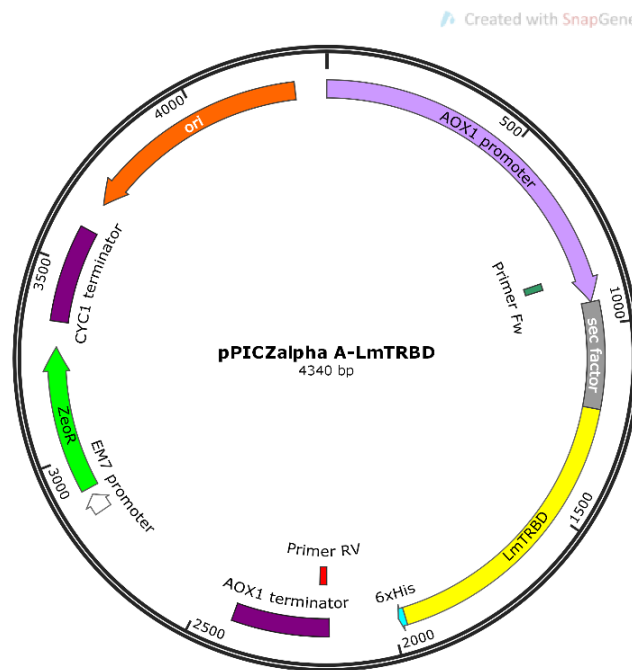


Figura 15. Mapa do vetor de expressão contendo o domínio TRBD de *L. major*

O domínio TRBD de *L. major* (amarelo) foi clonado no vetor pPiCZ- α em fusão com um sinal de secreção no extremo amino-terminal (cinza) e uma cauda de 6X Histidina (azul claro) no extremo carboxi-terminal. Imagem gerada com o programa Snap Gene Viewer 3.3.4.

3.6 Testes de mini-indução de proteína heteróloga em sistema bacteriano

Bactérias *Escherichia coli* da cepa BL21-DE3 códon⁺ ou da cepa BL21-Artic Express pRIL foram transformadas com o plasmídeo pET28a⁺ contendo a sequência de um domínio do TERT de *L. major*. Uma colônia transformante foi inoculada em 1 ml de meio Luria-Bertani (LB) [Tryptona 1%, Extrato de levedura 0,5%, NaCl 1%] com canamicina (50 μ g.ml⁻¹) e cloranfenicol (50 μ g.ml⁻¹) e incubada em agitação constante a 200 rpm a 37 °C em *shaker*. O pré-inóculo foi colocado em 50 ml de meio LB com canamicina (50 μ g.ml⁻¹) e cloranfenicol (50 μ g.ml⁻¹) e cultivado a 200 rpm a 37°C até atingir uma OD₆₀₀ (densidade óptica) de 0,6-0,8. Em seguida, foi induzida a expressão da proteína adicionando-se Isopropyl-B-D-Galactoside (IPTG) na concentração de 0,5-1 mM sob temperatura constante de 25 °C por 3 ou 4 horas. Durante este tempo, foi retirada uma amostra de 1 ml da cultura a cada hora para controle. Em seguida, as

células foram centrifugadas durante 10 minutos a 4.600 g e os precipitados foram ressuspensos em 1ml de tampão A (20mM fosfato de sódio pH 7,6, 500mM NaCl, 25mM Imidazol). Em seguida, o material ressuspenso foi submetido a sonicação por 10 sessões de 20 segundos e centrifugados a 4°C por 1 minuto a 16.000 g para separação das frações solúvel e insolúvel do extrato. As frações foram ressuspensas em 5x SDS *loading buffer* e separadas em gel de SDS-PAGE 12%.

3.7 Purificação do domínio TEN a partir da fração insolúvel do extrato proteico

Cada grama do precipitado de células bacterianas obtido como descrito no item 3.5, foi diluído em 15ml tampão de lise (Tris-HCL 50mM pH 7,0 ou 8,8, NaCl 50mM, EDTA 10mM) contendo 1mg.ml⁻¹ de lisozima (Invitrogen) e incubado no gelo por 30 minutos. Depois, a suspensão foi submetida a 30 sessões de 20 segundos de sonicação a 4 °C, seguida de 20 segundos de repouso. A amostra foi centrifugada a 4°C por 15 minutos a 4.600 g e após a centrifugação foi mantida em tampão A (20mM Glicina pH 7,6, 20mM NaCl, 6M uréia, 1mM β-mercaptoetanol).

A purificação do domínio TEN foi realizada utilizando-se cromatografia por troca iônica em uma coluna CM-Sepharose (Fast Flow, GE Healthcare) pré-equilibrada em tampão A acoplada ao sistema de cromatografia líquida de alta pressão (HPLC; AKTA Purifier™) utilizando-se um gradiente linear 0-100% de tampão B (20mM Glicina, 1 M NaCl, 6M uréia, 1mM β-mercaptoetanol pH 7,6) em 20 volumes de coluna. A purificação da proteína foi verificada por eletroforese em SDS-PAGE 12% e por *western blot*.

3.8 Purificação do domínio TEN a partir da fração solúvel do extrato proteico

Cada grama do precipitado de células bacterianas obtidos como descrito no item 3.5, foi diluído em 15ml tampão de lise (Tris-HCL 50mM pH 7,0; NaCl 50mM; EDTA

10mM) contendo $1\text{mg}\cdot\text{ml}^{-1}$ de lisozima (invitrogen), cocktail de inibidor de protease (invitrogen) 1x e 1% de N-lauryl sarcosine, as amostras foram incubadas a $37\text{ }^{\circ}\text{C}$ 30 minutos. Depois a suspensão foi submetida a 30 sessões de 20 segundos de sonicação a $4\text{ }^{\circ}\text{C}$ seguida de 20 segundos de descanso. A amostra foi centrifugada a $4\text{ }^{\circ}\text{C}$ por 15 minutos a 4.600 g e o pellet bacteriano foi descartado.

Foram feitas várias tentativas de purificar o domínio TEN a partir da fração solúvel, tanto por cromatografia de troca iônica, como por cromatografia de afinidade. Para a cromatografia de troca iônica, utilizou-se uma coluna CM-Sepharose (Fast Flow, GE Healthcare) pré-equilibrada em tampão A (Tris 20mM pH8,5; NaCl 50mM) acoplada ao sistema de cromatografia líquida de alta pressão (HPLC; AKTA Purifier™) utilizando-se um gradiente linear 0-100% de tampão B (Tris 50mM pH8,5; NaCl 50mM, Imidazole 500mM) em 15 volumes de coluna. Para a cromatografia de afinidade utilizou-se uma coluna HisTrap FF Crude (Fast Flow, GE Healthcare) pré-equilibrada em tampão A (Tris 50mM pH7,6; NaCl 50mM, Imidazole 25mM) acoplada ao sistema de cromatografia líquida de alta pressão (HPLC; AKTA Purifier™) utilizando-se um gradiente linear 0-100% de tampão B (Tris 50mM pH7,6; NaCl 50mM, Imidazole 500mM) em 15 volumes de coluna.

3.9 Re-enovelamento do domínio TEN purificado a partir da fração insolúvel

As frações da proteína recombinante referentes ao domínio TEN purificado segundo etapas descritas no item 3.7, foram reunidas. Em seguida, estas frações foram dialisadas contra tampões contendo concentrações decrescentes de uréia e β -mercaptoetanol utilizando-se os seguintes tampões:

I) 25mM Tris, 50mM NaCl, 4M uréia, 1mM β -mercaptoetanol pH 7,6 (*overnight*);

II) 25mM Tris, 50mM NaCl, 2M uréia, 1mM β -mercaptoetanol pH 7,6 (3 horas);

III) 25mM Tris, 50mM NaCl, 1mM β -mercaptoetanol pH 7,6 (*overnight*);

IV) 25mM Tris, 50mM NaCl, pH 7,6 (24 horas).

Após a diálise, as amostras de proteína foram concentradas utilizando-se concentradores Vivaspín 20 MWCO PES de 10kDa (GE Healthcare) com ciclos de centrifugação de 10 min a 2600 g a 4 °C.

3.10 Recuperação de proteínas recombinantes contidas em corpúsculos de inclusão

O isolamento de corpúsculos de inclusão foi utilizado para purificar o domínio RT recombinante. *Pellets* de bactérias transformadas, nas quais tinha sido induzida a expressão da proteína heteróloga, foram ressuspensos em solução contendo Tris-HCl 0,1M e EDTA 1mM pH 7,0, na presença de lisozima (Invitrogen) 0,5mg.ml⁻¹. Em seguida, os *pellets* foram incubados 30 minutos no gelo e sonicados por 25 ciclos de 30 segundos com 30 segundos de descanso. Após a sonicação, foi adicionado 3mM MgCl₂, DNase 10ug.ml⁻¹ (Sigma) e incubado 30 minutos a temperatura ambiente. Após o tratamento com DNase foi adicionado EDTA (20mM pH 8,0), Triton X-100 (2% v/v) e NaCl (0,5M) e incubado por mais 30 minutos a 4° C. As amostras foram centrifugadas 15 minutos a 4 °C por 4600 g e lavadas com solução contendo Tris-HCl 0,1M e EDTA 20mM pH 7,0. O *pellet* final foi ressuspensão em tampão A (20mM Glicina pH 7,6, 20mM NaCl, 7M uréia, 1mM β -mercaptoetanol) e a purificação foi conferida por eletroforese em SDS-PAGE 12% corados com azul de comassie e por *western blot*.

3.11 Western blot

Diferentes amostras proteicas foram fracionadas em géis SDS-PAGE 12% ou 15% e transferidos para membranas Amersham Hybond Low Fluorescence 0,2 mm PVDF (GE Healthcare), utilizando o sistema mini-PROTEAN Tetra system (Bio-Rad).

As membranas foram reveladas seguindo o protocolo da Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare) com algumas modificações: as incubações com os anticorpos primário e secundário foram realizadas em solução de bloqueio [PBS 1X; Tween 20 0,1%; Blotting-Grade Blocker (Bio-Rad) 1%].

Os títulos utilizados para os anticorpos foram os seguintes:

- Soro mono-específico policlonal anti-TERT-N-terminal obtido em coelho: 1/7.500;
- Anticorpo secundário anti-coelho obtido em cabra (Sigma): 1/12.500;
- Anticorpo primário anti-HIS-TAG obtido em camundongo (Bio-Rad): 1/6.000;
- Anticorpo secundário anti-camundongo obtido em cabra (Bio-Rad): 1/6.000.

3.12 Formação de duplex parcial de oligonucleotídeos para ensaio de CD e fluorescência

Os oligonucleotídeos OverhangFw e OverhangRv (Tabela 1) foram misturados em uma relação molar 1:1 a temperatura ambiente, aquecidos a 95°C por 10 minutos e deixados esfriar lentamente até atingir novamente a temperatura ambiente. O duplex parcial obtido é referido como 3' *G-Overhang*.

3.13 Espectroscopia de dicroísmo circular (CD)

A fim de se verificar o correto enovelamento dos domínios purificados, foram realizadas medidas de espectroscopia de dicroísmo circular (CD). As medidas de CD foram obtidas no intervalo de 200-260 nm utilizando-se um espectropolarímetro JASCO J-815 equipado com um controlador de temperatura de sistema Peltier. Os experimentos foram realizados a 293 K utilizando-se uma cubeta de caminho ótico de 0,5 nm, uma

velocidade de leitura de $100 \text{ nm}\cdot\text{min}^{-1}$, tempo de resposta de 1 s, um *bandwidth* de 2 nm e um *data pitch* de 0.5 nm. O espectro final corresponde à média de 20 acumulações, já subtraído da medida do tampão (branco) e normalizado através da transformação da unidade de miligraus para elipcidade residual média $[\theta]$ ($\text{degrees}\cdot\text{cm}^2\cdot\text{dmol}^{-1}\cdot\text{residue}^{-1}$). Os espectros de CD do tampão e dos DNAs têm sinais desprezíveis nas concentrações testadas. O domínio TEN foi analisado a $0,4 \text{ mg}\cdot\text{ml}^{-1}$ em tampão contendo 20 mM de glicina pH 7,6; 50mM de NaCl. O efeito da adição do DNA no espectro de CD foi avaliado mediante a adição de 150 pmol de cada um dos DNAs (Tabela 1) às amostras proteicas. A deconvolução do espectro de CD foi realizada utilizando o servidor online Dichroweb (WHITMORE; WALLACE, 2008) com o algoritmo CDSSTR e set de referencia 4 (SREERAMA; WOODY, 2000).

3.14 Espectroscopia de fluorescência estática

As medidas de espectroscopia de fluorescência estática foram realizadas em um espectrofluorímetro Hitachi F-7000 utilizando-se uma cubeta de caminho ótico de 10 mm e um *bandwidth* de 5,0 nm. Foram utilizados comprimentos de onda de excitação de 260, 280 e 295 nm e o comprimento de onda da fluorescência emitida foi detectada no intervalo de 310-410 nm na velocidade de 1200 nm/s. As amostras foram medidas na concentração de 0,3 mg/ml.

3.15 Cromatografia analítica de exclusão molecular

A cromatografia analítica de exclusão molecular do domínio TEN foi realizada a fim de se observar seu estado oligomérico em solução utilizando-se uma coluna Superdex 75 10/300 (GE Healthcare) equilibrada em tampão HEPES 40mM pH 7,4, KCl 100mM acoplada a um sistema de cromatografia líquida (HPLC; AKTA PurifierTM). Uma amostra de domínio TEN purificado ($V=1\text{ml}$; $[2,7\text{mg}\cdot\text{ml}^{-1}]$) foi

injetada num fluxo de 0,5 ml/min. Os padrões de peso molecular foram obtidos de um kit de calibração de gel filtração para baixos pesos moleculares (GE Healthcare) contendo os seguintes componentes: Blue dextran (2000 kDa), Conalbumina (75 kDa), BSA (66 kDa), Ovalbumina (43 kDa), Anidrase carbônica (29 kDa), Ribonuclease A (13,7 kDa) e Aprotinina (6,5 kDa). Os volumes de eluição de cada padrão foram medidos e utilizados para calcular o coeficiente de partição (K_{av}), que está definido como $K_{av}=(V_r-V_o)/(V_c-V_o)$, onde V_r =volume de retenção, V_o = Volume morto (Calculado em função do tempo de retenção do padrão blue dextran) e V_c =Volume de coluna. O K_{av} obtido para cada padrão foi tabulado contra o log da massa molecular para gerar uma curva padrão que foi utilizada para estimar a massa molecular da amostra purificada do domínio TEN e assim avaliar a formação de agregados proteicos.

3.16 Transformação de células competentes de *Pichia pastoris* KM71H

A transformação em *P. pastoris* da cepa KM71H foi realizada por eletroporação segundo o protocolo estabelecido por (CREGG, 2007). As células foram transformadas com 1µg de vetor linearizado com a enzima de restrição *PmeI* (invitrogen) e incubadas por três dias a 30°C em meio YPDS Agar com 100µ/ml de zeocina (invitrogen). Após esse tempo a placa foi examinada para a presença de colônias isoladas resistentes a zeocina.

3.17 Teste de mini-indução em leveduras

Uma colônia de *Pichia pastoris* Km71H foi utilizada para inocular 10 ml de meio *Buffered Glycerol-complex* (BMGY) [2% Peptona, 1% Extrato de levedura, 1.34% Fosfato de Potássio, pH 6.0, 100mM *Yeast Nitrogen Base* (sem aminoácidos), 0.4µg/ml Biotina, 0.5% Glicerol.] e incubada a 28°C com agitação constante até atingir OD₆₀₀ de 2-6 (16 horas aproximadamente). Foi coletada a amostra não induzida, tanto de meio

como de células. As culturas foram centrifugadas a 3000g por 5 minutos e ressuspensas a uma OD₆₀₀ final de 1 em 10ml de meio *Buffered Methanol-complex* (BMMY) [2% Peptona, 1% Extrato de levedura, 1.34% Fosfato de Potássio, pH 6.0, 100mM *Yeast Nitrogen Base* (sem aminoácidos), 0.4µg/mL Biotina, 0.5% Metanol.] para induzir a expressão do gene heterólogo.

A cada 24 horas, foram coletadas amostras tanto de meio como de células, e foi adicionado metanol 100% a uma concentração final de 0,5% v/v. As amostras foram coletadas durante 6 dias após a indução. A presença de proteína recombinante no meio foi confirmada por *Western blot*.

Tabela 1. Oligonucleotídeos utilizados neste trabalho

Oligonucleotídeo	Sequência
Fita G	5'-TTAGGGTTAGGGTTAGGGTTAGGGTTAGGG-3'
OverhangFw	5'-CTGGCCGTCGTTTTACTTAGGGTTAGGGTTAGG-3'
OverhangRv	5'-GTAAAACGACGGCCAG-3'
5' AOX	5' GACTGGTTCCAATTGACAAGC 3'
3' AOX	5' GCAAATGGCATTCTGACATCC 3'

4. RESULTADOS E DISCUSSÃO

4.1. O componente TERT de *Leishmania* é o mais basal entre os eucariotos

Foi gerado cladograma para se verificar a relação filogenética entre o TERT de *Leishmania* e os TERT de outros eucariotos. (Figura 16). As sequências utilizadas para construir o cladograma encontram-se no Apêndice 2.

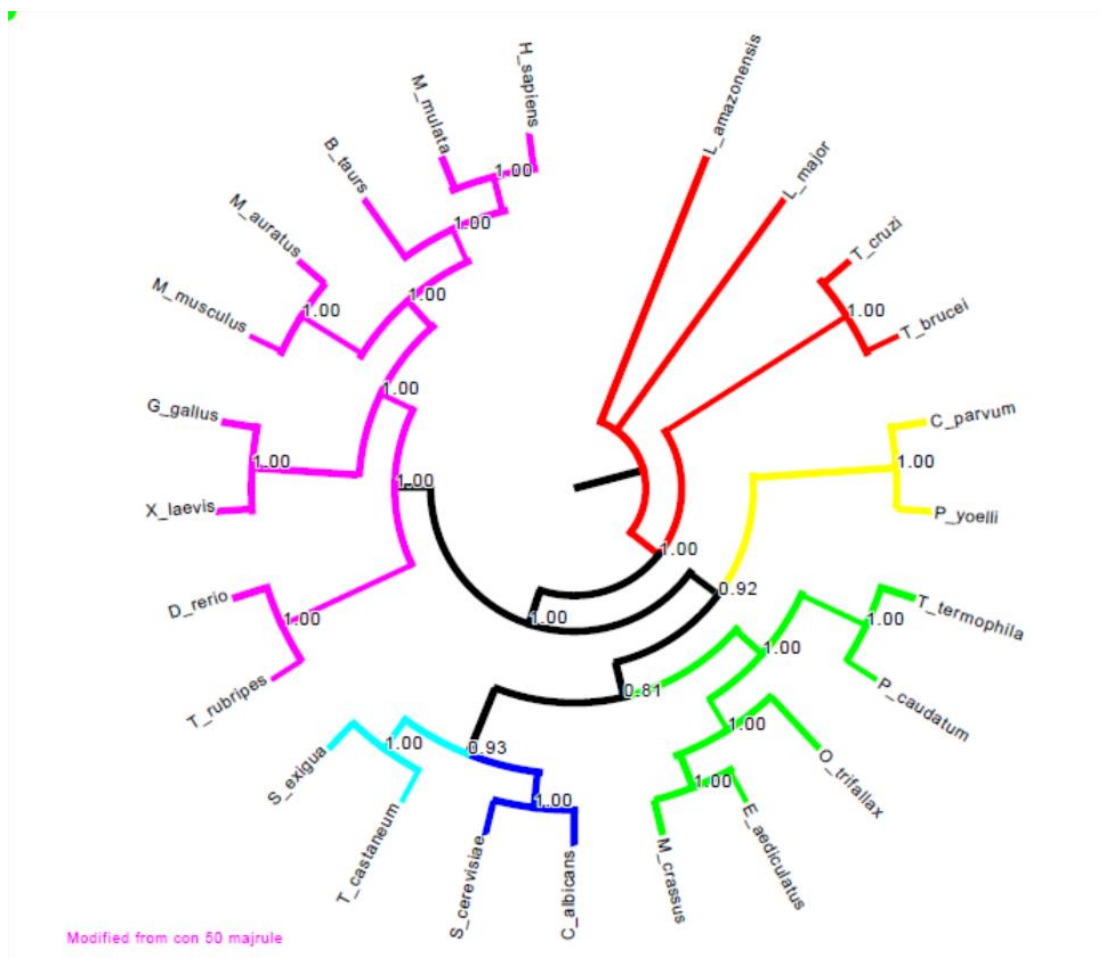


Figura 16. Cladograma dos componentes TERT de diferentes eucariotos

Cladograma gerado com o software Mr. Bayes v. 3. 1. 1. Valores de probabilidade posterior após 500.000 ciclos estão indicados nos nós. Diferentes grupos foram utilizados para se obter o cladograma (Tripanossomatídeos (vermelho); Apicomplexa (amarelo); Ciliados (verde); Fungos (azul escuro); Invertebrados (azul claro); Vertebrados (Magenta). Os componentes TERT das diferentes espécies de *Leishmania* sp. são as mais basais entre os eucariotos estudados.

Os resultados do cladograma mostram que o componente TERT de *Leishmania* sp. é o mais basal entre os eucariotos, o que está em concordância com estudos anteriores que mostravam que os componentes TERT das diferentes espécies deste gênero são bastante conservados entre si, mas são altamente divergentes quando comparados com as estruturas primárias dos TERTs de outros gêneros, mesmo quando comparados com espécies filogeneticamente mais próximas como *Trypanosoma* sp. (GIARDINI et al., 2006; DEY; CHAKRABARTI, 2018). Estes resultados indicam que os TERT de *Leishmania* apresentam características primitivas quando comparados com os TERT de outros eucariotos.

Para confirmar o grau de conservação dos componentes TERT dos Tripanossomatídeos foram feitas comparações da sequência de aminoácidos e alinhamentos múltiplos destas sequências (Apêndices 3 e 4) entre os organismos deste grupo, utilizando-se as ferramentas BLAST (SODING et al., 2005) e MUSCLE (EDGAR, 2004). Os resultados das análises de conservação mostraram que a sequência de aminoácidos do componente TERT é bastante conservado dentro do gênero *Leishmania*, mantendo uma identidade de 90% entre espécies do velho mundo, como *L. major* e *L. donovani*. Esta conservação é mantida quando se comparam espécies do subgênero *Leishmania* do velho e o novo mundo, já que o TERT de *L. amazonensis* compartilha uma identidade de aproximadamente 85% com as de *L. major* e *L. donovani*. Além disso, a conservação da sequência do componente TERT de *Leishmania* sp. é mantida mesmo entre subgêneros já que *L. panamensis* (subgênero *Viannia*) compartilha identidade de 76-78% quando comparado com as espécies do subgênero *Leishmania* mencionadas anteriormente (Tabela 2). A conservação da identidade na estrutura primária do componente TERT é perdida quando se compara o gênero *Leishmania* com o gênero *Trypanosoma*, evidenciado por uma queda nos valores

de identidade para entre 31 e 38% (Tabela 2). Interessantemente, a identidade do componente TERT entre espécies do gênero *Trypanosoma* é de aproximadamente 50% (Tabela 2), o que pode indicar que houve uma maior pressão seletiva para conservar a estrutura primária do componente TERT em *Leishmania* do que em *Trypanosoma*.

Tabela 2. Comparação entre as sequências primárias dos componentes TERT de protozoários dos gêneros *Leishmania* e *Trypanosoma*

	Identidade (%)					
	L. major	L. donovani	L. amazonensis	L. panamensis	T. cruzi	T. brucei
L. major	100	90	86	78	38	35
L. donovani		100	87	78	37	31
L. amazonensis			100	76	36	32
L. panamensis				100	36	31
T. cruzi					100	48
T. brucei						100

Tentativas de se comparar toda a estrutura primária do TERT de *Leishmania* com eucariotos mais distantes e mais derivados, como ciliados, leveduras e humanos, não foram bem sucedidas dado o baixo grau de conservação da estrutura primária das telomerasas em geral e também o grande tamanho das proteínas desta família (MASON et al., 2011). Para poder fazer estas comparações, foi necessário usar frações menores ou domínios individualizados da proteína. Os alinhamentos feitos em regiões conservadas foram refinados a partir dos dados de GIARDINI et al., (2006) e serão discutidos em detalhe em seções posteriores.

As comparações de estrutura primária entre diferentes organismos e o componente TERT de *L. major* permitiram delimitar os 4 domínios estruturais nesta proteína (Figura 17). Dada a alta conservação da sequência primária do componente TERT no gênero *Leishmania*, pode-se assumir que os resultados obtidos para uma espécie podem ser aplicados a todas as espécies do gênero cujas sequências já foram anotadas.

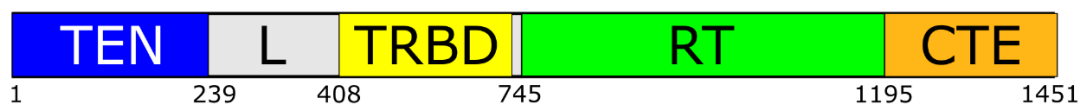


Figura 17. Delimitação dos domínios estruturais na estrutura primária de *L. major*

Os domínios estruturais do TERT de *L. major* foram delimitados por alinhamentos de sequência múltiplos e são mostrados na figura como uma representação gráfica. Abaixo se encontra o número de aminoácido no qual começa e acaba cada domínio.

Após a delimitação dos domínios do componente TERT de *L. major*, eles foram analisados individualmente, como será mostrado nas seções a seguir.

4.2 Caracterização estrutural do domínio TEN do TERT de *Leishmania* sp.

Alinhamentos de sequências de aminoácidos entre a região amino-terminal do TERT de *L. major* (aminoácidos 1-408) com as regiões amino-terminais dos TERT de diferentes organismos foram realizados inicialmente com o intuito de se verificar o nível de conservação dessa região entre os diferentes eucariotos. O melhor alinhamento obtido foi entre as sequências de *L. major* e as sequências dos TERT de diferentes vertebrados (Figura 18). Como reportado anteriormente (GIARDINI et al., 2006), o nível de conservação deste domínio é muito baixo. A região que compreende os primeiros 120 aminoácidos da porção amino-terminal, apresenta inserções próprias do gênero *Leishmania* (Figura 18). Já a região entre os aminoácidos 120 e 230 é a que apresenta o maior nível de conservação em termos de estrutura primária. Esta região entre os aminoácidos 120 e 230, contém o motivo GQ, o único motivo dentro do domínio TEN conservado entre os diferentes organismos. O motivo GQ encontra-se na região carboxi-terminal do domínio TEN (BOSOY et al., 2003; MORIARTY et al., 2004; JACOBS et al., 2006), e é possível se verificar que os três aminoácidos invariáveis deste motivo, que correspondem aos resíduos Glicina 192 (G-192), Glutamina 217 (Q-217) e Glicina 220 (G-220), são conservados em *L. major* (Figura

18). Em *T. thermophila* estes três aminoácidos possuem importância estrutural e são indispensáveis para a interação com o DNA genômico (JACOBS et al., 2006).

```

Imaj : MSASFPSIPGFAGPLSLKAFLEEYFGLHLTFVAVETASPSERAATAATPSAAEFRAIRDVVLPPNQSELVVVYVALHASSPPPTTAHASPTPPTPALGRAASATGFERLR : 112
Hsap : -MFRAPRCR-----FVRSLLRSHYREVLPPLATFVRRRLGPOGWRLVORGDFAAFRAI-----VACCLVCPVWLDARPPPAAPSERQVSSCLKEL-----VARVVC : 86
Mmus : -MTRAPRCR-----FVRSLLRSRYREVWPLATFVRRRLGPEGRRLVORGDGKIYRTL-----VACCLVCMHWGSCPPPADLSFHQVSSCLKEL-----VARVVC : 86
Maur : -MFRAPRCR-----FVRALLRSYRQVWPLATFVRRRLGPEGRQLVORGDGKVFRTL-----VRCCLVCPVWLDSCPPPADLSFHQVSSCLKEL-----VARVVC : 86

Imaj : PLTHQTVASSAHDTMTRONPADRONSSSSSSFAKHKSTRDPRASAFRLLYTNTSHRPLTDALLRHPWASFAACLGPAANGFIEMYQPIVIOLEAMAGGVQVIGPALK : 224
Hsap : RL-----CERGAKNVLAFGFALLDGAFFGGPEEAF--TTSVRSYLENT---VTDALRGSGAWGLLIRRVGDDVLVHLLARCALFV--LVAPSCAYQVGGSELY : 176
Mmus : RL-----CERNERNVLAFGFELLNFAFGGPEMAF--TTSVRSYLENT---VIETLRVSGAWMLLISRVGDDLLVYLLAHCALMI--LVPPSCAYQVGGSELY : 176
Maur : RL-----CERGERNVLTGFGALLNGACGGPEMTF--TTSVRSYLENS---VTESLRVSGAWMLLINRVGDDLLVYLLARCALMI--LVPPSCAYQVGGSELY : 176

Imaj : HTA-----FERSESAAP-----SSEMKGPIPOGSGCISSRGVKRAIDTRECSEPLPQKQRRVEAEPKAKAGRRLCRGVAEDSRPCRDFSSSVNR : 312
Hsap : QLGAATQAREPPHAS--GPRRLG-----CERAWNHSVREAGVPLG---LEALGARRRGGSSASPSLPLP--KRFRRGAFFEPERTVGGCGSWAHPGRTRGPSDRGFVVS : 274
Mmus : QICATTDIWEVSASVSRPTRPVGRNFTNLRFLQCIKSSSRCEAPKPLA---LPSRGTKRHISLTSTSVESA--KKARCYPVPRVEEGPHRCVLTPTSCSKSWVPSARSPEVP : 283
Maur : QICATAETWESVSRIRPTRPVGRNFTHLGSTRVNRNSSHQEAWKPPP---LPSREAKRSISITNRSVPS--KKARCDLAPRLEKGBYRCAVPTESDKTWVENPAKSHAVP : 283

Imaj : AAMAPTWAAALRTDVPTRRLYSARISDGDGAGDGS-----GS-LPLPAGETEAQWL-----ERPHSHLHRAICAPL-----PARRAY----- : 385
Hsap : PARPAE-EATSELEGALSGTRHSHPSVGRCPHAGPPSTSRPWPDTPCPPVYAETKHFYLYSSGD--KECLRPSFLLSLRSPLTGARRLVETIFLCSRPWMPGTPRRLRRI PQ : 384
Mmus : ---TAE-KDLSKGVSDLSL--SGSVCKEKPSSTSLLSFPRQNAFQLRE--FLETIRHFLYSRGDGCERLNPSSELLSNLOHNLIGARRLVEIIFLCSRPRTSGPLCRTHRLSR : 389
Maur : ISRTTK-EDLSSGVKAPGLSR--SGSVCYKPKPSSTSLQSELQNAFQLRE--YLETIRFLYSREGGRERLNPSSELLNQLQPSLTGARRLVEIIFLCSRPRTSGPLCGRRRLS- : 391

```

Figura 18. Alinhamento múltiplo de seqüências entre os domínios TEN dos TERT de *Leishmania* e mamíferos

O alinhamento de seqüências mostra uma baixa conservação na seqüência primária do TEN de *L. major* quando comparado ao dos mamíferos. Na região amino-terminal da seqüência de *Leishmania* existem diversas inserções (triângulos) características do gênero, enquanto a região carboxi-terminal mostra uma maior conservação. Os três aminoácidos invariáveis do motivo GQ (G-192; Q-217; G-220) estão destacados com asteriscos. (Lmaj: *L. major* (AAO67514.1); Hsap: *Homo sapiens* (NP_937983.2); Mmus: *Mus musculus* (NP_033380.1); Maur: *Mesocricetus auratus* (NP_001268556.1)) Resíduos destacados em preto mostram conservação em todos os organismos e em cinza, conservação na maioria dos organismos.

Para melhor se evidenciar o grau de conservação entre os domínios TEN do TERT de *L. major* e o domínio TEN de outros eucariotos cujas estruturas terciárias já foram resolvidas (JACOBS et al., 2006; POLSHAKOV et al., 2016), foram realizadas tentativas de predição de estrutura terciária utilizando-se os programas Phyre2 (KELLEY et al., 2015) e HHpred (SODING et al., 2005). Porém, devido à baixa conservação da estrutura primária compartilhada entre os mesmos e os espúrios valores de E-value obtidos, não foi possível se obter a predição de estrutura terciária. Conforme mostrado na Tabela 3, a quantidade de aminoácidos modelada utilizando os diferentes programas é muito baixa (menos do que 50 aminoácidos), especialmente no caso da predição feita por Phyre2, onde cabe destacar que a estrutura usada para gerar o modelo não corresponde a uma telomerase e que o *E-value* é muito alto (Tabela 3).

Tabela 3. Homologia estrutural entre o domínio TEN de LmTERT com estruturas cristalográficas depositados em PDB

Domínio TEN					
	Região da LmTERT	Identidade	Score	Modelo homólogo	E-value
HHpred	163-177	15%	40,42	Telomerase reverse transcriptase. <i>T. thermophila</i> (2B2A)	0,64
Phyre	156-200	23%	18,03	Putative Transcriptional Regulator. <i>R. jostii</i> . (2G7G)	15

Na tentativa de se compensar o insucesso na obtenção do modelo estrutural do TEN de *L. major*, foi realizada a predição *in silico* da estrutura secundária dos primeiros 408 aminoácidos do componente TERT de *L. major*, onde está contido o domínio TEN. Esta predição foi realizada usando o algoritmo GOR V (Figura 19) e indicou que o domínio TEN do componente TERT de *L. major* pode apresentar uma estrutura enriquecida em regiões desenoveladas, já que prediz 56,6% de elementos desordenados e 28,2 % de hélices- α na constituição do mesmo.

Existem duas estruturas cristalográficas disponíveis para o domínio TEN, o de *T. thermophila* (JACOBS et al., 2006) e o de *O. polymorpha* (POLSHAKOV et al., 2016). Análises de conservação de estrutura primária entre o domínio TEN de *L. major* e os domínios TEN destes dois organismos mostram que não há uma identidade de sequência significativa (~ 12%) e os valores de E-value são espúrios, impedindo fazer comparações de conservação da estrutura primária. No entanto, quando se comparam os resultados obtidos de predição de estrutura secundária para o domínio TEN de *L. major* (Figura 19) com as estruturas tridimensionais resolvidas para *T. thermophila* (JACOBS et al., 2006) e *O. polymorpha* (POLSHAKOV et al., 2016) é possível perceber que a estrutura de *O. polymorpha* (Figura 6A) apresenta um alto conteúdo de hélices- α e regiões desenoveladas, similar ao predito para *L. major*. Enquanto a estrutura de *T. thermophila* (Figura 6B), apresenta poucas regiões desenoveladas e um alto conteúdo de hélices α e folhas- β . Isto sugere que o domínio TEN de *Leishmania* é estruturalmente mais parecido com o de *O. polymorpha* do que com o de *T. thermophila*. A relevância biológica deste fato deve ser resolvida posteriormente.

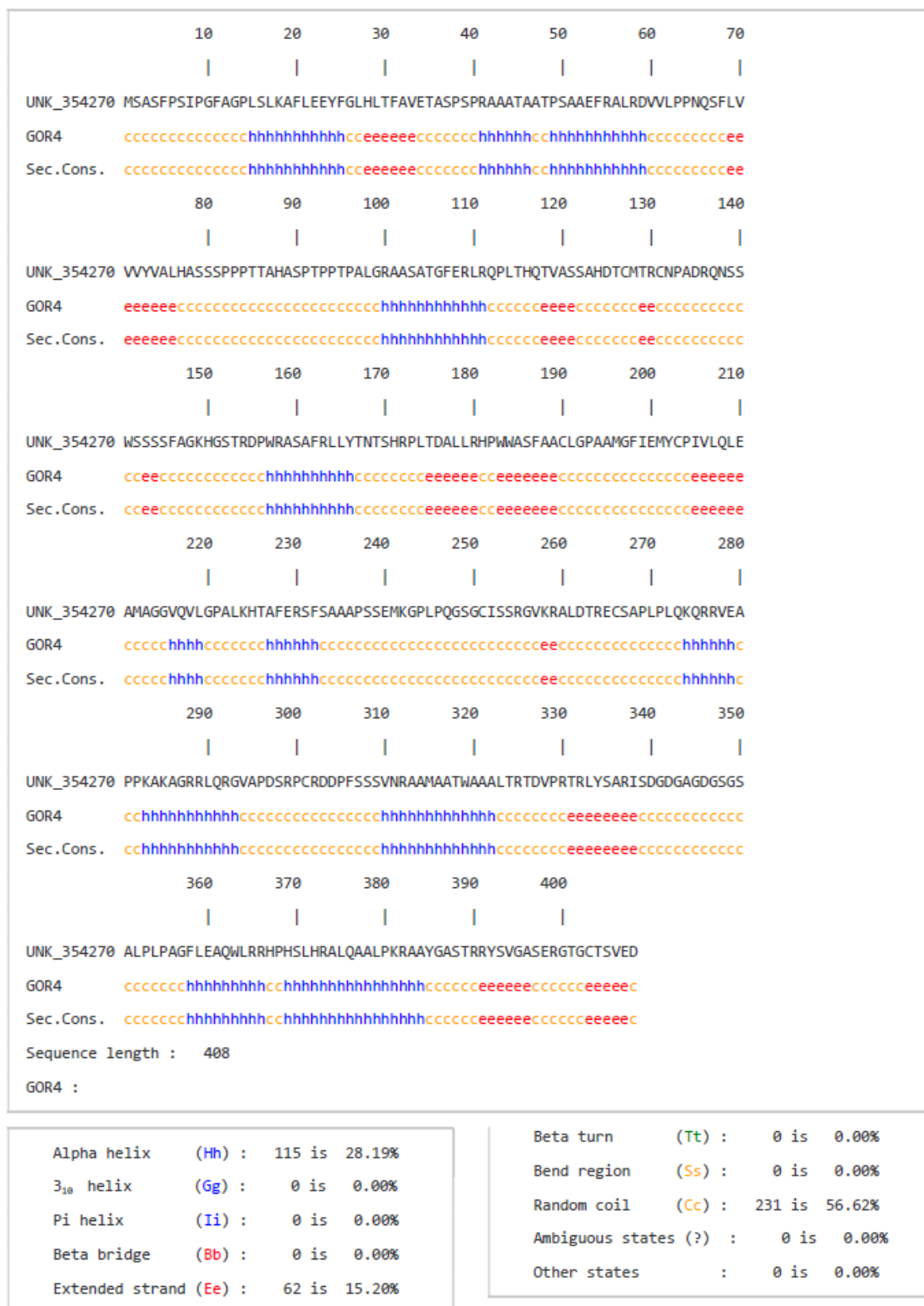


Figura 19. Predição de estrutura secundária do domínio TEN de *L. major*

Representação da estrutura secundária do domínio TEN de *L. major* predita usando o algoritmo GORV. A predição de estrutura secundária mostra principalmente a presença de elementos desordenados (56,62%) e hélices- α (28,19%).

4.3 Obtenção do domínio TEN recombinante e estudos estruturais com este domínio

Os extratos obtidos após realizar o teste de mini-indução e solubilidade foram fracionados em SDS-PAGE 15% corado com azul de comassie (Figura 20). No gel, observa-se que o domínio TEN está sendo expresso em corpúsculos de inclusão, já que a banda de massa molecular esperada, que começa ser evidenciada após uma hora de indução está exclusivamente presente na fração insolúvel do extrato proteico bacteriano (Figura 20 a partir de uma hora de indução).

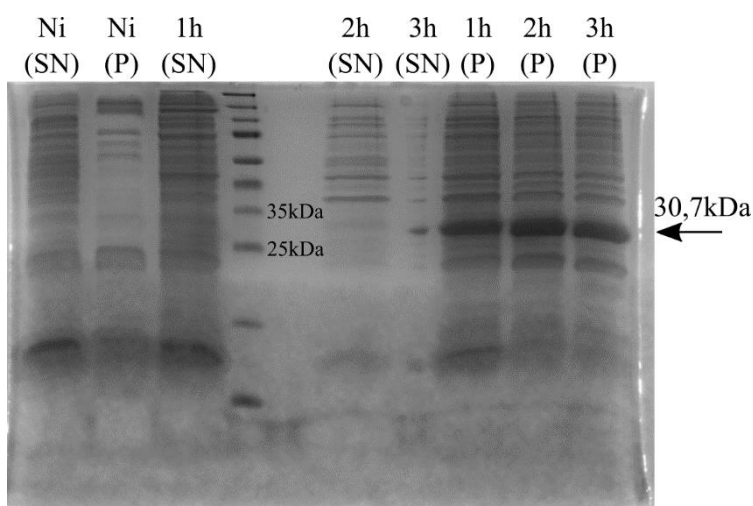


Figura 20. Teste de mini-indução de bactérias *E. coli* BL21 DE3 códon plus transformadas com pET28a-domínio TEN

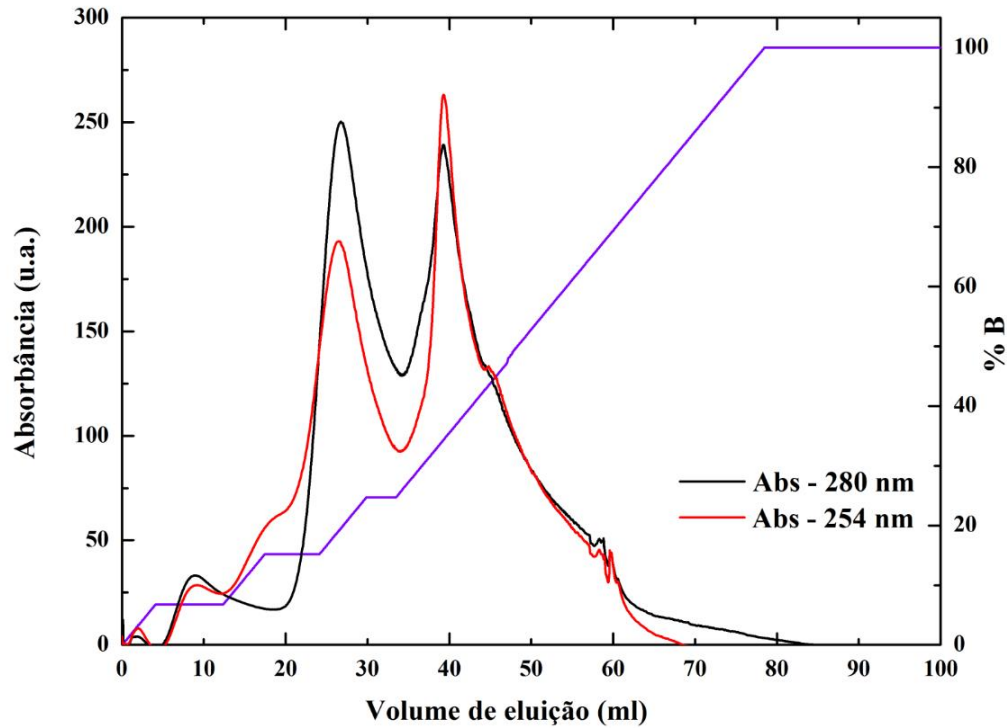
Gel de poliacrilamida 15% corado com azul de comassie, mostrando extratos proteicos de bactérias *E. coli* DE3 códon⁺ transformadas com pET28a-TEN. As amostras do gel correspondem a extratos não induzidos (N.I) e depois de 1, 2 e 3 horas de indução. Para cada amostra foi obtida a fração solúvel (SN) e insolúvel (P) do extrato bacteriano. M.P.M: Marcador de peso molecular Spectra Multicolor Broad Range Protein Ladder (Thermo Fisher scientific). A seta aponta a banda com massa esperada correspondente ao domínio TEN recombinante (30,7kDa) O resultado mostra que o domínio TEN está sendo expresso na fração insolúvel do extrato proteico bacteriano.

A experiência do nosso grupo de pesquisa mostra que as proteínas de *Leishmania* quando expressas em sistema heterólogo bacteriano tendem a permanecer em corpúsculos de inclusão (LIRA; SIQUEIRA NETO, DE; et al., 2007a; LIRA; SIQUEIRA NETO; et al., 2007b; SIQUEIRA-NETO et al., 2007; LIRA et al., 2009; SILVA, DA et al., 2010; PAVANI et al., 2014). Porém, há vários antecedentes que

mostram que estas proteínas podem ser purificadas em condições desnaturantes, com alta concentrações de ureia e posteriormente re-enoveladas por diálise escalonada (LIRA; SIQUEIRA NETO, DE; et al., 2007; LIRA; SIQUEIRA NETO; et al., 2007; LIRA et al., 2009; PAVANI et al., 2014). As proteínas purificadas nestas condições têm sido expressas em larga escala, e obtidas com alto grau de pureza. Adicionalmente as proteínas re-enoveladas têm sido usadas com sucesso em estudos bioquímicos e estruturais (SIQUEIRA-NETO et al., 2007; LIRA et al., 2009; SILVEIRA, DA et al., 2013; PAVANI et al., 2014). Por isso foi decidido purificar o domínio TEN a partir da fração insolúvel realizando pequenas modificações a partir de um protocolo já disponível para purificação deste tipo de proteínas (LIRA et al., 2009).

O primeiro passo para se obter o TEN recombinante foi purificar a proteína em condições desnaturantes utilizando-se cromatografia de troca aniônica (HiTrap CM Sepharose FF) em um gradiente linear de 20-500 mM de NaCl (Figura 21A). O cromatograma apresenta um primeiro pico que se inicia quando a porcentagem de tampão B é de 15% (~150mM NaCl) e um segundo pico que se inicia quando a porcentagem de tampão B é de aproximadamente 25% (~250mM NaCl) (Figura 21A). As frações eluídas nessa etapa de purificação, correspondentes aos picos do cromatograma, foram avaliadas em gel SDS-PAGE 12 % (Figura 21B). Os géis mostram que uma única banda com massa molecular similar à esperada para o domínio TEN recombinante (30,7 kDa) é eluída numa porcentagem entre 12 e 40% de tampão B (Figura 21B frações 23-56). Estes resultados indicam que a purificação do domínio foi bem-sucedida utilizando-se coluna de troca iônica.

A



B

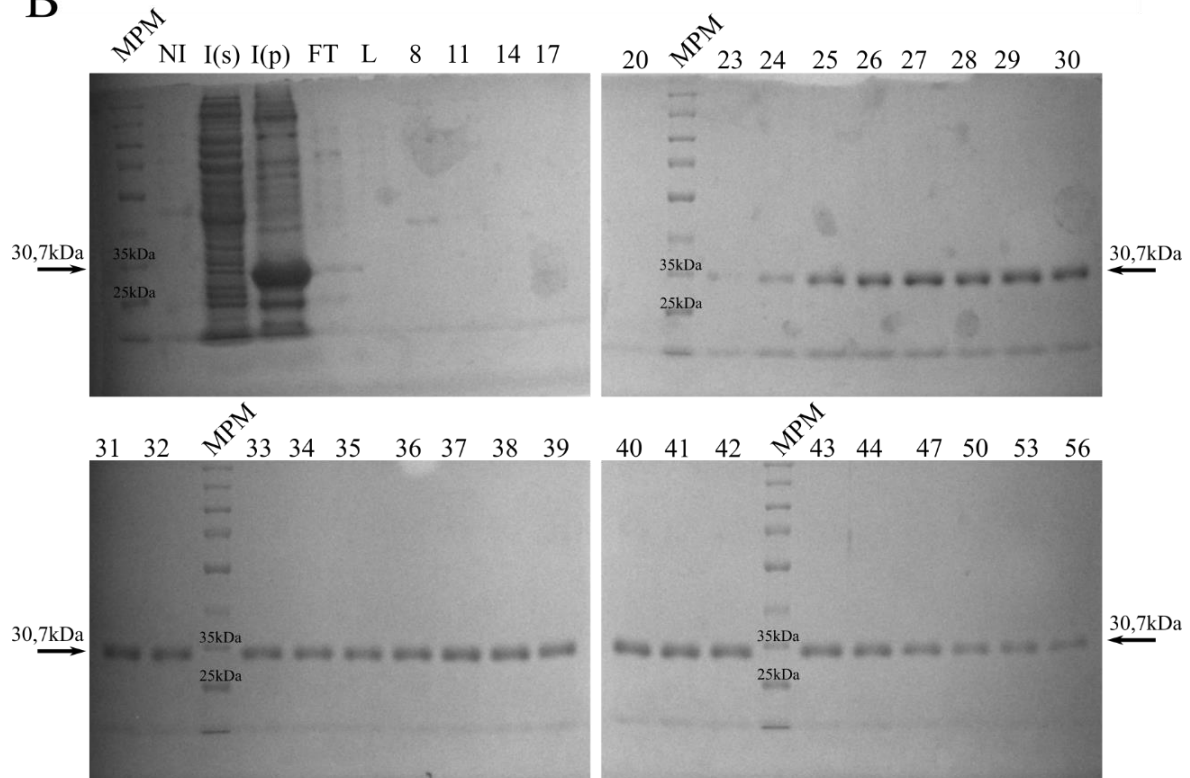


Figura 21. Purificação do domínio TEN do TERT de *L. major* a partir da fração insolúvel do extrato proteico de bactérias *E. coli* DE3 códon plus, por cromatografia de troca aniônica

(A) Cromatograma da purificação, mostrando os picos de absorvância a 280 (preto) e 254 (vermelho) eluídos no gradiente com tampão B (azul). Os picos se apresentam numa porcentagem de 15 a 40% de Tampão B (150-400mM). (B) SDS-PAGE 12% corados com azul de Coomassie contendo as frações eluídas da coluna. MPM: marcador de peso molecular Spectra™ Multicolor Broad Range Protein Ladder (Thermo Fisher Scientific), NI: extrato proteico de bactérias não induzidas, I(s): fração solúvel de extrato proteico de bactérias induzidas, I(p): fração insolúvel do extrato proteico de bactérias induzidas, FT: *flowthrough*, proteínas que não ligam na resina, Lav: proteínas eluídas na lavagem da coluna, 8-56: número do eluato segundo o gráfico mostrado em (A). Nas frações correspondente aos picos observados no cromatograma, pode se observar uma única banda com peso molecular esperado, correspondente ao domínio TEN recombinante purificado.

Tendo se obtido a proteína purificada em condições desnaturantes, foi necessário proceder com seu re-enovelamento. O processo de re-enovelamento é muito importante, pois para se realizar estudos bioquímicos e estruturais é indispensável se ter a proteína corretamente enovelada, já que a estrutura está diretamente relacionada com a função. O re-enovelamento foi realizado por diálise escalonada para a retirada completa da uréia e do β -mercaptoetanol presentes nos tampões de purificação. Em estudos anteriores (LIRA et al., 2009), foi mostrado que a adição de heparina sódica durante a diálise de proteínas que se ligam a DNA auxilia no correto enovelamento e evita a formação de agregados proteicos. Esta estratégia tem sido usada com sucesso na obtenção de outras proteínas do complexo telomérico de *Leishmania* que são expressas na fração insolúvel do extrato bacteriano e que interagem com DNA telomérico especificamente (LIRA et al., 2009; PAVANI et al., 2014). Por isso, esta estratégia também foi utilizada durante o re-enovelamento do domínio TEN, porém, percebeu-se que o processo de re-enovelamento deste domínio não era alterado pela presença de heparina sódica (dados não mostrados).

Ensaio de espectroscopia de dicroísmo circular e de fluorescência estática utilizando diferentes comprimentos de onda foram realizados para se verificar o estado de enovelamento da proteína após a diálise (Figura 22).

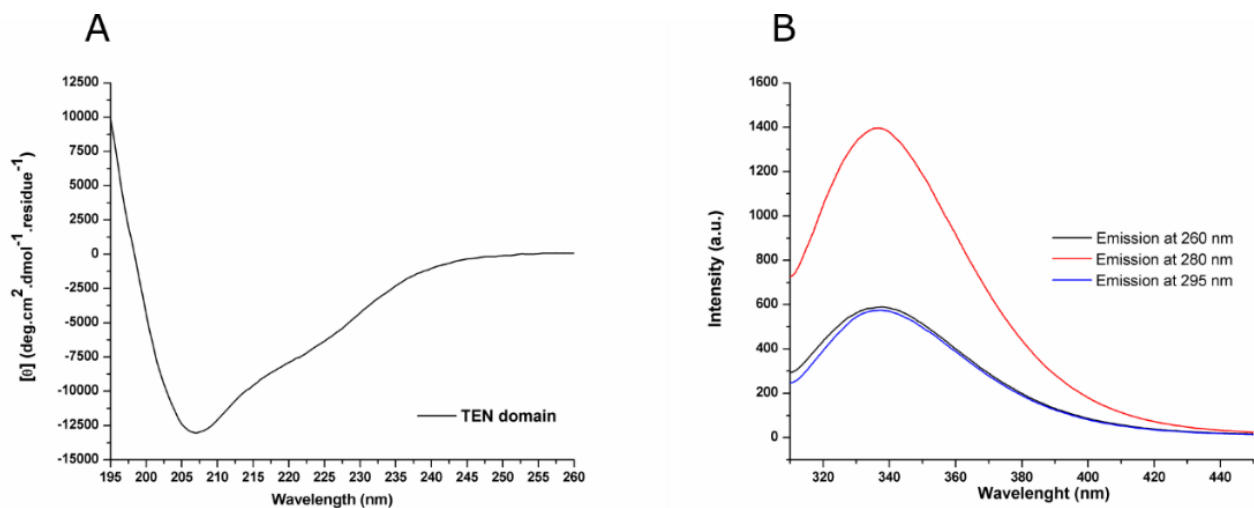


Figura 22. Avaliação do estado de enovelamento do domínio TEN purificado

(A) espectro de dicroísmo circular do domínio TEN purificado. A análise de desconvolução de dicroísmo circular mostra que o domínio TEN purificado é uma proteína composta por aproximadamente 70% de elementos desenovelados e hélices- α o que está em concordância com a predição de estrutura secundária obtida pelo programa GORV (SEN et al., 2005). (B) Espectro de fluorescência utilizando diferentes comprimentos de onda de excitação. Os espectros de fluorescência sugerem que os triptofanos da proteína estão num ambiente mais hidrofóbico em comparação com os resíduos de tirosina, o que é esperado para proteínas enoveladas ou parcialmente enoveladas. Os resultados obtidos sugerem que após o renovelamento, o domínio TEN possui uma estrutura terciária similar à estrutura nativa.

O espectro de dicroísmo circular do domínio TEN apresenta um valor mínimo de $[\theta]$ a 207 nm (Figura 22A) indicando que a proteína possui grandes quantidades de elementos desordenados combinados com hélices- α . De fato, a desconvolução da curva de CD do domínio TEN que indica que 70% destes dois elementos estão presentes na composição da estrutura secundária deste domínio, corrobora a predição *in silico* de estrutura secundária mostrada anteriormente (Figura 19).

O espectro de fluorescência estática obtido a partir do domínio TEN, mostra que a intensidade de fluorescência emitida é maior no comprimento de onda de excitação a 280 nm, já que nesse comprimento de onda os resíduos de tirosina e triptofano são excitados juntamente. Quando utilizado o comprimento de onda de excitação de 295 nm, que é capaz de excitar apenas os resíduos de triptofano, o máximo de intensidade fluorescente é observada a 333 nm, ao contrário dos comprimentos de onda (260 e 280 nm) que excitam também as tirosinas, e o máximo de intensidade fluorescente

observada é a 339 nm. Estes dados indicam que os resíduos de triptofano encontram-se em um ambiente mais hidrofóbico que os resíduos de tirosina, como é esperado em proteínas enoveladas ou parcialmente enoveladas (Figura 22) (MUNISHKINA; FINK, 2007).

Para confirmar que de fato o domínio TEN havia sido purificado, foi realizado um ensaio de *western blot*, utilizando-se como anticorpo primário um anticorpo anti-HIS (Bio-Rad), que reconhece a cauda de histidina fusionada à região amino-terminal do domínio TEN heterólogo (Figura 23).

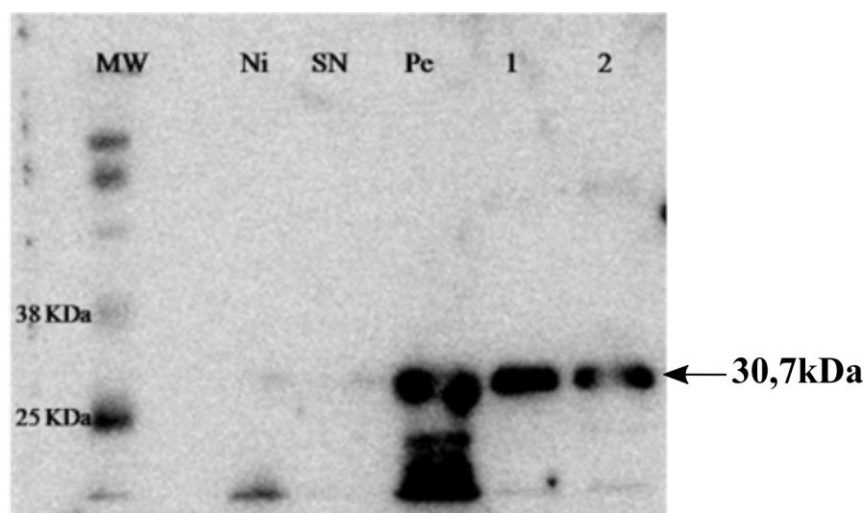


Figura 23. Western blot confirmando a purificação do domínio TEN recombinante

Western blot para confirmar a identidade do domínio TEN recombinante. MW: marcador de peso molecular Western Sure® Pre-stained Chemiluminescent Protein Ladder (Li-cor), Ni: extrato proteico de bactérias não induzidas, SN: fração solúvel do extrato proteico de bactérias induzidas, Pe: fração insolúvel do extrato proteico de bactérias induzidas, 1 e 2 amostras da proteína purificada. A banda reconhecida pelo anticorpo (30,7kDa), está apontada por uma seta e aparece exclusivamente na fração insolúvel do extrato induzido e amostra de proteína purificada, mostrando que corresponde ao domínio TEN da telomerase de *L. major*.

O resultado do *western blot* confirma que a banda purificada, de aproximadamente 30 kDa é efetivamente a proteína heteróloga que contém a cauda de histidina, já que a marcação só aparece no pellet do extrato induzido e nas frações

correspondentes à proteína purificada (Figura 23). Com esse último resultado concluiu-se que a proteína purificada é o domínio TEN da telomerase de *L. major*.

Após a obtenção do domínio TEN purificado e enovelado, foi testada sua capacidade de interação com o DNA telomérico. Para isso foram utilizados oligonucleotídeos sintéticos contendo as sequências da fita rica em G e um DNA dupla-fita parcial (3' *G-Overhang*) (Tabela 1) (Figura 2). Os testes foram realizados por espectroscopia de dicroísmo circular e de fluorescência estática (Figura 24).

Os espectros de dicroísmo circular da proteína purificada e da proteína purificada incubada com a fita G telomérica ou com o 3' *G-Overhang* apresentam uma superposição quase completa (Figura 24). A espectroscopia de fluorescência, comparando a proteína purificada e a proteína purificada interagindo com o 3' *G-Overhang* (Figura 24C) mostra que os espectros de emissão fluorescente a 280 nm da proteína, na presença e na ausência de DNA, são praticamente idênticos. Estes dados sugerem que não há interação entre o domínio proteico e o DNA, ou, que se há interação esta não altera a estrutura secundária da proteína. Estudos anteriores realizados com proteínas de *Leishmania* que interagem com DNA telomérico mostram que esta interação muda a conformação estrutural das proteínas de tal forma que esta mudança pode ser observada nos ensaios de espectrometria de dicroísmo circular (PAVANI et al., 2014; MOREA et al., 2017). Isto sugere que nos ensaios apresentados na Figura 24, não se observa interação entre o domínio TEN de *L. major* e o DNA telomérico.

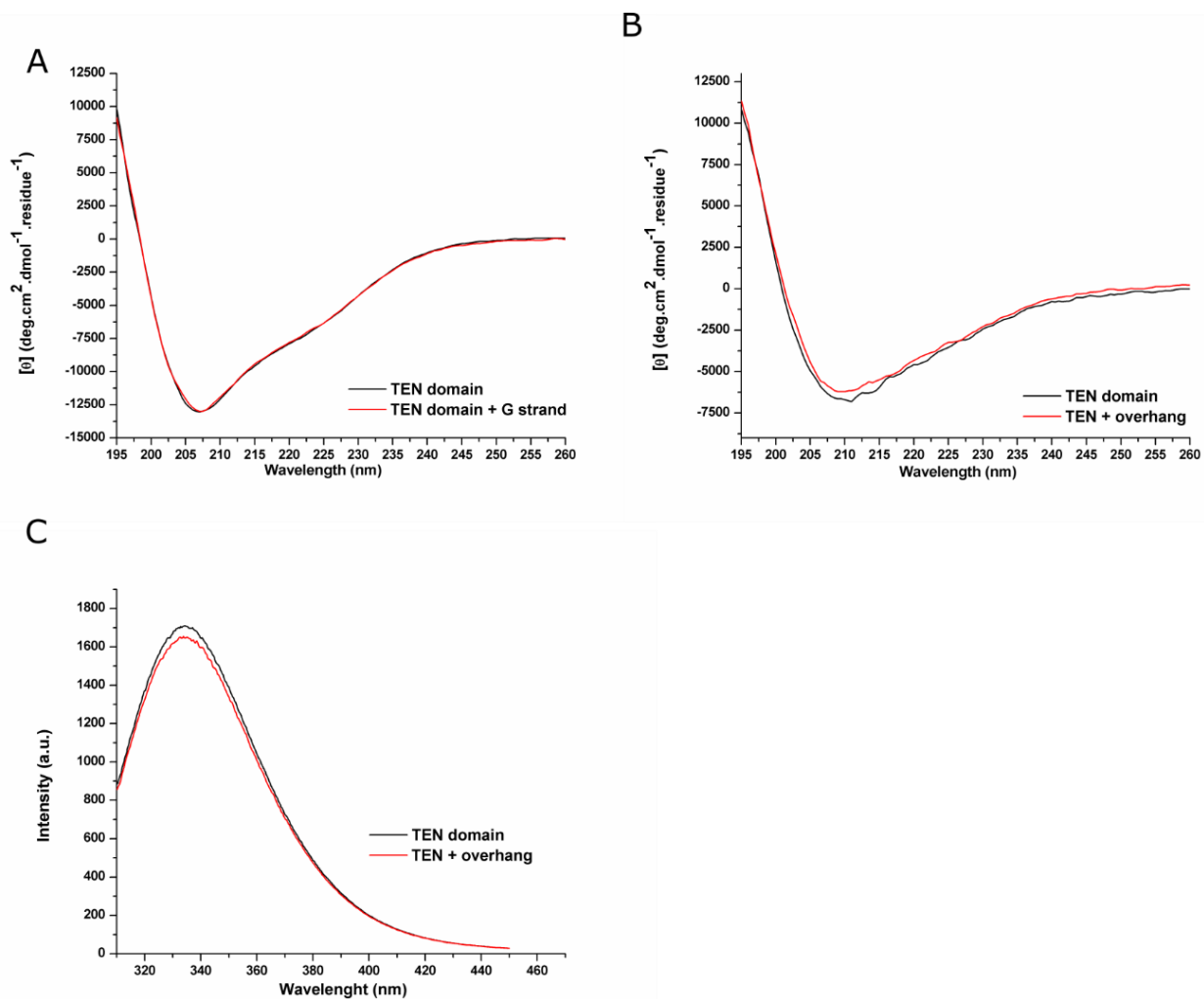


Figura 24. Espectros de Dicroísmo circular e fluorescência estática do domínio TEN purificado, na presença de DNA telomérico

(A-B) Comparação do espectro de dicroísmo circular entre o domínio TEN purificado e o domínio TEN com a fita G telomérica (A) e com o 3'G-Overhang (B). Os resultados do dicroísmo circular mostram que os espectros são virtualmente idênticos. (C) Comparação entre os espectros de fluorescência do domínio TEN purificado e o domínio TEN incubado com o duplex parcial telomérico (3'G-Overhang), neste caso os espectros também são virtualmente idênticos. Os resultados obtidos por dicroísmo circular e espectrometria de fluorescência sugerem que o domínio TEN da telomerase de *L. major* não interage com o DNA telomérico.

A interação entre o domínio TEN e o DNA telomérico tem sido amplamente descrita em outros organismos (FRIEDMAN; CECH, 1999; XIA et al., 2000; BOSOY et al., 2003; JACOBS et al., 2006), já que o domínio TEN tem sido caracterizado como o adaptador entre o DNA telomérico e o restante do complexo ribonucleoproteico da telomerase (JIANG et al., 2015). Por esta razão, os resultados aqui mostrados que sugerem que não há interação entre o domínio TEN do TERT de *L. major* e os DNAs

teloméricos são inesperados. Contudo, diversas explicações para este dado são possíveis, além da inexistência verdadeira de uma interação entre ambos. A primeira é que o domínio TEN talvez necessite de outros domínios do TERT de *L. major* para posterior interação com o DNA telomérico, ou ainda, de outra proteína telomérica para atingir a conformação que lhe permite se associar ao DNA telomérico. Esta hipótese se mostra interessante ao considerar que o domínio TEN isolado é um fragmento proteico enriquecido em regiões desenoveladas. Também é possível que a proteína não esteja enovelando corretamente durante a diálise.

Uma forma de verificar se o processo de re-enovelamento está sendo efetivo é avaliar se a proteína está formando agregados proteicos, já que a formação de agregados é comum quando a proteína não consegue voltar a sua estrutura terciária nativa durante o re-enovelamento (PATRA et al., 2000; VALLEJO; RINAS, 2004; SINGH; PANDA, 2005; SINGH et al., 2015). Por este motivo, antes de se realizar mais testes de interação com o domínio TEN foi decidido realizar uma cromatografia analítica de exclusão por tamanho, para avaliar o estado de agregação da proteína depois do re-enovelamento.

Para isto, foi utilizada uma coluna Superdex 75 10/300 (GE Healthcare) equilibrada em tampão HEPES 40mM pH 7,4, KCl 100mM acoplada a um sistema de cromatografia líquida (HPLC; AKTA PurifierTM) (Figura 25). Nestas condições, o volume de eluição teórico esperado calculado a partir da equação mostrada nos materiais e métodos para o domínio TEN recombinante, cujo peso é de aproximadamente 30 kDa é de 11,19 ml. Porém, observamos que a amostra foi eluída muito antes, em três picos com volumes de eluição de 6,89 ml, 7,78 ml e 8,91 ml respectivamente, os quais encontram-se muito próximos do volume morto da coluna (7,87ml), indicando que a proteína está formando agregados de elevada massa

molecular. Uma pequena fração da amostra foi eluída num volume similar ao volume de eluição esperado para o domínio TEN monomérico (Figura 25A).

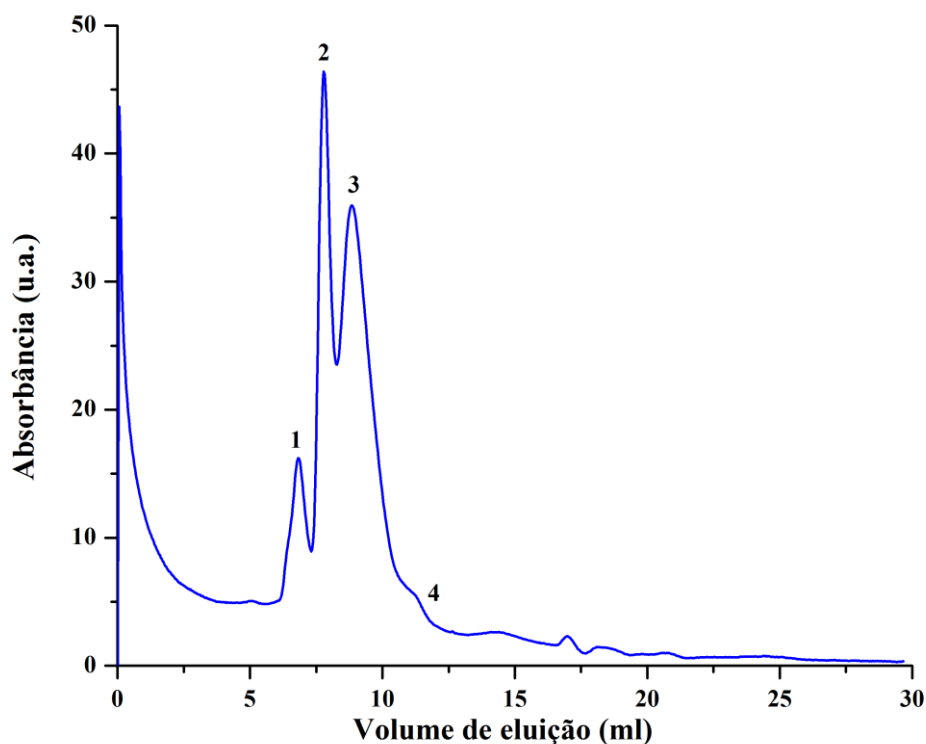


Figura 25. Cromatografia analítica de exclusão por tamanho do domínio TEN

Picos de eluição do domínio TEN purificado passando pela coluna de exclusão molecular. A amostra foi eluída em 3 picos de volume diferente, destacado com números de 1-4 na figura. O pico 1 corresponde a um volume de eluição de 6,89ml, o pico 2 corresponde a um volume de eluição de 7,78ml e o pico 3 corresponde a um volume de eluição de 8,91ml. Todos estes volumes estão abaixo do volume de eluição estimado para o domínio TEN monomérico e por tanto mostram que a proteína está formando agregados. O número 4 na figura mostra o volume de eluição esperado para o domínio TEN monomérico (11,19 ml), há um “ombro” no gráfico que sugere que uma fração minoritária da amostra encontra-se na forma monomérica.

Ao se analisar os três picos obtidos, surgiu a hipótese de que nem todos os complexos de alto peso molecular se devem a agregados proteicos, mas somente o pico de eluição de 6,89 ml. Os outros dois picos podem corresponder a contaminação com ácidos nucleicos bacterianos. Um indício de que esta hipótese pode ser verdadeira, é a alta absorbância a 256 nm observada no cromatograma da purificação (Figura 21A). Caso esta hipótese seja verdadeira, ela explicaria também por que não há uma interação

aparente entre o DNA telomérico de *L. major* e o domínio TEN, pois a amostra proteica já estaria saturada de DNA.

A fim de se comprovar esta hipótese, os extratos bacterianos correspondentes ao *input* da purificação foram fracionados em gel de agarose 1%, mostrando que havia uma grande quantidade de DNA na mistura proteica que era aplicada na coluna (Figura 26A).

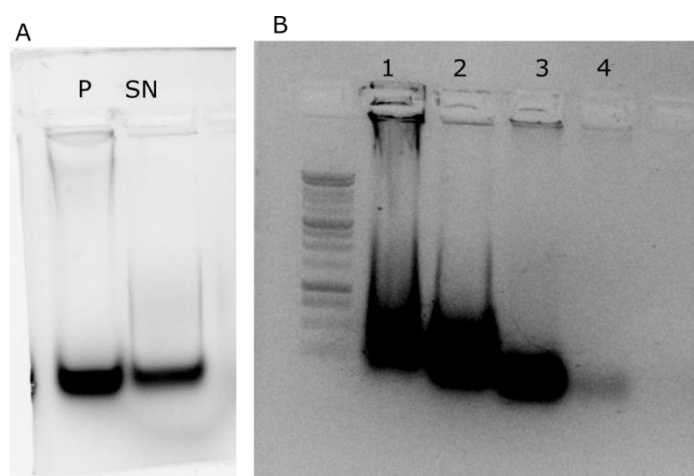


Figura 26. Comprovação de contaminação com DNA bacteriano em amostras proteicas contendo o domínio TEN.

Géis de agarose 1% corados com brometo de etídeo mostrando o DNA presente nas amostras da proteína TEN recombinante. (A) Presença de DNA nos extratos bacterianos de bactérias induzidas P: pellet, SN: sobrenadante. A Figura mostra que os extratos ainda estavam contaminados após o tratamento com DNase. (B) Otimização do protocolo para retirar o DNA nas amostras proteicas. 1. Extrato bacteriano bruto. 2. Sobrenadante do extrato após a sonicação. 3. Pellet bacteriano antes de tratar com DNase. 4. Pellet bacteriano depois de tratar com DNase. O gel mostra que o tratamento com DNase após a centrifugação é suficiente para retirar o DNA contaminante.

Para reduzir a quantidade de DNA presente na amostra proteica esta foi tratada com DNase antes da sonicação (Secção 3.5). No entanto, os géis de agarose mostraram que o tratamento da amostra bruta não era suficiente para se eliminar o DNA presente (Figura 26A P e SN). Para solucionar este problema, um segundo tratamento com DNase foi feito exclusivamente com a fração insolúvel do extrato, antes de adicionar o tampão com uréia e desnaturar as proteínas. Os resultados no gel de agarose permitem

concluir que este segundo tratamento com DNase reduziu consideravelmente a quantidade de DNA presente no extrato que foi aplicado na coluna (Figura 26B Linha4).

Deste modo, o domínio TEN, sem contaminação com DNA bacteriano, foi purificado. Entretanto não foi possível realizar ensaios de caracterização com o extrato livre de DNA bacteriano, pois a proteína precipitou durante o processo de re-novelamento por diálise (Resultados não mostrados).

Algumas proteínas têm uma maior probabilidade de precipitar se não estão interagindo com seu ligante (SINGH; PANDA, 2005). Neste sentido, é possível que a falta de DNA esteja fazendo com que o domínio TEN não consiga se re-novelar após a purificação. Em estudos anteriores, este problema tinha sido resolvido com a adição de heparina sódica na amostra proteica durante a diálise (LIRA et al., 2009). Porém, a adição de heparina ($100\mu\text{g}\cdot\text{ml}^{-1}$) à amostra de TEN não foi suficiente para impedir sua precipitação. No momento, outros protocolos de re-novelamento estão sendo testados, visando à obtenção do domínio TEN enovelado e na ausência de DNA contaminante, permitindo assim, a realização de ensaios de caracterização bioquímica.

Devido à possibilidade de que o domínio TEN não estivesse sendo enovelado corretamente durante o processo de diálise decidiu-se investir em protocolos que permitissem a solubilização da proteína a partir dos corpúsculos de inclusão sem que esta fosse desnaturada. Foram testados vários protocolos que permitiam que proteínas expressas na forma insolúvel passassem a ser expressas na fração solúvel do extrato bacteriano (SINGH; PANDA, 2005; SINGH et al., 2015).

De todos os protocolos testados o único que foi efetivo na solubilização do domínio TEN foi revisado por SINGH et al. (2015). Este protocolo utiliza a adição de 1% do detergente N-Lauryl Sarcosine ao tampão de lise antes do processo de sonicação

para aumentar a solubilidade das proteínas expressas em corpúsculos de inclusão (Figura 27). Graças à adição do detergente, aproximadamente metade do domínio TEN expresso foi solubilizado, como pode ser visto na Figura 27A (Linhas P e SN) onde uma banda com o peso molecular esperado (30,7 kDa) aparece na fração solúvel do extrato. Verifica-se também que a mesma não está presente quando a lise é feita na ausência de N-Lauryl sarcosine (Figura 27B linha SN).

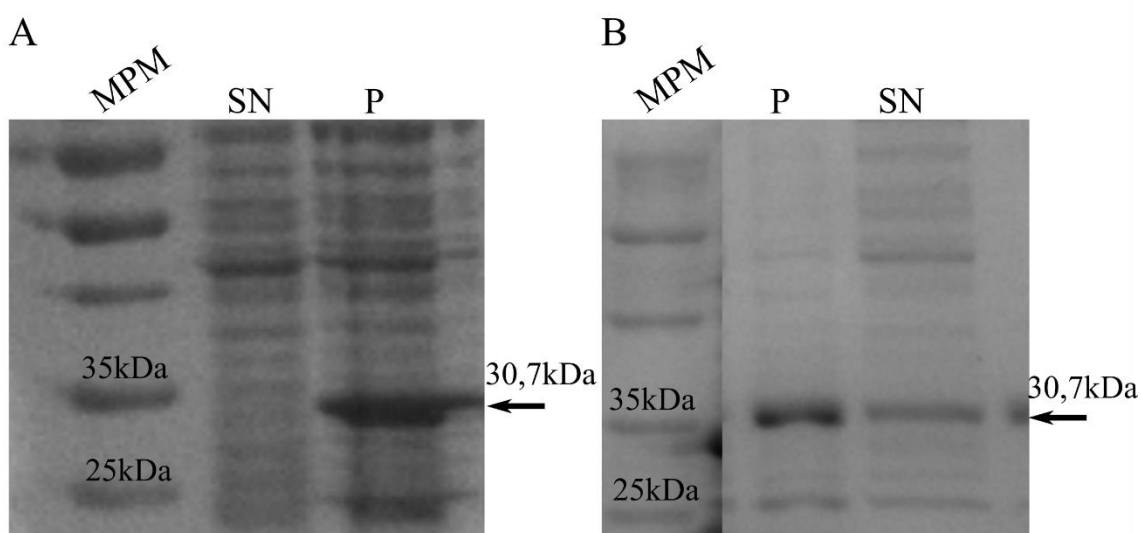


Figura 27. Tratamento do pellet bacteriano com N-Lauryl sarcosine durante a lise

SDS-PAGE 12% corados com azul de comassie mostrando a solubilidade do domínio TEN quando a lise é realizada na ausência (A) e presença (B) do detergente N-Lauryl sarcosine. MPM: Marcador de peso molecular (PageRuler™ Prestained Protein Ladder, Thermo Fisher Scientific). SN: fração solúvel do extrato bacteriano. P: Fração insolúvel do extrato bacteriano. A seta indica o peso esperado do domínio TEN (30,7kDa). Os geis mostram que quando a lise bacteriana é feita na ausência de N-Lauryl sarcosine o domínio TEN recombinante é expresso exclusivamente na fração insolúvel. Quando a lise é realizada na presença de N-Lauryl sarcosine, aparece uma banda do peso esperado na fração solúvel e a intensidade da banda presente na fração insolúvel diminui, indicando que uma parte do domínio TEN recombinante está sendo expressa na fração solúvel.

Tentou-se purificar o domínio TEN a partir da fração solúvel do extrato proteico, mediante cromatografia de troca iônica, já que esta aproximação tinha funcionado na purificação do domínio TEN a partir do extrato insolúvel (Figura 21). No entanto, esta tentativa não foi bem-sucedida na purificação do domínio TEN a partir da fração solúvel, como pode ser visto na Figura 28, já que a proteína não ligou na resina e foi

eluída no *flow-through*. Isto é evidenciando pelo fato que não há proteínas do tamanho esperado eluídas durante o gradiente linear e que a banda de massa esperada presente no *input* está presente com a mesma intensidade no *flow-through*.

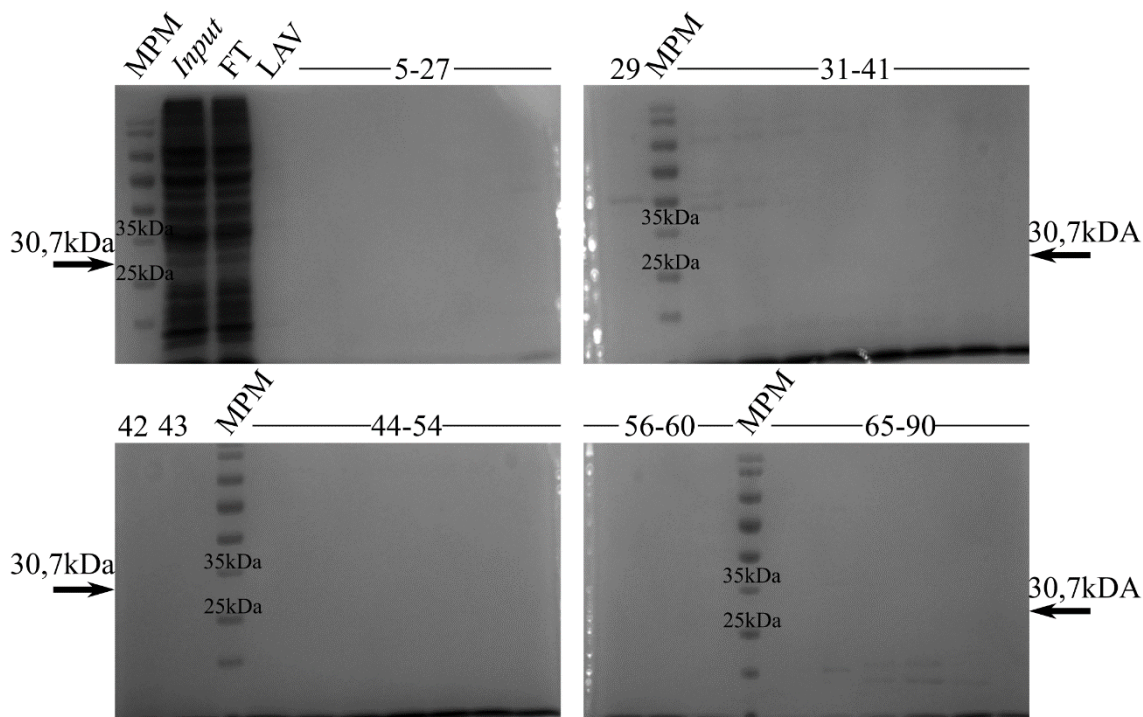


Figura 28. Tentativa de purificação do domínio TEN expresso a partir da fração solúvel do extrato proteico mediante cromatografia de troca iônica

SDS-PAGE 12% corados com azul de comassie mostrando a purificação do domínio TEN a partir da fração solúvel do extrato proteico mediante cromatografia de troca iônica. MPM: Marcador de peso molecular (PageRuler™ Prestained Protein Ladder, Thermo Fisher Scientific). *Input*: amostra injetada na coluna. FT: *Flow-through*, proteínas que não ligam na coluna. LAV: proteínas eluídas na lavagem da coluna, 5-90: número da fração eluída no gradiente linear de tampão B. As setas indicam o peso esperado do domínio TEN heterólogo, mas pode se ver que nenhuma banda desse tamanho foi eluída da coluna. Adicionalmente, pode se ver que a banda do peso esperado no *input* está presente com a mesma intensidade no *flow-through*, o que indica que a proteína não ligou na coluna.

Dado o insucesso nas tentativas de purificação por troca iônica tentou-se purificar o domínio TEN expresso na fração solúvel por cromatografia de afinidade, já que ele é expresso em fusão com uma calda de histidina no extremo amino-terminal. No entanto, como mostrado na Figura 29 obtiveram-se resultados similares aos obtidos por

cromatografia de troca iônica. O domínio TEN não se ligou na coluna já que a banda do tamanho esperado está presente unicamente no *input* e no *flow-through*.

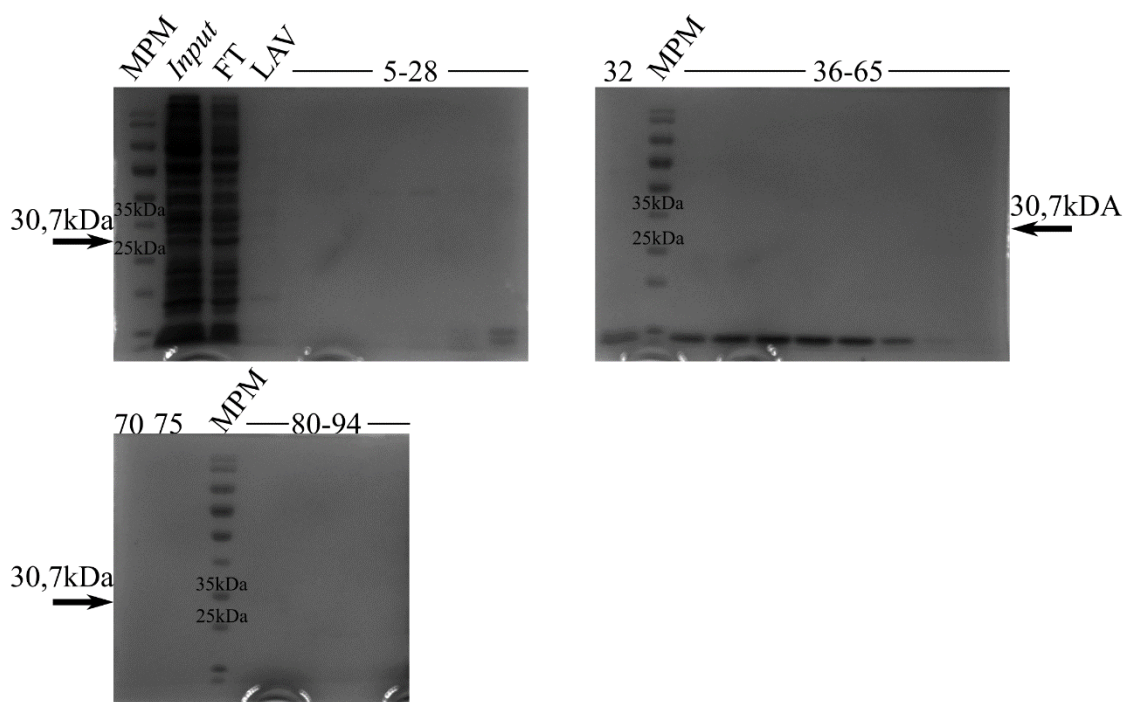


Figura 29. Tentativa de purificação do domínio TEN expresso a partir da fração solúvel do extrato proteico mediante cromatografia de afinidade

SDS-PAGE 12% corados com azul de comassie mostrando a purificação do domínio TEN a partir da fração solúvel do extrato proteico mediante cromatografia de afinidade. MPM: Marcador de peso molecular (PageRuler™ Prestained Protein Ladder, Thermo Fisher Scientific). *Input*: amostra injetada na coluna. FT: *Flow-through*, proteínas que não ligam na coluna. LAV: proteínas eluídas na lavagem da coluna, 5-94: número da fração eluída no gradiente linear de tampão B. As setas indicam o peso esperado do domínio TEN heterólogo, mas pode se ver que nenhuma banda desse tamanho foi eluída da coluna. Adicionalmente, pode se ver que a banda do peso esperado no *input* está presente com a mesma intensidade no *flow-through*, o que indica que a proteína não ligou na coluna.

Para garantir que o domínio TEN estava sendo expresso na fração solúvel antes de se tentar novas alternativas para sua purificação, foi realizado um *western blot* com os extratos proteicos utilizados na cromatografia por troca iônica. Este *western blot* foi revelado com soro policlonal específico para a região amino-terminal do TERT de *L. amazonensis*. (Figura 30). O *western blot* mostrou que o domínio TEN está sendo expresso na fração solúvel, já que o anticorpo reconheceu uma única banda de peso molecular esperado tanto no *input* como no *flow-through* (Figura 30, *input* e FT), a qual

não está presente no extrato proteico não induzido (Figura 30 NI). Este resultado confirma que a banda presente na fração solúvel é o domínio TEN heterólogo e não um produto inespecífico.

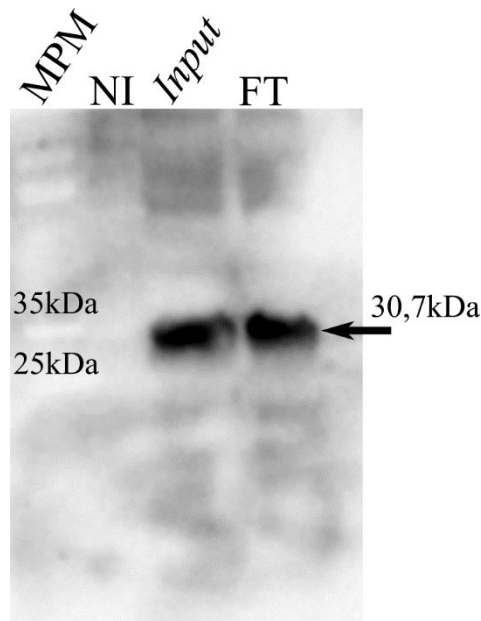


Figura 30. Western blot comprovando que o domínio TEN foi expresso na fração solúvel do extrato proteico

Diferentes extratos proteicos foram fracionados em gel SDS-PAGE 12%, transferido para membrana PVDF e relevado usando soro policlonal específico para a região amino-terminal do TERT de *L. amazonensis*. MPM: marcador de peso molecular (PageRuler™ Prestained Protein Ladder, Thermo Fisher Scientific), Input: amostra injetada na coluna. Flow-through: proteínas que não ligam na coluna. A banda (30,7kDa) reconhecida pelo soro anti-TERT, está apontada por uma seta e aparece exclusivamente nos extratos onde a expressão da proteína heteróloga foi induzida, confirmando a presença do domínio TEN na fração solúvel do extrato bacteriano.

Após a confirmação da expressão do domínio TEN na fração solúvel, estão sendo explorados novos protocolos de purificação que permitam purificar o domínio TEN em larga escala.

Dada a baixa caracterização estrutural disponível para o domínio TEN dos diferentes TERT, incluído o de *Leishmania*, qualquer dado experimental obtido a partir da purificação deste domínio será de grande importância para o aprofundamento do

entendimento deste domínio no mecanismo de ação da telomerase de diferentes organismos.

4.4 Caracterização estrutural do domínio TRB do TERT de *Leishmania* sp.

De modo geral, o domínio TRB apresenta alta conservação, tanto de estrutura primária quanto terciária entre as diferentes espécies de eucariotos. Este domínio apresenta três motivos conservados que são de importância na interação com o TER e na montagem do complexo RNP das telomerasas.

Para se avaliar a conservação do domínio TRB em *Leishmania* sp. foram realizados alinhamentos múltiplos entre as sequências do domínio TRB de vários grupos de eucariotos. Como mostrado na Figura 31, os três motivos que caracterizam o domínio TRB, demarcados no alinhamento (CP, QPF e T), são conservados em diferentes grupos de organismos eucariotos mais derivados como os mamíferos e os ciliados. O domínio TRB de *L. major* apresenta algumas substituições nas sequências dos motivos QFP e T, ainda que haja uma boa conservação geral entre eles. Porém, ao se analisar a sequência que caracteriza o motivo CP, nota-se que a conservação é muito baixa ou inexistente, o que dificulta o alinhamento entre as sequências. Nota-se também que, dependendo tanto da delimitação do domínio, quanto dos organismos utilizados para gerar o alinhamento, a posição deste motivo varia (resultados não mostrados), sugerindo fortemente que este motivo não é conservado no domínio TRB de *Leishmania* sp.

Vale a pena se destacar que as substituições apresentadas no motivo T da TERT de *L. major* e que parecem ser próprias do gênero, correspondem a aminoácidos conservados que nas proteínas homólogas dos demais eucariotos (ciliados, leveduras de brotamento e vertebrados) são de importância para a atividade de telomerase (ROUDA;

SKORDALAKES, 2007). Estas substituições serão discutidas com maiores detalhes posteriormente.

Entretanto, estudos de mutação sítio-dirigida em outros organismos eucariotos não encontraram nenhum resíduo dentro do motivo QFP que fosse indispensável para a atividade telomerase (BRYAN et al., 2000; LAI et al., 2001; ROUDA; SKORDALAKES, 2007). Deste modo, não é possível discutir as implicações da conservação deste domínio em *L. major*.

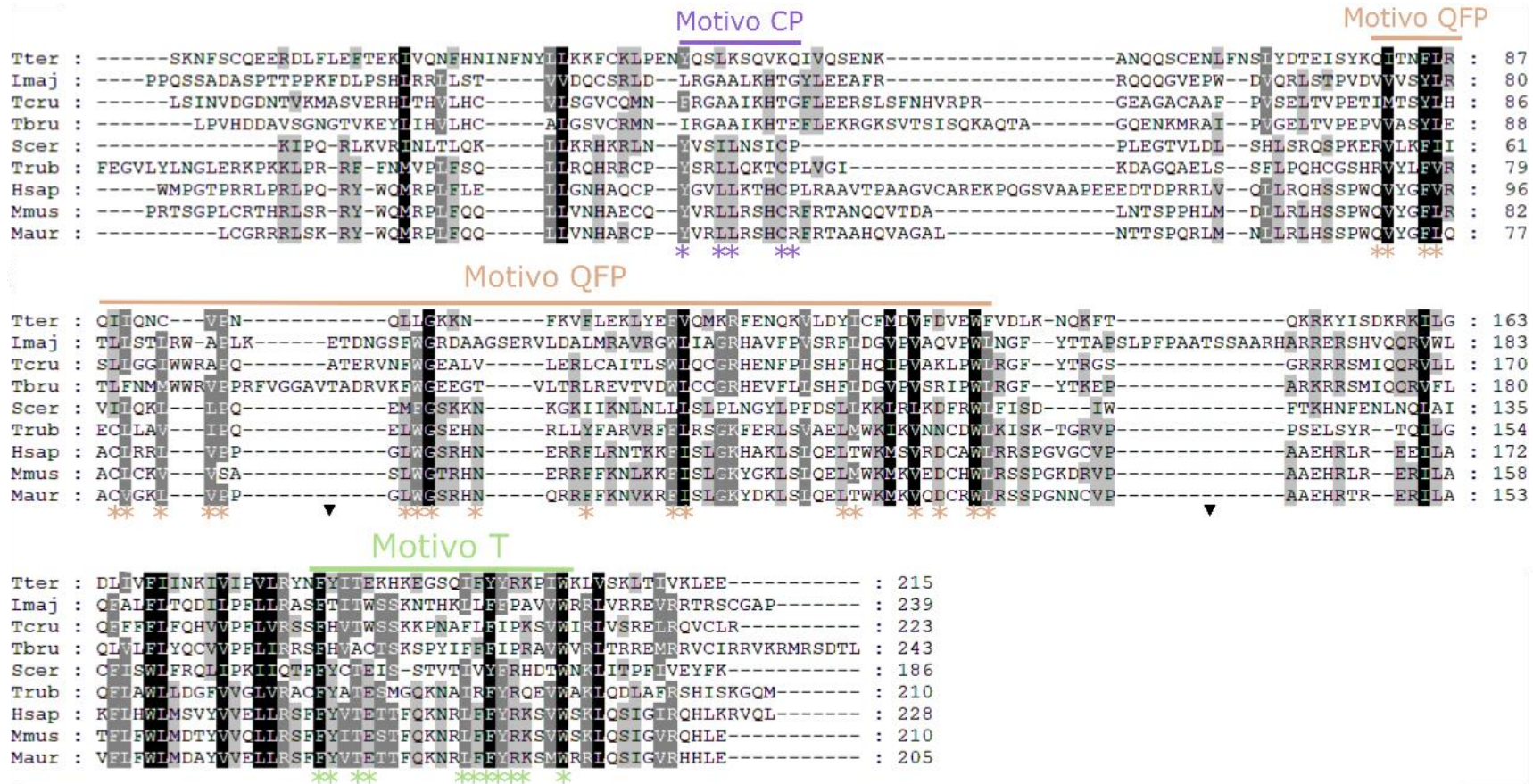


Figura 31. Alinhamento de seqüências do domínio TRB e demarcação da posição dos motivos

O alinhamento mostra o domínio TRB e seus três motivos conservados, CP, QFP e T. Resíduos de importância funcional nos motivos são mostrados com asteriscos e inserções de seqüências exclusivas de um único organismo são destacadas por triângulos. *L. major* mostra substituições em resíduos presentes nos motivos T e QFP, e um grau de conservação muito baixo ou inexistente na seqüência do motivo CP. O domínio TRB foi isolado a partir da seqüência dos TERT dos seguintes organismos: *Lmaj*: *L. major* (AAO67514.1); *Tter*: *Tetrahymena thermophila* (XP_001010651.3); *Tcru*: *Trypanosoma cruzi* (XP_819198.1); *Tbru*: *Trypanosoma brucei* (XP_829083.1) *Scer*: *Saccharomyces cerevisiae* (NP_013422.2); *Trub*: *Takifugu rubripes* (XP_003969013.1); *Hsap*: *Homo sapiens* (NP_937983.2); *Mmus*: *Mus musculus* (NP_033380.1); *Maur*: *Mesocricetus auratus* (NP_001268556.1)). Resíduos destacados em preto denotam conservação em todos os organismos e em cinza conservação na maioria dos organismos.

A melhor predição de estrutura terciária obtida para do domínio TRB apresenta 232 aminoácidos no total (resíduos 424 a 656) e foi modelado a partir das estruturas cristalográficas do domínio TRB de *T. rubripes* (4LMO) (HARKISHEIMER et al., 2013) e do TERT de *T. castaneum* (3DU5) (GILLIS et al., 2008) (Figura 32A), este modelo foi refinado por dinâmica molecular (MD). O modelo obtido após 150 ns de MD mostra uma conservação estrutural parcial do domínio TRB quando alinhado estruturalmente (sobreposição de Ca) com o motivo de *T. rubripes* (Figura 32B).

Após 150 ns de MD, a estrutura apresenta estabilidade energética, como mostrado no gráfico r.m.s.d. (*root mean square deviation*) (Figura 32C). O gráfico de r.m.s.f (*root mean square fluctuation*) mostra que a região correspondente aos primeiros 75 aminoácidos da região amino-terminal apresenta flutuações entre 0,5 e 0,6 nm ao longo da dinâmica molecular, comparado com uma flutuação de aproximadamente 0,2 nm no restante da molécula (Figura 32D). Esses dados mostram que essa região é a mais flexível deste domínio.

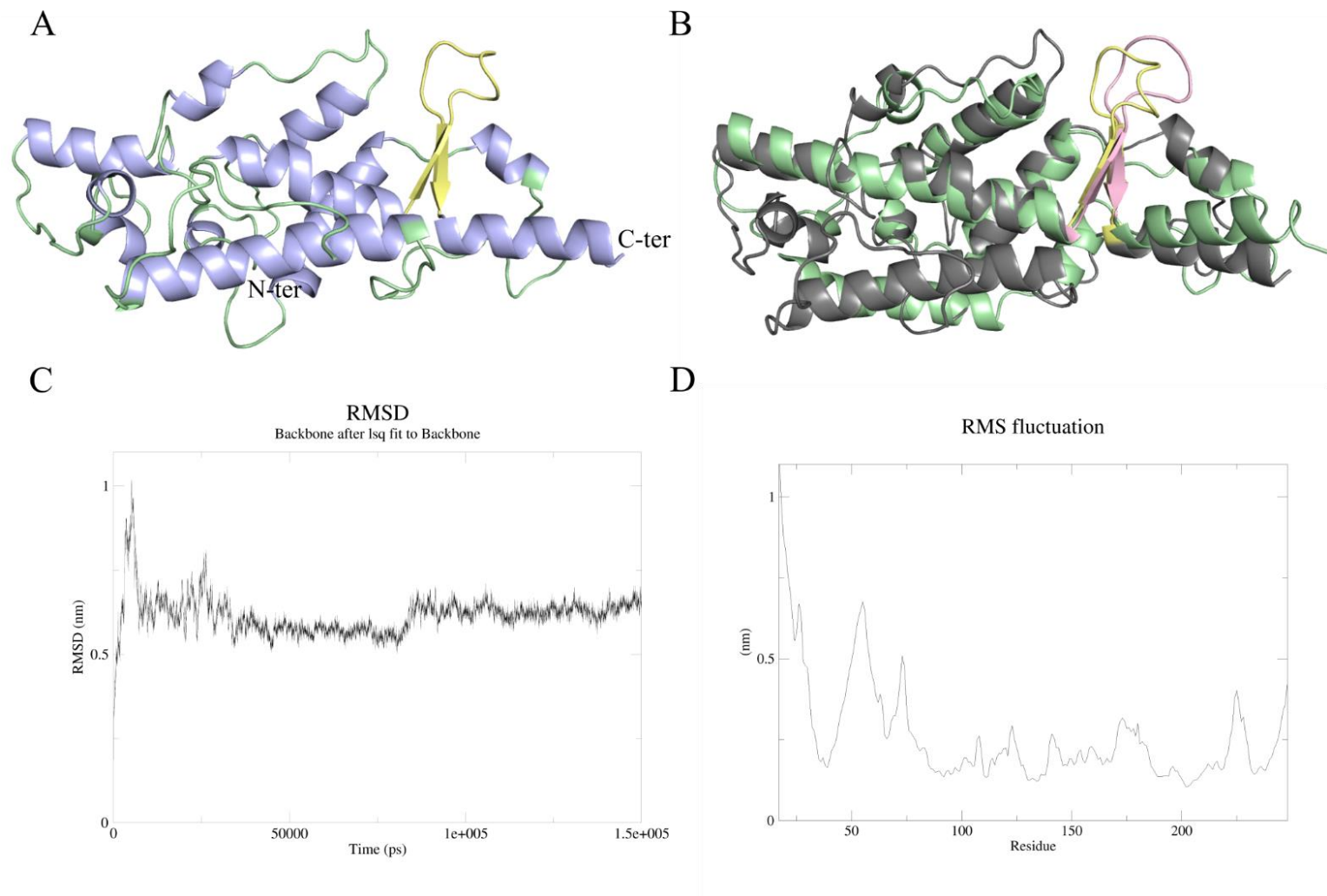


Figura 32. Predição da estrutura terciária do domínio TRB de *L. major*

(A) Representação do modelo estrutural *in silico* do domínio TRB da LmTERT (resíduos 424 a 656) obtido por técnicas de *threading* a partir de modelos cristalográficos previamente reportados 4LMO (domínio TRB, TERT *T. rubripes*) e 3DU5 (TERT *T. castaneum*) e 150 ns de simulação de dinâmica molecular (MD). As hélices- α estão destacadas em azul, os loops em verde e o β -hairpin característicos deste domínio em amarelo. (B) sobreposição de Ca da estrutura cristalográfica do domínio TRB de *T. rubripes* (4LMO) (verde) e o modelo *in silico* do domínio TRB de *L. major* (Cinza). A superposição das duas estruturas apresenta conservação geral da estrutura terciária, especialmente na região próxima ao β -hairpin (em rosa em, *T. rubripes*, em amarelo em, *L. major*). (C) *Root mean square deviation* (r.m.s.d.) da simulação de MD executada por GROMACS (Groningen Machine for Chemical Simulation) v.4.5.3 using GROMOS 96 53a6 force field. (D) *Root mean square fluctuation* (r.m.s.f.) da simulação de MD executada por GROMACS (Groningen Machine for Chemical Simulation) v.4.5.3 utilizando GROMOS 96 53a6 force field. As representações das proteínas em *cartoon* foram geradas com o programa PyMol.

Dado que a estrutura cristalográfica usada como molde principal para se gerar o modelo tridimensional do domínio TRB de *L. major* foi a estrutura cristalográfica do domínio TRB de *T. rubripes*, decidiu-se realizar estudos estruturais comparativos entre elas. A sobreposição de C α do modelo predito do domínio TRB de *L. major* com a estrutura tridimensional de *T. rubripes* evidencia uma conservação estrutural dos seus elementos de estrutura secundária. A sobreposição de C α entre ambas as estruturas apresenta um valor de 3,3 Å de desvio quadrático médio (*root mean square deviation*; r.m.s.d). Entretanto, este valor de r.m.s.d. atinge o baixo valor de 1,5 Å na região do β -*hairpin*. A sobreposição de C α do modelo predito do domínio TRB de *L. major* e de domínios TRB cristalográficos de *T. thermophila* e *T. castaneum* apresentam valores de r.m.s.d. semelhantes tanto para toda a estrutura quanto para a região do β -*hairpin*.

4.4.1 Análise estrutural do motivo T em *L. major*

O próximo passo foi verificar em que local da estrutura terciária estavam localizados os resíduos do motivo T que foram previamente reportados como de importância na atividade telomerase, para se averiguar quais substituições presentes em *Leishmania* poderiam ter implicações sobre a catálise enzimática (Figura 33). Os aminoácidos que foram examinados são apresentados na Tabela 4. Estes aminoácidos causam uma perda severa de atividade telomerase (80-100%) quando mutados por alaninas em outros organismos, como *T. thermophila* e *S. cerevisiae* (BRYAN et al., 2000; LAI et al., 2001). Nestes casos, a perda da atividade enzimática provavelmente é provocada pela falta de interação entre TERT e TER, principalmente, devido ao não reconhecimento do elemento TBE do TER (LAI et al., 2001; ROUDA; SKORDALAKES, 2007). No caso de *Leishmania* sp. essas substituições no motivo T talvez expliquem a baixa atividade da enzima detectada *in vitro* nos extratos do parasito, comparados com extratos de *T. brucei* (CANO et al., 1999; GIARDINI et al., 2011).

Tabela 4. Resíduos conservados no motivo T de diferentes organismos, o asterisco denota que o resíduo não é conservado em *L. major* quando comparado com os eucariotos derivados

Espécie	Resíduos						
L. major	F626	T627*	T639	W630*	P642*	A643*	W646
T. rubripes	F511	Y512	T514	E515	R527	Q528	W531
T. thermophila	F476	Y477	T479	E480	R492	K493	W496
S. cerevisiae	F381	Y382	T384	E385	R397	Q398	W401
T. cruzi	F473	H474*	T476	W477*	P489*	K490	W493
T. brucei	F489	H490*	A492*	C493*	P505*	R506	W509

A Figura 33 apresenta em detalhe os aminoácidos representados na Tabela 4, em *L. major*, *T. rubripes* e *T. thermophila*. Esta representação estrutural permite apreciar que as substituições naturais de *L. major* se encontram no característico β -hairpin e que conseqüentemente podem alterar a organização geral do motivo T.

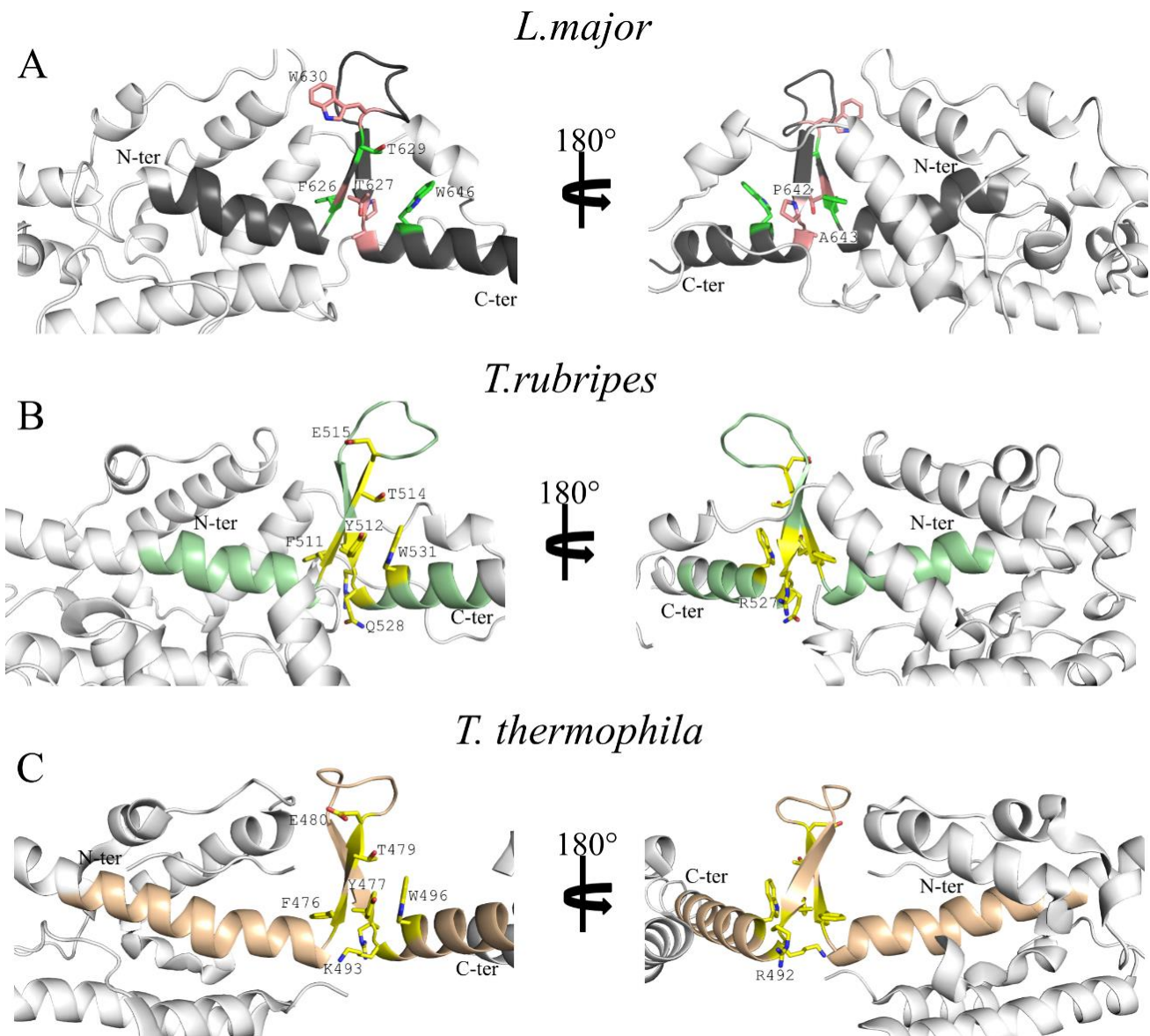


Figura 33. Modificações no motivo T causadas pelas substituições presentes em *L. major*

Ampliação do motivo T em *L. major* (A), *T. rubripes* (B) e *T. thermophila* (C). Os resíduos importantes estão mostrados em sticks amarelos em *T. rubripes* e *T. thermophila*. Em *L. major* os resíduos em verde estão conservados em relação ao motivo consenso, enquanto os resíduos em rosa não são conservados. Pode se ver que as substituições presentes em *L. major* encontram-se no β -hairpin. As representações das proteínas foram geradas com o programa PyMol.

Uma das possíveis alterações estruturais que existe no motivo T em *L. major* se dá na formação de uma pinça hidrofóbica que é mediada por aminoácidos aromáticos. Em *T. rubripes*, esta pinça é estabilizada pelos resíduos Y512 e W530 que têm os anéis aromáticos das suas cadeias laterais enfrentadas formando uma pinça hidrofóbica (*hydrophobic pincer*) que estabiliza a formação do β -hairpin (Figura 34B) (JANSSON et al., 2015). Em *T. thermophila*, a pinça está formada pelos aminoácidos W496 e Y477 que estão numa conformação similar à observada para *T. rubripes* (Figura 34C)

(ROUDA; SKORDALAKES, 2007). Como no caso de *L. major* existe uma treonina (T628) no lugar da tirosina (Y), não é possível se dar a estabilização da pinça hidrofóbica por parte destes dois aminoácidos. No entanto, na estrutura de *L. major* (Figura 34A) o triptofano que deveria formar a pinça (W646) está numa orientação diferente e se encontra no mesmo plano que um outro resíduo aromático da folha- β mais carboxi-terminal (F642). É possível que a estabilização da pinça hidrofóbica em *L. major* seja dada entre estes dois resíduos, sugerindo que o arranjo desta pinça hidrofóbica e, conseqüentemente, o modo de interação com o TER, em *L. major* é diferente do que nos organismos caracterizados previamente.

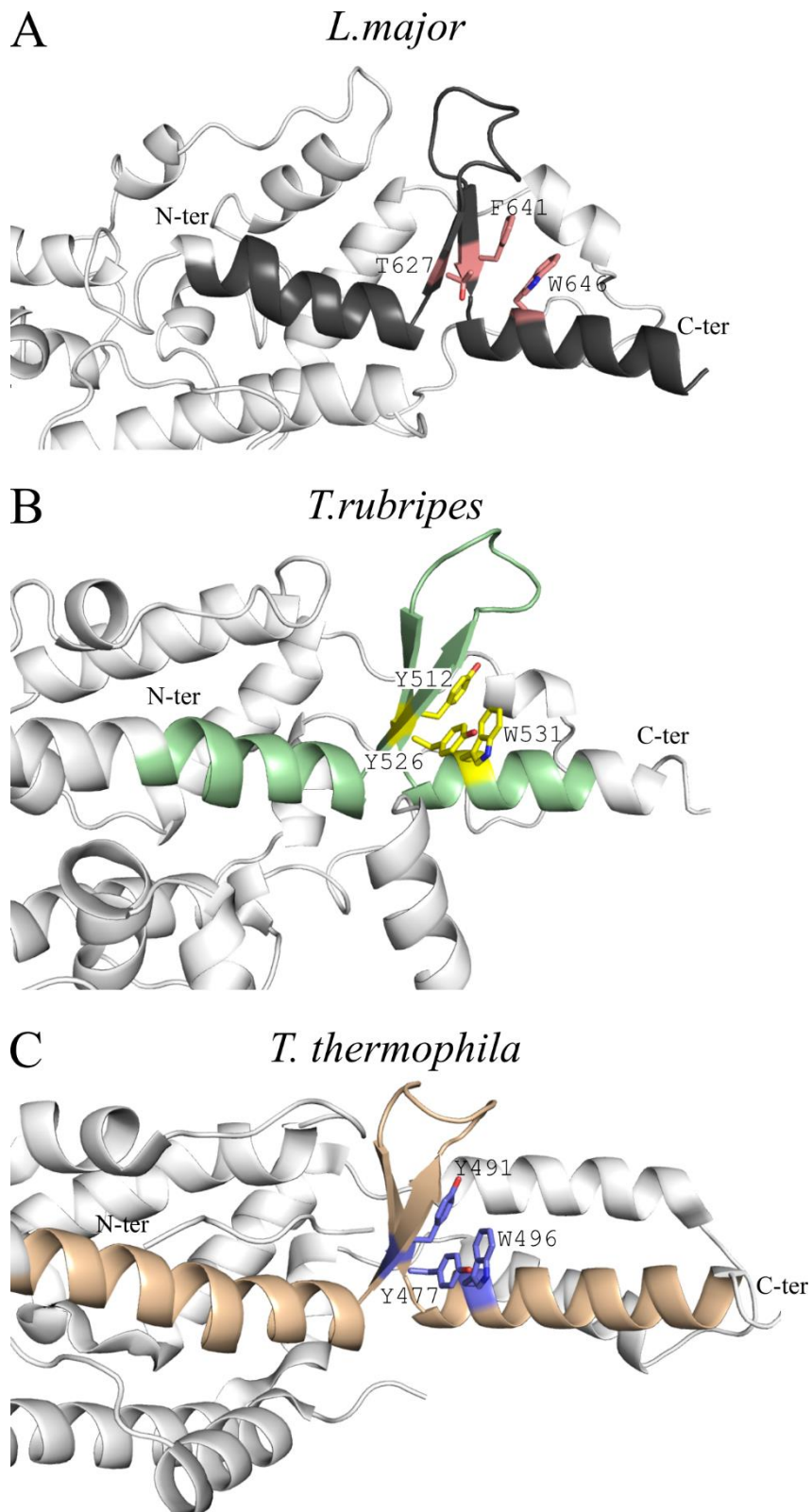


Figura 34. Detalhe da pinça hidrofóbica no motivo T

Ampliação do motivo T em *L. major* (A), *T. rubripes* (B) e *T. thermophila* (C). Os resíduos envolvidos em formar a pinça hidrofóbica estão destacados em sticks. Pode se ver que em *T. rubripes* e *T. thermophila* o triptofano presente na hélice- α mais carboxi-terminal está no mesmo plano que uma tirosina presente na folha- β mais amino-terminal, enquanto que em *L. major* a orientação do triptofano está alterada e este se encontra no mesmo plano que uma fenilalanina presente na folha- β mais carboxi-terminal. As representações da proteína em *cartoon* foram geradas com o programa PyMol.

4.4.2 Análise estrutural do motivo CP em *L. major*

Dentro dos resíduos presentes no motivo T que não são conservados em *L. major* (Tabela 4), cabe destacar os resíduos P643 e A644 (Figura 33A). Em *T. rubripes* estes resíduos correspondem a R527 e Q528, dois resíduos com cadeias laterais longas que interagem com outros dois resíduos conservados no motivo CP (L381 e C384) (Figura 35B) (HARKISHEIMER et al., 2013). Estas interações também estão presentes em *T. thermophila* sendo mediadas pelos resíduos R492 e K493 no motivo T e C331 e L333 no motivo CP (Figura 35C) (ROUDA; SKORDALAKES, 2007). A interação entre os motivos T e CP têm sido caracterizada como de importância estrutural na conformação da superfície característica do domínio TRB, que forma duas cavidades chamadas de T e CP *pockets* que foram propostas como de importância funcional para a montagem do complexo e a atividade de telomerase (ROUDA; SKORDALAKES, 2007; HARKISHEIMER et al., 2013; JANSSON et al., 2015).

No entanto, *L. major* não apresenta conservação em nenhum destes aminoácidos, e no seu lugar apresenta resíduos com cadeias laterais muito pequenas o que dificultaria sua interação com resíduos de outros motivos (Figura 35A). Adicionalmente, o motivo CP não apresenta conservação de estrutura primária em *L. major* (Figura 31), o que indica que os aminoácidos que interagiriam com o motivo T também não são conservados e que possivelmente a interação entre os dois motivos não exista, ou seja feita por outros resíduos em *L. major*.

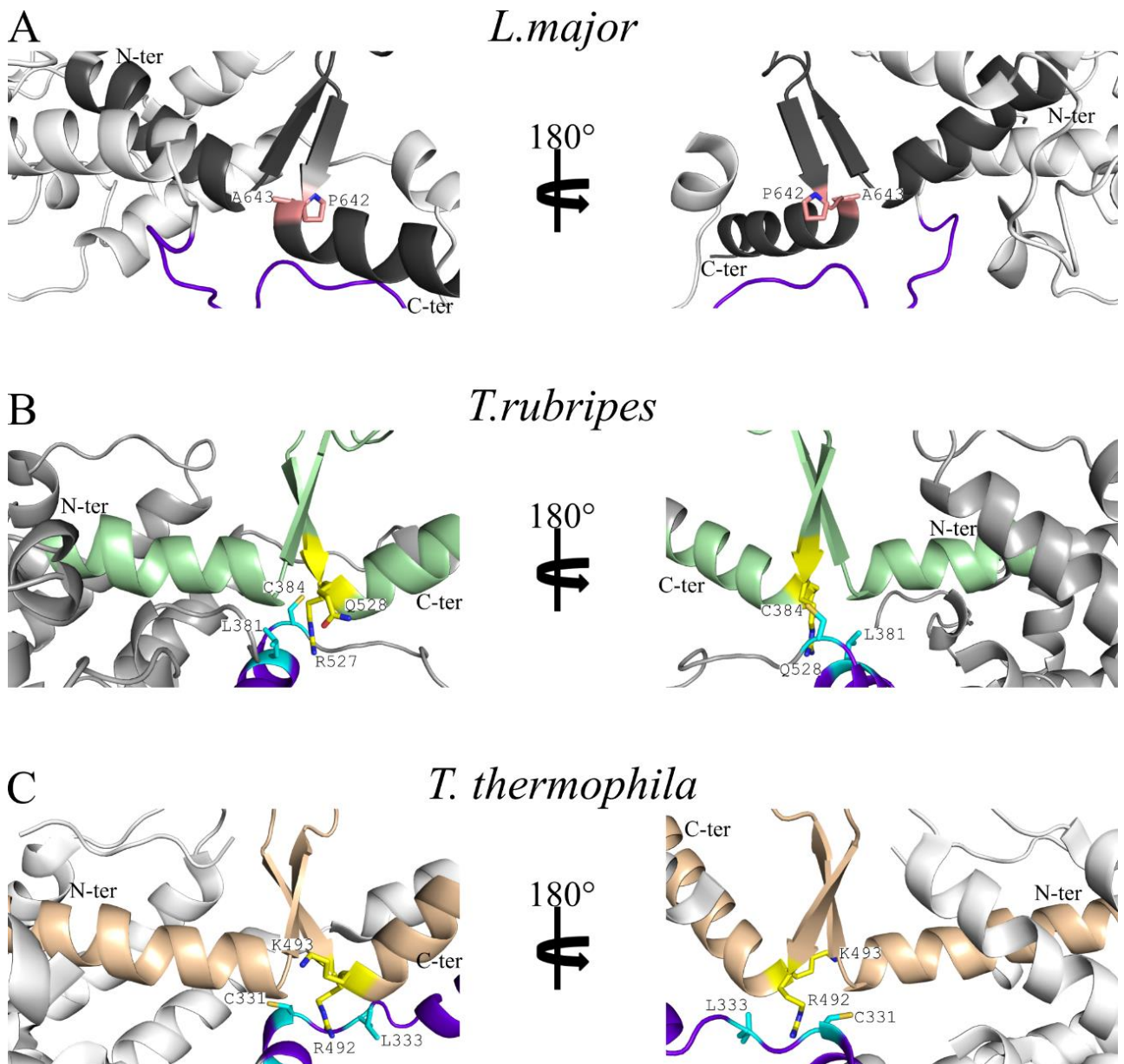


Figura 35. Detalhes da interação entre o motivo CP e o motivo T

Ampliação dos motivos T e CP em *L. major* (A), *T. rubripes* (B) e *T. thermophila* (C). Os resíduos importantes estão mostrados em *sticks*. Em *L. major* o motivo T está demarcado em cinza e o motivo CP está demarcado em roxo. Na estrutura de *T. rubripes* o motivo T está demarcado em verde e o motivo CP está demarcado em roxo. Na estrutura de *T. thermophila* o motivo T está demarcado em marrom e o motivo CP está demarcado em roxo. Na estrutura de *T. rubripes* e *T. thermophila* há resíduos que medeiam a interação entre os motivos T e CP, esta interação não é conservada em *L. major*. As representações das proteínas foram geradas com o programa PyMol.

Em termos de estrutura terciária, o domínio CP é formado por duas pequenas hélices- α em *T. thermophila* e vertebrados (ROUDA; SKORDALAKES, 2007; HARKISHEIMER et al., 2013). Este motivo é parcialmente conservado em *L. major* já que a hélice mais amino-terminal está conservada, mas no lugar da segunda hélice há

uma região desenovelada (Figura 36), diferente de uma região rica em hélices- α , como em outros organismos. O elemento desenovelado abrange os resíduos 56 a 76, esta região coincide com a região de maior flutuação do r.m.s.f. da dinâmica molecular (Figura 32D), sendo que a estrutura primária desta região (RGA AKHTG) está enriquecida em aminoácidos flexíveis, contendo duas alaninas e duas glicinas. Três destes resíduos encontram-se formando um *cluster* GAA no começo da região desordenada. Estes dados indicam que esta região é altamente flexível. Estes resultados, aliados à reorientação da pinça hidrofóbica no motivo T, sugerem uma modificação no mecanismo de interação com o componente TER em *Leishmania*.

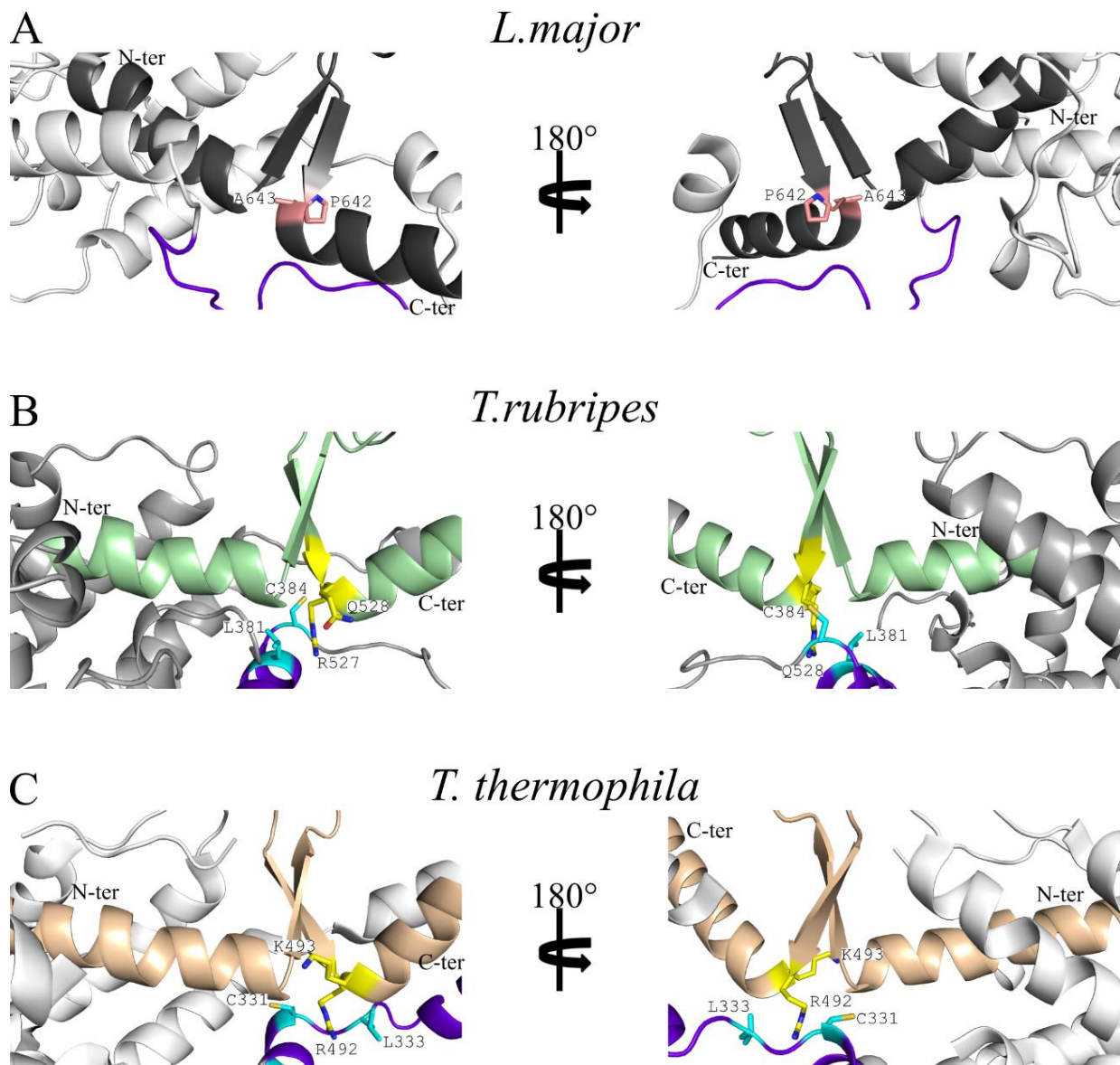


Figura 36. Detalhe dos motivos T e CP

Representações em *cartoon* do domínio TRB de *L. major* (A), *T. rubripes* (B) e *T. thermophila* (C). Em *T. rubripes* e *T. thermophila* o motivo CP está formado pela junção de duas hélices- α (azul) a segunda das quais está muito próxima do motivo T (amarelo). No caso de *L. major* a segunda hélice- α é substituída por uma região desenovelada (Laranja), cuja proximidade com o motivo T (rosa) é difícil de se inferir. As representações das proteínas foram geradas com o programa PyMol.

4.4.3 Análise estrutural do mecanismo de interação entre o domínio TRB e o componente TER em *L. major*

Com o intuito de se estudar mais detalhadamente a possível interação do domínio TRB de *L. major* com o TER, foi feito um estudo de *docking* molecular entre a estrutura predita *in silico* para o domínio TRB de *L. major* e a estrutura tridimensional do elemento TBE de *T. thermophila*. A informação sobre a possível interação entre o

domínio TRB de *L. major* e o elemento TBE de *T. thermophila*, cuja estrutura está disponível (PDB ID: 5C9H), permite fazer inferências entre a possível interação destes dois componentes em *L. major* já que o elemento TBE de *L. major* predito por VASCONCELOS et al., (2014) é similar ao elemento TBE de *T. thermophila* descrito por LICHT; COLLINS, (1999), tanto em termos de estrutura secundária como em termos de sequência de nucleotídeos (VASCONCELOS et al., 2014).

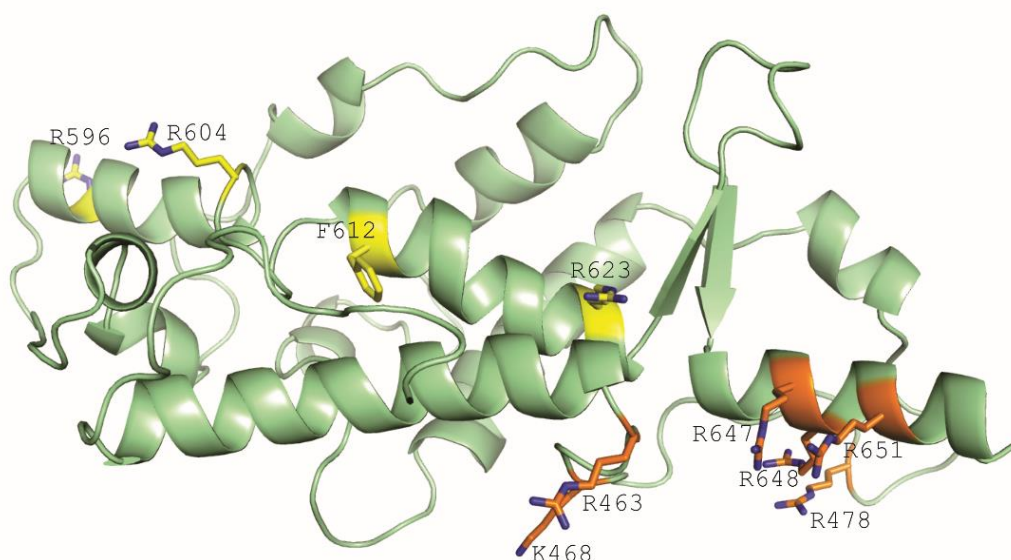


Figura 37. Resíduos do modelo tridimensional do TRBD de *L. major* escolhidos como alvo para *docking* molecular

Representação em *cartoon* de predição de estrutura terciária do domínio TRB de *L. major*. Os aminoácidos escolhidos como alvo para se fazer *docking* molecular de um único resíduo são representados em *sticks*. Os resíduos destacados em laranja são os que apresentaram melhores valores energéticos e de HADDOCK score nos *dockings* moleculares usando um único resíduo como alvo e foram usados conjuntamente como alvo para se fazer o *docking* molecular definitivo (*dockingI*). A representação da proteína foi gerada com o programa PyMol.

Para se fazer o *docking* molecular, primeiramente foram selecionados resíduos próximos localizados na região de interação entre o domínio TRB e o TBE em *T. thermophila*, assim como resíduos localizados em regiões próximas. Todos os resíduos selecionados compartilham propriedades químicas (aminoácidos carregados positivamente e aminoácidos aromáticos) com os resíduos que medeiam a interação

entre o domínio TRB e o elemento TBE em *T. thermophila* (Figura 37). Foram feitos 10 *dockings* moleculares iniciais usando como alvo cada um dos resíduos selecionados anteriormente e mais um raio de 5,0 Å e a totalidade do elemento TBE de *T. thermophila*.

Tabela 5. Valores obtidos para cada um dos alvos após o *Docking* molecular

Alvo	Tamanho de <i>cluster</i>	HADDOCK score	RMSD	Energia de desolvatação	Energia de Van de Waals	Energia eletrostática	Energia de violação de restrições	Área de superfície enterrada	Z-score
R 463	79	-86.2 +/- 4.1	1.2 +/- 1.1	29.1 +/- 2.2	-38.8 +/- 5.5	-383.8 +/- 20.7	2.3 +/- 0.47	1101.3 +/- 79.0	-2.1
R 478	47	-72.8 +/- 1.4	9.5 +/- 0.0	26.7 +/- 8.3	-29.3 +/- 2.4	-352.6 +/- 54.5	3.8 +/- 0.44	770.3 +/- 73.3	-1.2
K 468	77	-88.5 +/- 3.1	5.2 +/- 0.2	31.7 +/- 5.4	-39.3 +/- 5.3	-405.9 +/- 21.5	2.5 +/- 1.14	1329.0 +/- 185.6	-1.2
R 596	156	-97.6 +/- 3.0	1.0 +/- 0.6	15.2 +/- 5.0	-40.8 +/- 4.7	-361.0 +/- 9.7	2.0 +/- 0.47	922.9 +/- 44.9	-1.6
R 604	8	-88.2 +/- 18.6	0.9 +/- 0.5	29.0 +/- 4.3	-41.2 +/- 11.5	-381.2 +/- 45.9	2.3 +/- 0.97	1091.6 +/- 221.1	-1.7
F 612	105	-58.5 +/- 1.9	12.7 +/- 0.2	3.8 +/- 3.0	-39.3 +/- 2.7	-136.7 +/- 9.1	44.6 +/- 0.51	908.4 +/- 52.4	-1.4
R 623	150	-91.6 +/- 7.1	0.7 +/- 0.4	17.1 +/- 5.9	-43.3 +/- 6.1	-328.2 +/- 38.0	2.0 +/- 0.28	1348.3 +/- 150.4	-1.0
R647	17	-97.8 +/- 4.1	8.3 +/- 0.6	32.7 +/- 8.7	-38.1 +/- 7.0	-463.8 +/- 40.0	3.0 +/- 0.75	1258.4 +/- 246.2	-1.4
R 648	68	-116.7 +/- 13.8	0.9 +/- 0.6	35.0 +/- 9.7	-47.9 +/- 8.1	-520.4 +/- 77.3	2.2 +/- 0.22	1351.7 +/- 201.5	-2.0
R 651	52	-81.5 +/- 2.4	7.2 +/- 0.3	28.8 +/- 4.4	-31.9 +/- 5.4	-394.2 +/- 19.0	3.2 +/- 0.15	961.9 +/- 220.1	-1.3
R 463, K 468, R 478, R647, R 648, R 651	79	-143.7 +/- 3.4	1.1 +/- 0.7	42.6 +/- 4.5	-82.2 +/- 3.1	-535.0 +/- 32.9	28.8 +/- 14.67	2241.7 +/- 106.6	-1.8

Os valores energéticos e de HADDOCK *score* obtidos para o melhor grupo (*cluster*) de soluções para cada alvo são apresentados na Tabela 5. A partir dos resultados obtidos para cada um dos resíduos isolados, foram selecionados os resíduos que apresentaram os melhores valores energéticos e de HADDOCK *score* para se fazer um *docking* molecular da região mais provável de interação entre o domínio TRB de *L. major* e o elemento TBE de *T. thermophila*. Os aminoácidos selecionados foram: R463, K468, R478, R647, R648 e R651 (Figura 37) (Tabela 5). Posteriormente, o *docking* correspondente a estes aminoácidos será referido como *docking 1*.

Os valores energéticos e de HADDOCK *score* obtidos para o melhor grupo (*cluster*) de soluções do *docking 1* são apresentados na Tabela 5. Foram feitos outros ensaios de *docking* molecular, usando-se diferentes combinações de resíduos, mas nenhum deles apresentou resultados energéticos e de HADDOCK *score* tão bons quanto os obtidos para o *docking 1* (resultados não mostrados). A melhor solução do *docking 1* é apresentada na Figura 38A, mostrando que o elemento TBE interage com a região próxima à região de interação de *T. thermophila* (Figura 38B) embora a orientação deste elemento seja diferente.

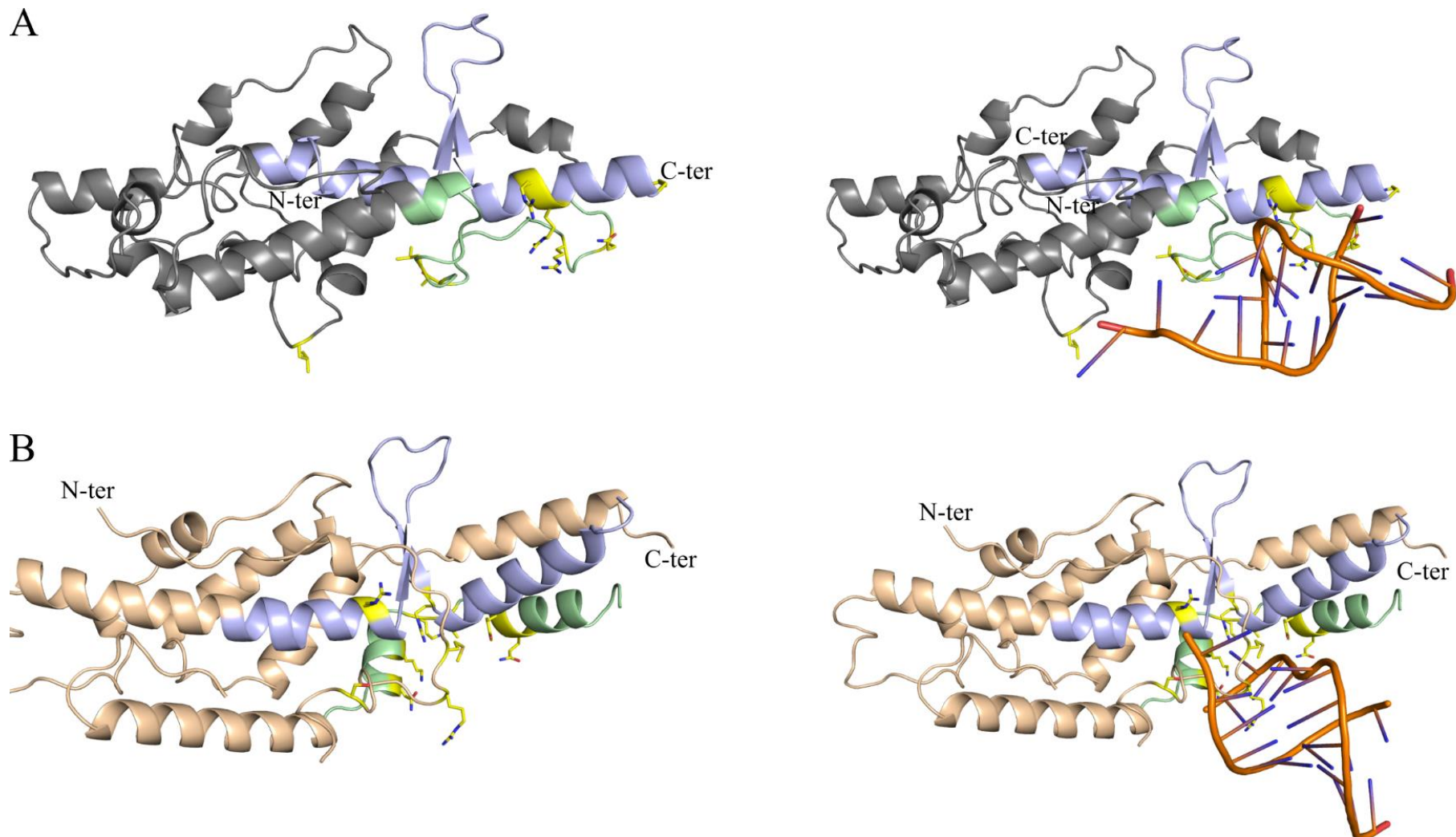


Figura 38. Detalhe da região de interação entre o componente TBE e o domínio TRB

(A). Representação em *cartoon* da melhor colução encontrada no *docking* realizado entre o domínio TRB de *L. major* e o elemento TBE de *T. thermophila*. O motivo T está destacado em azul e a região homóloga ao motivo CP está destacada em verde, os aminoácidos preditos de interagir diretamente com o RNA formando enlace de hidrogênio estão destacados como *sticks* amarelos. O modelo é apresentado com e sem o RNA. (B) Representação em *cartoon* do detalhe da interação entre o componente TBE e o domínio TRB de *T. thermophila* PDB:5C9H, O motivo T está destacado em azul e a região homóloga ao motivo CP está destacada em verde, os aminoácidos preditos de interagirem diretamente com o RNA formando enlace de hidrogênio estão destacados como *sticks* amarelos. O modelo é apresentado com e sem o RNA. É possível ver que a estrutura da região de interação predita no *docking* não é conservada em *L. major* e adicionalmente, que o elemento TBE interage com outras regiões ricas em hélices- α e aminoácidos carregados positivamente presentes na estrutura. As representações das proteínas e os ácidos nucleicos foram geradas com o programa PyMol.

Os resultados do *docking I* permitem inferir que região predita de interação entre o domínio TRB de *L. major* e o elemento TBE não é completamente conservada quando comparada com *T. thermophila* (Figura 38). Os dados do *docking I* também sugerem que a interação entre o componente TER e o elemento TBE não é estabilizada por aminoácidos das mesmas regiões da proteína em *L. major* e *T. thermophila*, já que o TBE em *L. major* parece interagir com aminoácidos presentes na hélice- α mais carboxi-terminal do domínio que é rica em aminoácidos de carga positiva (Figura 38A).

Para aprimorar a comparação estrutural da interação entre o domínio TRB e o elemento TBE em *T. thermophila* e *L. major* foram feitas análises de interações hidrofóbicas e eletrostáticas utilizando o programa LIGPLOT (WALLACE et al., 1995). Os resultados destas análises são mostrados nas Figuras 39-42 e são sumarizadas na Tabela 6.

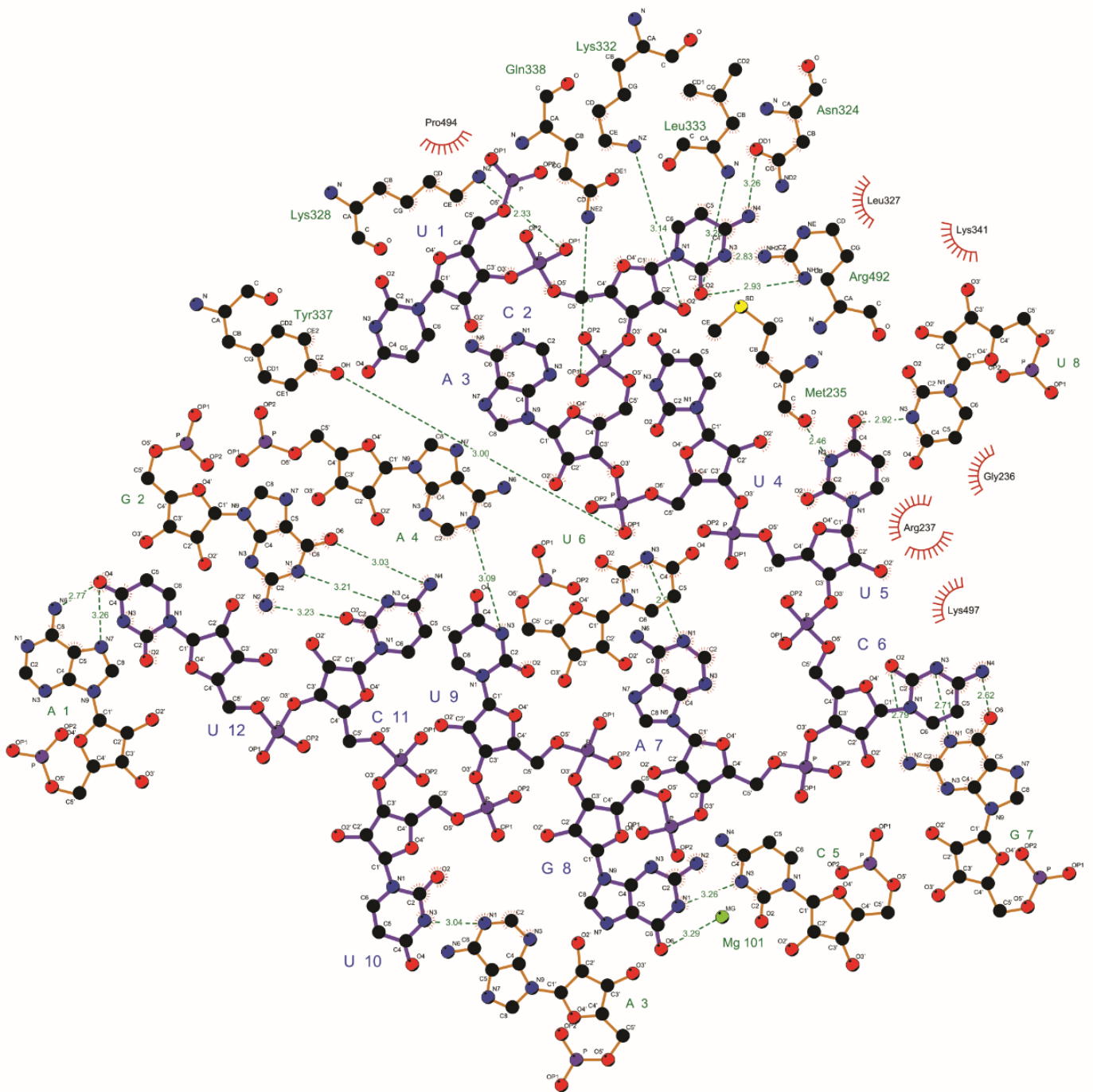


Figura 39. Análise de interações entre o domínio TRB de *T. thermophila* e a fita A do elemento TBE
 Diagrama de interações por ligação de hidrogênio (traço verde) e interações hidrofóbicas (pestanas vermelhas) entre o domínio TRB de *T. thermophila* e a fita A do elemento TBE do mesmo organismo (5C9H). Figura gerada com o programa LIGPLOT.

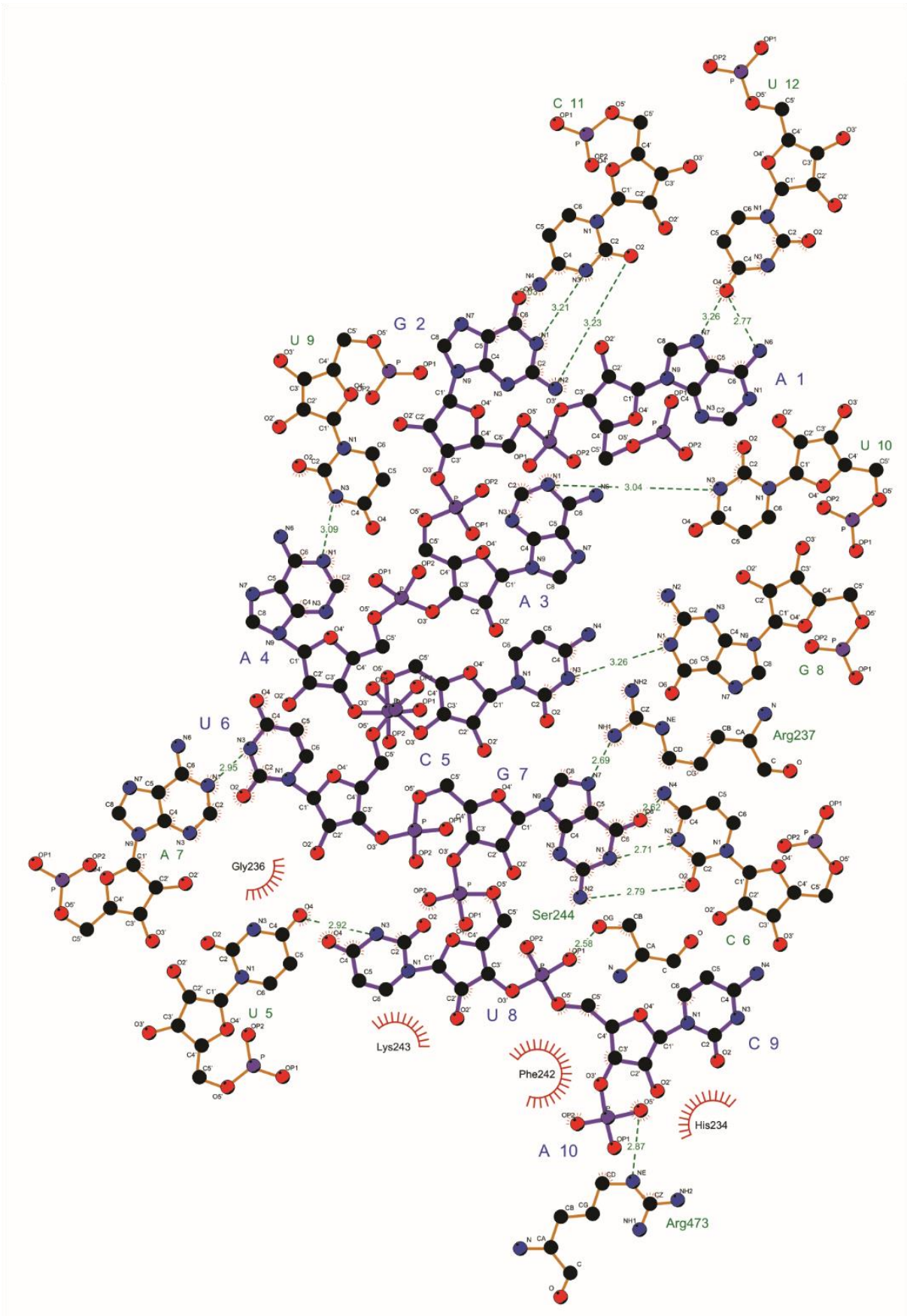


Figura 40. Análise de interações entre o domínio TRB de *T. thermophila* e a fita B do elemento TBE
 Diagrama de interações por ligação de hidrogênio (traço verde) e interações hidrofóbicas (pestanas
 vermelhas) entre o domínio TRB de *T. thermophila* e a fita B do elemento TBE do mesmo organismo
 (5C9H). Figura gerada com o programa LIGPLOT.

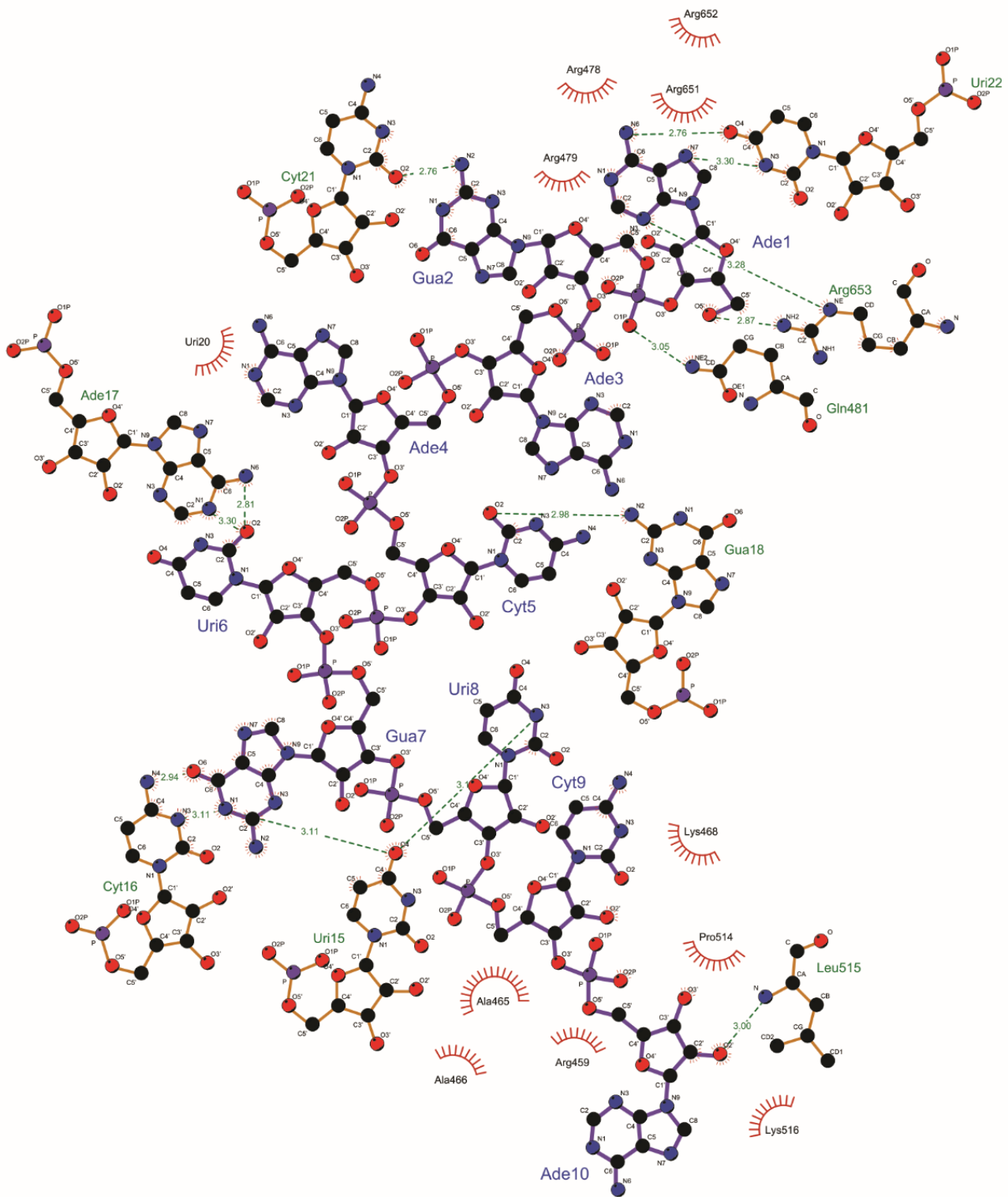


Figura 41. Análise de interações entre o domínio TRB de *L. major* e a fita A do elemento TBE
 Diagrama de interações por ligação de hidrogênio (traço verde) e interações hidrofóbicas (pestanas vermelhas) entre o domínio TRB de *L. major* e a fita A do elemento TBE *T. thermophila* (*Docking 1*).
 Figura gerada com o programa LIGPLOT.

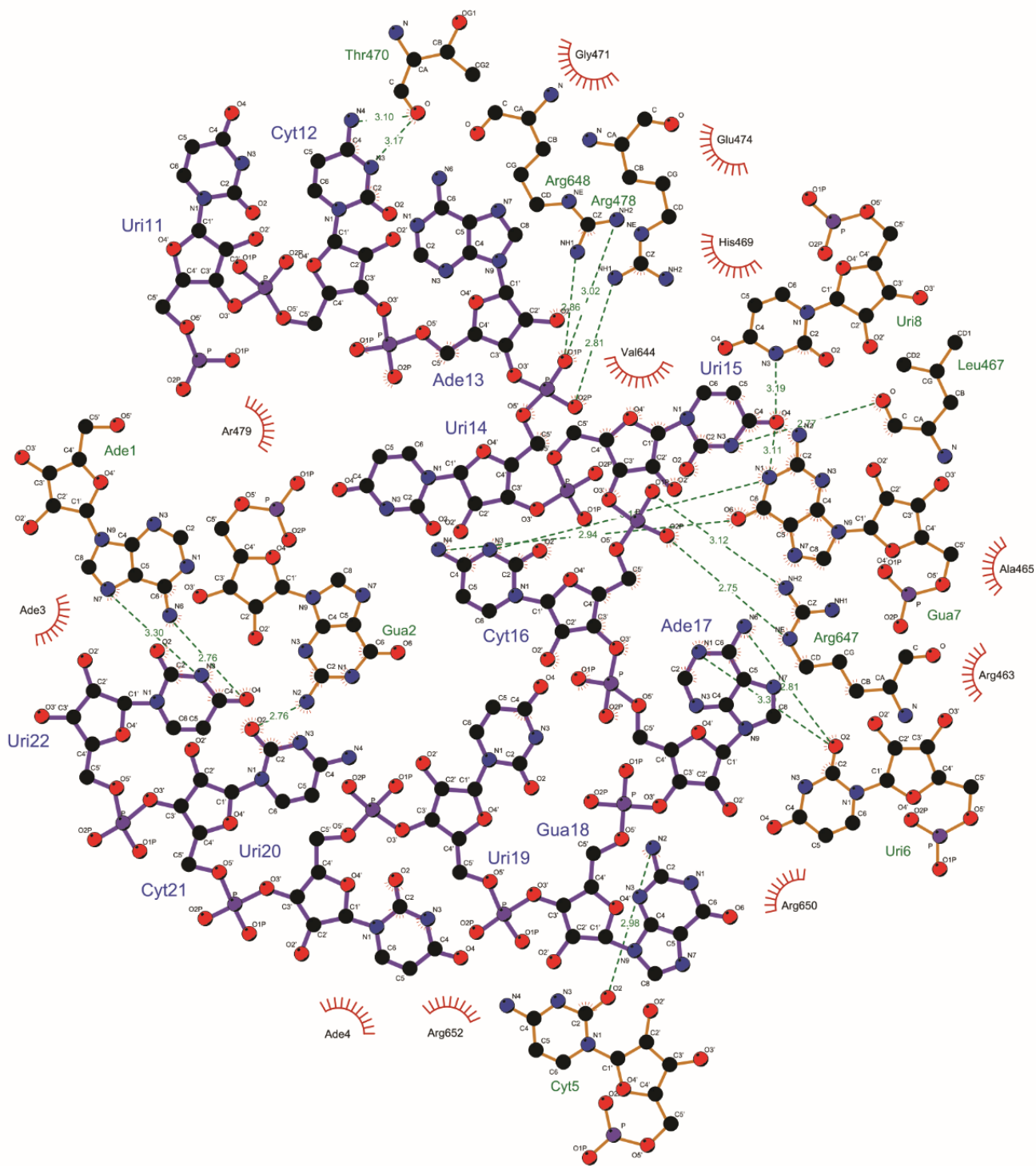


Figura 42. Análise de interações entre o domínio TRB de *L. major* e a fita B do elemento TBE
 Diagrama de interações por ligação de hidrogênio (traço verde) e interações hidrofóbicas (pestanas vermelhas) entre o domínio TRB de *L. major* e a fita B do elemento TBE *T. thermophila* (*Docking 1*).
 Figura gerada com o programa LIGPLOT.

Tabela 6. Resumos das interações hidrostáticas e ligações de hidrogênio entre o componente TBE e o domínio TBD em *T. thermophila* e *L. major*

Estrutura	Ligações de hidrogênio (LH)	Distâncias das LH (Å)	Resíduos envolvidos em interações hidrofóbicas
<i>T. thermophila</i> TRBD (5C9H)	Met235(O)-(N3)Ura5	2,46	
	Arg237(NH1)-(N7)Gua7	2,69	
	Ser244(OG)-(OP1)Cyt9	2,58	
	Asn324(OD1)-(N4)Cyt2	3,26	
	Lys328(NZ)-(OP1)Cyt2	2,33	
	Lys332(NZ)-(O2')Cyt2	3,14	His234; Gly236; Arg237;
	Leu333(N)-(O2)Cyt2	3,28	Phe242; Lys243; Leu327;
	Tyr337(OH)-(OP1)Ura4	3,00	Lys341; Pro494; Lys497
	Gln338(NE2)-(OP1)Ade3	2,80	
	Arg473(NE)-(O5')Ade10	2,87	
	Arg492(NH1)-(O2)Cyt2	2,93	
	Arg492(NH2)-(N2)Cyt2	2,83	
<i>L. major</i> TRBD modelo <i>in silico</i>	Leu467(O)-(N3)Ura15	2,77	
	Thr470(O)-(N4)Cyt12	3,10	
	Arg478(NH1)-(OP2)Ade13	2,81	
	Gln481(NE2)-(OP1)Ade2	3,05	Arg459; Arg463; Ala465
	Leu515(N)-(O2')Ade10	3,00	Ala466; Lys468; His469;
	Arg647(NE)-(O3')Cyt16	2,75	Gly471; Glu474; Arg478;
	Arg647(NH2)-(OP1)Ade17	3,12	Arg479; Pro514; Lys516;
	Arg648(NH1)-(OP1)Ura14	2,86	Val644; Arg650; Arg651;
	Arg648(NH2)-(OP1)Ura14	3,02	Arg652
	Arg653(NE)-(N3)Ade1	3,28	
Arg653(NH2)-(O5')Ade1	2,87		

A Tabela 6 mostra que a interação entre o domínio TRB e o elemento TBE em *T. thermophila* é mediada por 12 ligações de hidrogênio e 9 resíduos aparentemente envolvidos em interações hidrofóbicas, como já previamente apresentado por JANSSEN et al., (2015). A interação predita no *docking 1* entre o domínio TRB de *L. major* e o elemento TBE de *T. thermophila* é mediado por 10 ligações de hidrogênio e 15 resíduos aparecem envolvidos em interações hidrofóbicas. Estes dados mostram que, embora a interação não esteja sendo mediada pelos mesmos aminoácidos, as interações

envolvidas entre as proteínas e o ácido nucleico são similares nos dois organismos. Estes resultados reforçam a solução encontrada no *docking 1*.

Até o momento, foi mostrado que o motivo CP não é conservado estruturalmente em *Leishmania* e que o motivo T apresenta substituições em aminoácidos de importância para estabelecer a interação entre o domínio TRB e o TER. Também, que a maior quantidade de aminoácidos que interagem diretamente com o componente TER em outros organismos não é conservada em *L. major* e que a região do motivo CP que interage com o TBE em *T. thermophila*, não apresenta conservação de estrutura terciária. Também foi predito que o elemento TBE do TER de *T. thermophila* pode interagir com outras regiões dentro do domínio TRB de *L. major*.

A interação entre os motivos T e CP do domínio TRB com a região TBE do TER é importante para a definição do limite do molde na telomerase de vertebrados e ciliados (CHEN; GREIDER, 2003; JANSSON et al., 2015). Esta definição, garante que as sequências das repetições teloméricas sejam mantidas, já que impede que regiões além do molde do TER sejam retrotranscritas e adicionadas ao terminal dos cromossomos durante a elongação dos telômeros (CHEN; GREIDER, 2003). Como mencionado anteriormente, o gênero *Leishmania* apresenta repetições teloméricas bem definidas, compostas pela sequência TTAGGG (FU; BARKER, 1998; CONTE; CANO, 2005) e o componente TER de *Leishmania* possui um elemento TBE (VASCONCELOS et al., 2014). Isto indica que o complexo telomerase do parasita também possui um mecanismo que delimita o tamanho da sequência molde que é inserida nos terminais de cromossomos. No entanto, os resultados obtidos neste trabalho sugerem que o mecanismo estrutural para a interação entre o componente TERT e o TBE não é conservado em *Leishmania* e não segue os moldes descritos para outros eucariotos.

É possível que o mecanismo de interação entre o domínio TRB e o componente TER em *Leishmania* ocorra em regiões diferentes do domínio, como sugerido pelos resultados do *docking*. Outra hipótese é que outro domínio na estrutura do TERT do parasito atue auxiliando neste processo. Tanto ciliados quanto vertebrados apresentam resíduos conservados posicionados na região do *linker*, que estão envolvidos na definição do limite do molde no TER (HARKISHEIMER et al., 2013; JANSSON et al., 2015). O motivo CP2 (Figura 43A) específico de ciliados (R226, Q228, F230, Y231, C232, H234, R237 em *T. thermophila*) faz parte do T/CP *pocket* e é importante para a definição do limite do molde (JANSSON et al., 2015). Nos vertebrados o motivo TFYL (T300, F303, L304, Y305 em *T. rubripes*) (Figura 43B) desenvolve a mesma função que o motivo CP2 em ciliados (HARKISHEIMER et al., 2013). Entretanto, nenhum destes motivos é conservado em *Leishmania* (Figura 43).

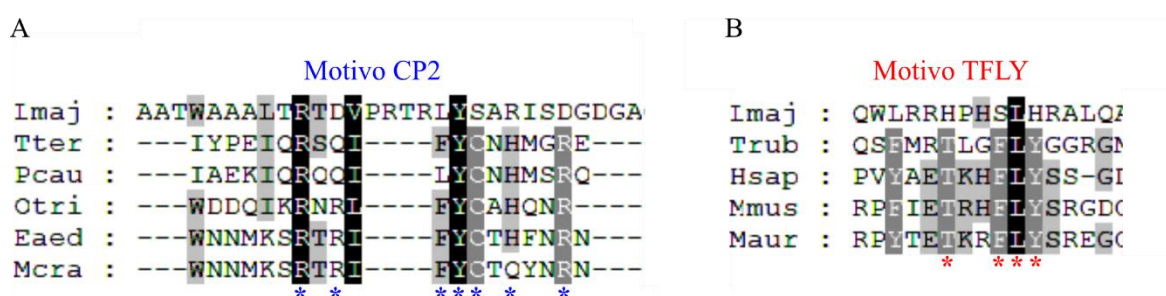


Figura 43. Motivos CP2 e TFLY

Alinhamento múltiplo de *L. major* com ciliados (A) e vertebrados (B), mostrando os domínios CP2 e TFLY (os resíduos conservados estão demarcados com asteriscos). Estes domínios não são conservados em *Leishmania*. (Lmaj: *L. major* (AAO67514.1); Hsap: *Homo sapiens* (NP_937983.2); Mmus: *Mus musculus* (NP_033380.1); Maur: *Mesocricetus auratus* (NP_001268556.1); Tter: *Tetrahymena thermophila* (XP_001010651.3); Pcau: *Panamecium caudatum* (BAB18587.1); Otri: *Oxytricha trifallax* (EJY64832.1); Eaed: *Euplotes aediculatus* (O00939.1); Mcra: *Monoeuplotes Crassus* (AAP42378.1)).

Em termos evolutivos e devido à posição basal do TERT de *Leishmania* sp. (Figura 16), estes resultados sugerem que o motivo T e o componente TBE do TER apresentam maior conservação nos eucariotos primitivos do que o motivo CP.

4.5 Expressão do domínio TRB de *L. major* em modelo heterólogo

Com o intuito de se validar experimentalmente os resultados obtidos *in silico*, o fragmento TRB (aminoácidos 408-745), foi primeiramente clonado no vetor de expressão bacteriano pET28a+. Porém, após diversas tentativas de se obter o domínio na forma recombinante, utilizando diferentes cepas de bactérias e alterando variáveis como temperatura de crescimento e concentração de IPTG, o domínio recombinante não foi expresso. Por esta razão, optou-se por alterar o sistema de expressão para a levedura *P. pastoris*, que por ser um eucarioto, é mais provável de expressar proteínas eucariotas, que não são expressas em sistema bacteriano. Além disso, nesse sistema existe a possibilidade de se obter a proteína recombinante secretada no meio de cultura, o que garante que ela seja expressa na forma solúvel facilitando a purificação diretamente a partir do meio.

O vetor pPiCZ- α contendo o domínio TRB clonado (Figura 15) foi utilizado para transformar bactérias DH5- α , a partir das quais foi obtido DNA plasmidial por mini-prep. O vetor foi então linearizado e utilizado para transformar *P. pastoris* da cepa Km71H.

O vetor pPiCZ- α linearizado pode se inserir em diferentes locais no genoma das leveduras, fazendo com que os níveis de expressão da proteína heteróloga variem dependendo do local de inserção. Assim, um primeiro teste de mini-indução foi realizado utilizando-se 4 colônias transformantes (colônias 2, 8, 12 e 13). Estas colônias foram escolhidas por mostrarem resistência a maiores concentrações do antibiótico zeocina, o que indica indiretamente uma maior expressão da proteína recombinante.

A expressão da proteína heteróloga foi induzida em culturas de leveduras em meio líquido, por uma semana, e amostras de sobrenadante do meio de cultura foram

coletadas todos os dias. Com estas amostras foi realizado um ensaio de *western blot* utilizando-se soro anti-TERT de *L. amazonensis* para conferir a expressão e secreção da proteína. Como mostrado na Figura 44 (linhas 6, 8 e 9), sobrenadantes de cultura das colônias 2 e 8 contêm a proteína recombinante. Os sobrenadantes de cultura das colônias 12 e 13 não mostraram expressão da proteína heteróloga (resultados não mostrados). O *western blot* mostra uma banda do tamanho esperado (39,3 kDa) no sobrenadante da colônia 8, após 96 e 120 horas de indução, enquanto que no sobrenadante da colônia 2, a proteína recombinante é expressa após 144 horas de indução. Porém, nessa amostra o soro também reconhece uma banda inespecífica ou produto de degradação da proteína expressa (Figura 44, linha 6, asterisco). Por essa razão decidiu-se padronizar o protocolo de purificação a partir da colônia 8.

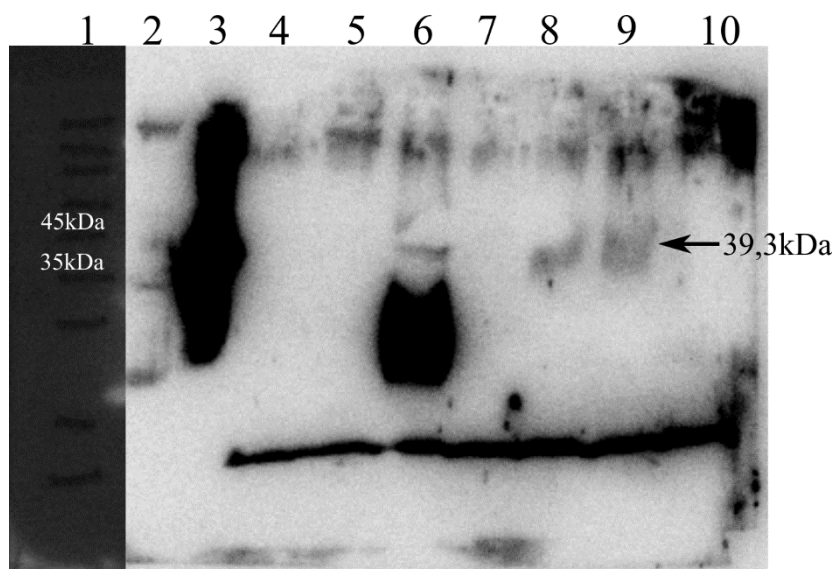


Figura 44. Western blot da mini indução do domínio TRB de *L. major* em *P. pastoris*

Amostras do sobrenadante do meio de cultura de leveduras induzidas foram fracionadas em SDS-PGE 15%, transferidos a membrana PVDF e reveladas usando soro anti-TERT de *L. amazonensis* produzido em coelho e anticorpo secundário anti coelho produzido em cabra, usando método de quimioluminescência por ECL plus. 1: marcador de peso molecular PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific). 2: extrato de bactérias não induzidas (controle negativo); 3: extrato de bactérias induzidas com o domínio TEN de *L. major* (controle positivo). 4: Sobrenadante da cultura de leveduras, colônia 2 não induzida. 5: Sobrenadante da cultura de leveduras, colônia 2, 120 horas após indução. 6: Sobrenadante da cultura de leveduras, colônia 2, 144 horas após indução. 7: Sobrenadante da cultura de leveduras, colônia 8 não induzida. 8: Sobrenadante da cultura de leveduras, colônia 8, 96 horas após indução. 9: Sobrenadante da cultura de leveduras, colônia 8, 120 horas após indução. 10: Sobrenadante da cultura de leveduras, colônia 8, 144 horas após indução. A seta aponta a banda do tamanho esperado (39,3 kDa) correspondente ao domínio TRB recombinante, o asterisco aponta a banda inespecífica reconhecida na colônia 2.

A fim de expressar o domínio TRD em grande escala a partir de cultura de *P. pastoris*, foram realizados vários ensaios, no entanto não foi possível se obter o domínio em quantidades suficientes para a realização dos ensaios de cromatografia para a sua purificação. Provavelmente a dificuldade para se expressar o domínio TRB de *L. major* em leveduras e bactérias deva-se ao fato dele ser tóxico ou muito instável, por isso em nenhum dos dois casos se obteve a proteína recombinante em quantidades suficientes para ser purificado.

4.6 Caracterização do domínio RT de *Leishmania* sp.

O domínio RT mostra maior conservação de estrutura primária no gênero *Leishmania*, o que permitiu usar as sequências do domínio RT das telomerasas de diferentes organismos para encontrar o gene que codifica a telomerase de *Leishmania* sp. antes mesmo que o genoma do parasito estivesse anotado. E com isso foi possível se clonar e caracterizar o gene de diferentes espécies do parasito como descrito em GIARDINI et al., (2006). A conservação na sequência de aminoácidos deste domínio pode ser observada no alinhamento múltiplo de sequências realizado utilizando-se a ferramenta MUSCLE (Figura 45), a pesar de se apresentarem algumas inserções específicas de Tripanosomatídeos. O alinhamento também permitiu delimitar os sete motivos da região catalítica compartilhados entre TERT e retrotranscriptases clássicas (1-2 e A-E), assim como os dois motivos exclusivos dos TERT (3 e IFD) (Figura 45).

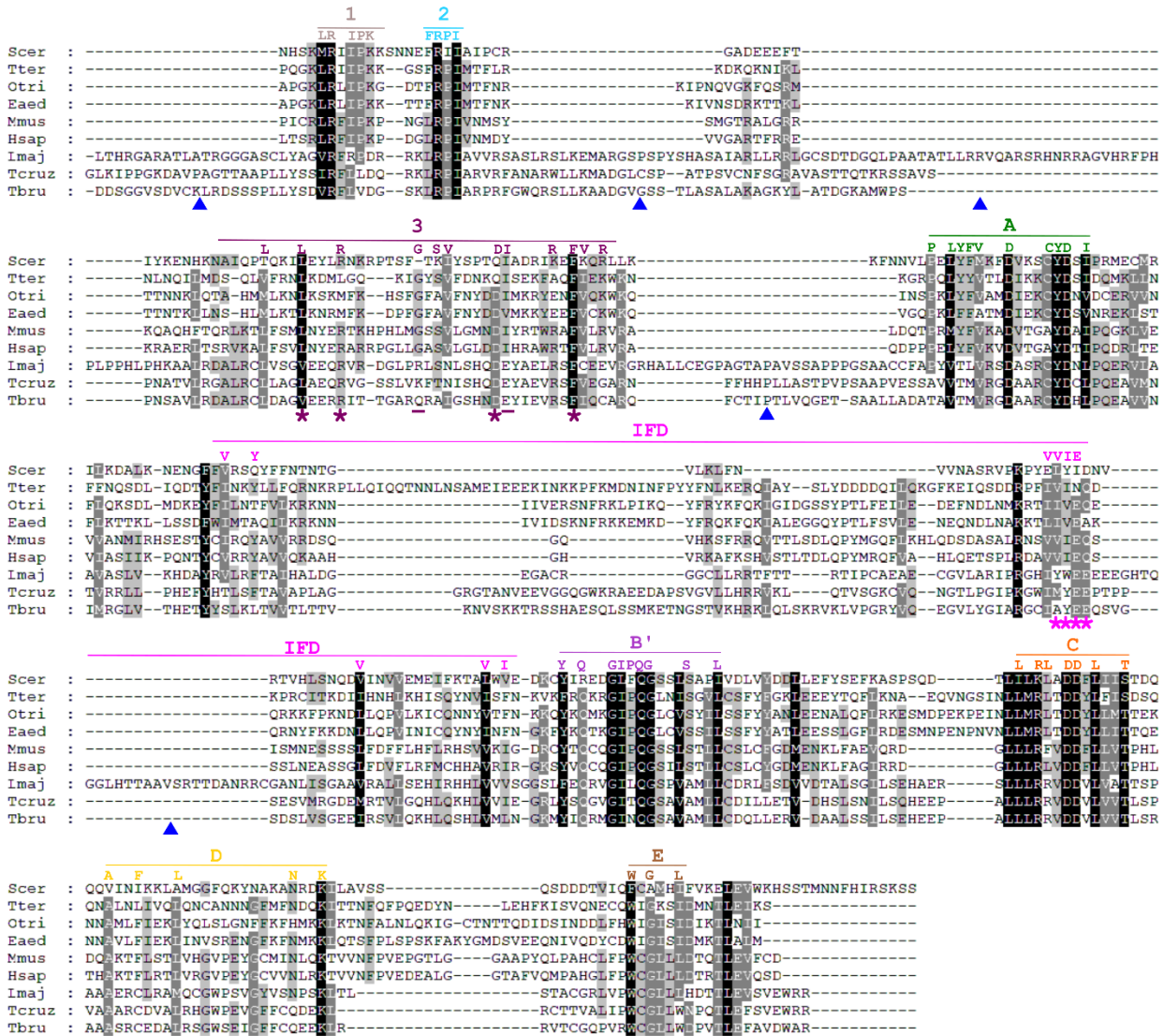


Figura 45. Alinhamento de sequência múltiplo do domínio RT

Motivos conservados no domínio RT foram identificados em LmTERT por alinhamento múltiplo de sequência utilizando a sequência de aminoácidos do TERT de *L. major* e TERTs de outros organismos. Os diferentes motivos (1, 2, 3, A, IFD, B', C, D e E) são destacados, inserções são denotadas com triângulos. Uma sequência consenso é mostrada acima de cada domínio. No motivo 3, asteriscos abaixo da sequências denotam resíduos de importância funcional conservados em *Leishmania* e os sinais “menos” denotam resíduos de importância funcional não conservados em *Leishmania*. No motivo IFD, os asteriscos denotam o motivo de importância para a processividade do TERT de *S. cerevisiae*. O domínio RT foi delimitado a partir de GIARDINI et al., (2006) e das sequências do TERT dos seguintes organismos: (Scer: *Saccharomyces cerevisiae* (NP_013422.2); Tter: *Tetrahymena thermophila* (XP_001010651.3); Otri: *Oxytricha trifallax* (EJY64832.1); Eaed: *Euplotes aediculatus* (O00939.1); Mmus: *Mus musculus* (NP_033380.1); Hsap: *Homo sapiens* (NP_937983.2); Lmaj: *L. major* (AAO67514.1); Tcruz: *Trypanosoma cruzi* (XP_819198.1); Tbru: *Trypanosoma brucei* (XP_829083.1)). Resíduos destacados em preto mostram conservação em todos os organismos e em cinza conservação na maioria dos organismos.

4.6.1 Análise dos motivos IFD e 3

Os motivos 3 e IFD foram caracterizados em humanos e *S. cerevisiae* respectivamente como motivos importantes para a velocidade na adição de nucleotídeos e a processividade da atividade de telomerase (LUE et al., 2003; XIE et al., 2010). Dado que a atividade de telomerase em *L. major* é baixa e pouco processiva, é interessante ver o nível de conservação dos dois motivos neste organismo.

Tabela 7. Resíduos do motivo 3 importantes para a processividade do TERT humanos, o asterisco denota um resíduo não conservado

Organismo	Resíduos					
<i>S. cerevisiae</i>	L482	R686	T493*	Q501*	I502	F509
<i>T. thermophila</i>	L575	L579*	G584	D592	I593	F600
<i>H. sapiens</i>	L665	R669	G677	D685	I686	F693
<i>L. major</i>	V859	R863	R870*	D878	E879*	F886
<i>T. brucei</i>	V666	R670	Q677*	D685	E686*	F693

A Tabela 7 mostra o nível de conservação dos diferentes resíduos do motivo 3 que são importantes para a processividade da enzima. Em termos gerais, estes resíduos são conservados em *L. major*. Os dois únicos resíduos que não são conservados (R870 e E879) (Tabela 7) também não são conservados em *T. brucei*, cuja atividade de telomerase e processividade são maiores do que em *Leishmania* sp..

O *cluster* L589, Y590, I591, D592 faz parte do motivo IFD de *S. cerevisiae* e é de importância para a processividade da enzima neste organismo. Este *cluster* apresenta uma conservação parcial em *L. major* (Y1008, W1009, E1010, E1011) (Figura 45, marcado por asteriscos), em *T. brucei*, o motivo IFD apresenta uma conservação similar. Juntos, estes dados sugerem que a falta de processividade do TERT de *L. major* não é devida à baixa conservação nestes motivos. Como não existem dados estruturais referentes aos motivos 3 e IFD eles não serão discutidos no ambiente tridimensional.

4.6.2 Caracterização estrutural da fração catalítica do TERT de *L. major*

Para se compreender o arranjo estrutural do centro catalítico da telomerase de *L. major* e principalmente, se os domínios TRB, RT e CTE adquirem a configuração anelar similar à estrutura característica de outras telomerases (GILLIS et al., 2008; JIANG et al., 2015), foi realizada uma predição de estrutura terciária dos três domínios juntos. O modelo obtido para *L. major*, foi gerado a partir da estrutura cristalográfica do TERT de *T. castaneum* que é a única estrutura cristalográfica de um TERT inteiro resolvida atualmente. O modelo obtido compreende a região que contém a fração carboxi-terminal do domínio TRB, a totalidade do domínio RT e a fração amino-terminal do domínio CTE, correspondendo aos resíduos 502 a 1254. A fim de se obterem modelos *in silico* mais verossímeis, o modelo foi submetido a simulações de dinâmica molecular utilizando GROMACS (Groningen Machine for Chemical Simulation) v.4.5.3 campo de força GROMOS 96 53 a 6 por 100ns.

A Figura 46 mostra a superfície das duas estruturas, onde é perceptível que a estrutura de *L. major* está incompleta e por isso, talvez, a cavidade onde se acomoda o híbrido de ácidos nucléicos não é tão aparente. Entretanto, a conformação anelar e a proximidade entre os domínios TRB e CTE são conservados, o que sugere que os domínios do TERT em *Leishmania* se acomodam em uma conformação anelar de modo similar aos outros TERT descritos (GILLIS et al., 2008; JIANG et al., 2015).

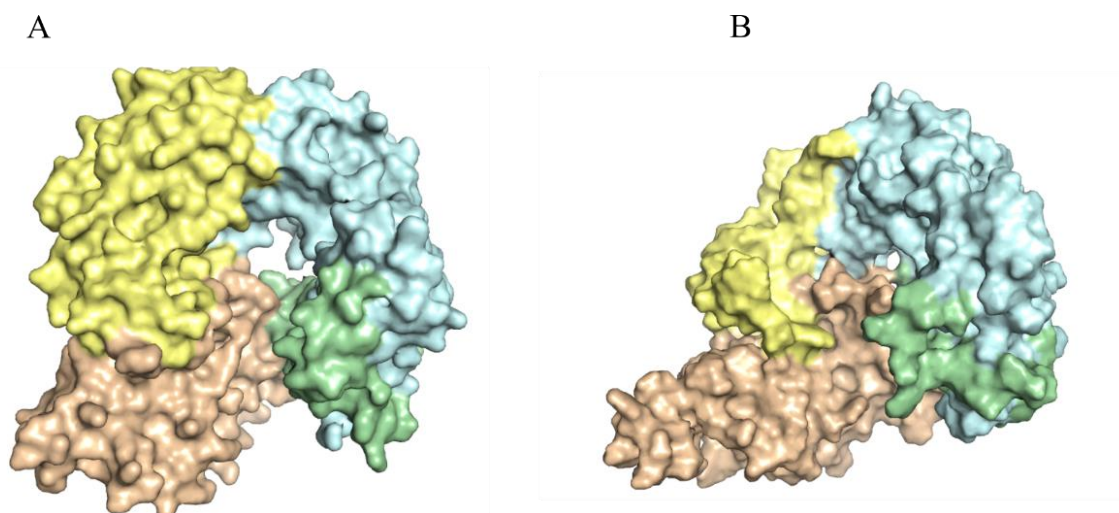


Figura 46. Comparação das superfícies da telomerase de *T. castaneum* e de *L. major*

Representação da superfície da estrutura cristalográfica de *T. castaneum* (A) e da predição de estrutura terciária de *L. major* (B). Ambas mostram a conformação anelar criada pelos domínios TRB (marrom), RT, (subdomínio fingers em verde e palm em azul) e CTE (amarelo). A cavidade onde se acomoda o híbrido de ácidos nucleicos não é tão aparente em *L. major*, provavelmente porque o modelo não mostra os domínios TRB e CTE completos. As representações da superfície das proteínas foram geradas com o programa PyMol.

Devido à baixa informação estrutural e funcional disponível sobre o domínio CTE e a impossibilidade de se gerar uma predição de estrutura terciária deste domínio, ele não foi estudado individualmente no presente trabalho.

4.6.3 Caracterização estrutural do domínio RT de *L. major*

A fim de se estudar a conservação estrutural do domínio RT, este foi isolado da dinâmica molecular da predição de estrutura terciária dos três domínios juntos (TRB, RT e CTE).

O primeiro passo foi verificar a conservação estrutural entre o domínio de *L. major* e o de *T. castaneum*. Para isso, foi realizada uma sobreposição de carbonos α das duas estruturas (Figura 47). A sobreposição de $C\alpha$ entre os dois domínios revela que há uma relativa conservação estrutural (2,8 Å de r.m.s.d), e que os dois modelos apresentam os mesmos componentes estruturais.

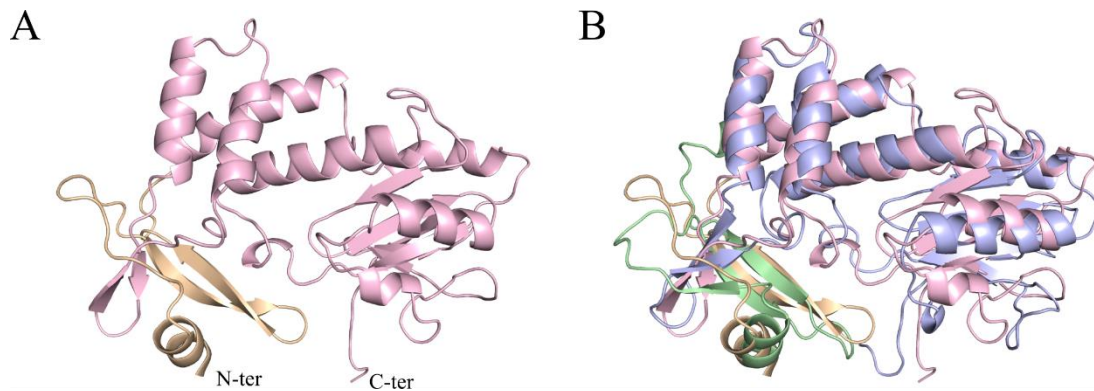


Figura 47. Estrutura terciária do domínio RT de *L. major*

(A) Representação em *cartoon* do modelo estrutural *in silico* do RT de *L. major* (resíduos 745 a 1195) obtido por técnicas de *threading* a partir do modelo cristalográfico previamente reportado 3DU5 (TERT *T. castaneum*) e 100 ns de simulação de dinâmica molecular (MD). O subdomínio dos dedos (*fingers*) está representado em marrom e o subdomínio da palma (*palm*) está representado em rosa (B) superposição de $C\alpha$ da estrutura cristalográfica do domínio TRB de *T. castaneum* (3DU5) (Subdomínio da palma (*palm*) em azul e subdomínio dos dedos (*fingers*) em verde) e o modelo *in silico* do domínio TRB de *L. major* (Subdomínio da palma (*palm*) em rosa e subdomínio dos dedos (*fingers*) em marrom). A sobreposição $C\alpha$ é muito boa, as duas estruturas compartilham elementos estruturais. As representações das proteínas foram geradas com o programa PyMol.

Em seguida, verificou-se se os motivos presentes no domínio RT eram conservados estruturalmente em *L. major* (Figura 48). Para isso, os motivos foram destacados dentro da estrutura de *T. castaneum* (Figura 48) e no modelo de *L. major* (Figura 48). Embora existam diferenças na orientação de alguns dos motivos, é possível se observar que os 7 motivos característicos do domínio RT, apresentam conservação estrutural. Verifica-se também que tanto em *T. castaneum* como em *L. major*, todos os motivos estão formados pelos mesmos elementos de estrutura secundária.

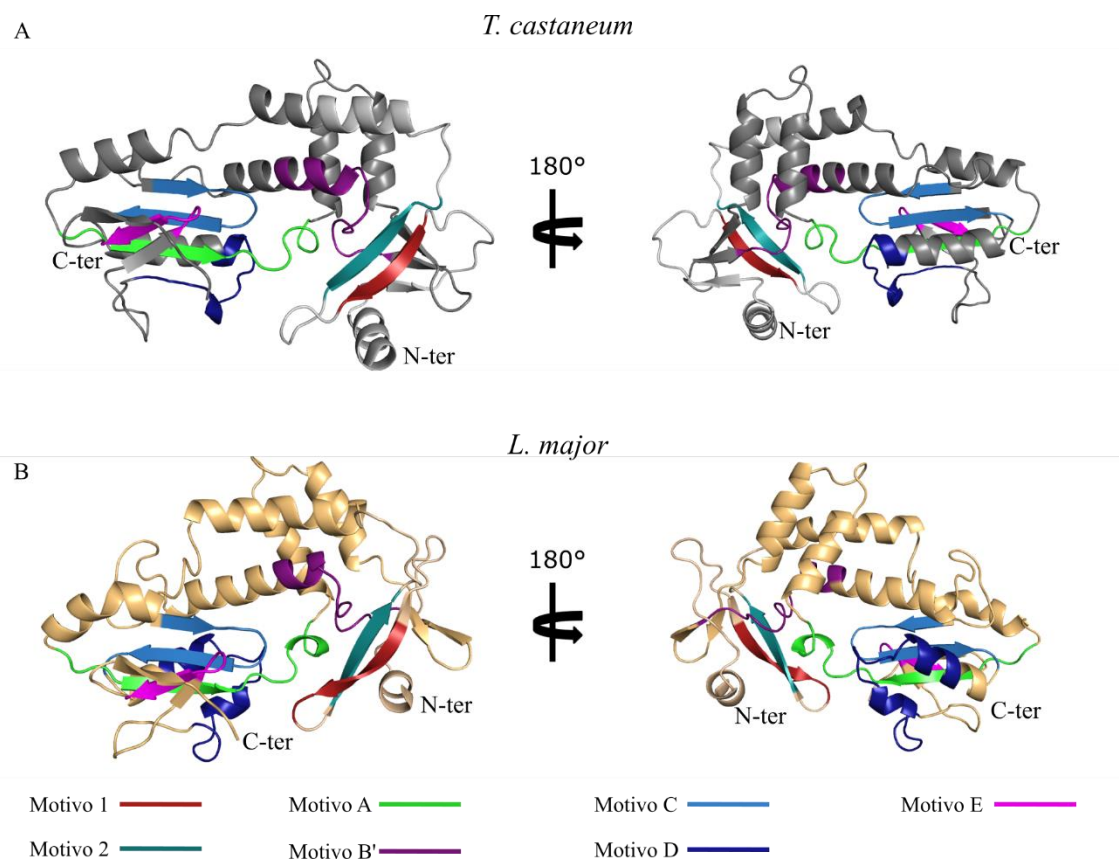


Figura 48. Conservação estrutural dos motivos do domínio RT em *L. major*

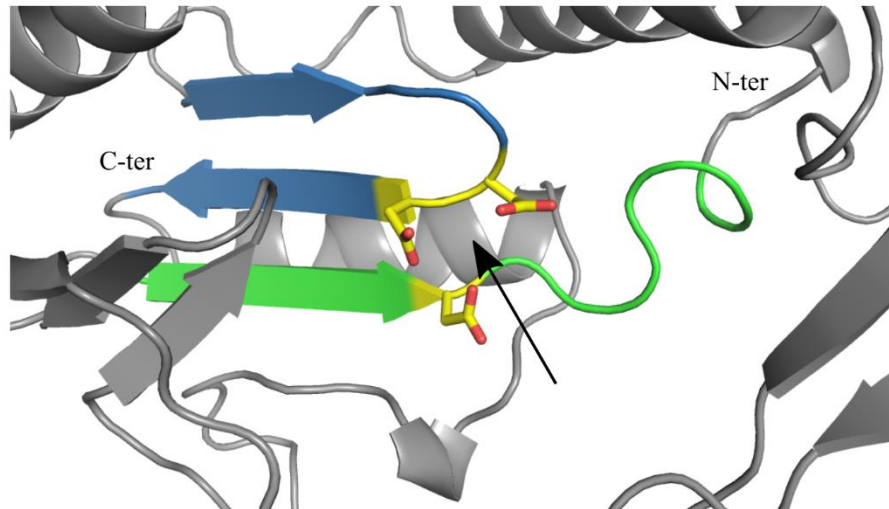
Representação em *cartoon* da estrutura do domínio RT de *T. castaneum* (A), e de *L. major* (B). Nas duas estruturas os motivos conservados estão destacados na mesma cor e pode se observar que em todos os casos estão compostos pelos mesmos elementos estruturais. As representações das proteínas foram geradas com o programa PyMol.

Uma região de grande importância dentro do domínio RT é o sítio catalítico da enzima. Uma característica própria de todas as polimerases é o uso de íons metálicos bivalentes como cofatores na transferência de nucleotídeos estabilizados por resíduos ácidos (D, aspartato ou E, glutamato) localizados próximos do sítio ativo (MITCHELL et al., 2010). No caso dos TERT, estes resíduos são 3 aspartatos (D), 1 localizado no motivo A e 2 localizados no motivo C, que estabilizam um íon de Mg^{2+} (WEINRICH et al., 1997; LINGNER et al., 1997). A importância destes resíduos foi mostrada anteriormente, pois a substituição dos mesmos por alaninas elimina a atividade de telomerase (WYATT et al., 2010). Adicionalmente, comparações estruturais entre o domínio RT de *T. castaneum* e a transcriptase reversa do vírus HIV mostraram que a

localização dos três aspartatos (D251, D343, D344 em *T. castaneum*) próximos ao sítio ativo é conservada, e que sua disposição permite inferir que eles estabilizam um íon de Mg^{2+} (Figura 49A) (GILLIS et al., 2008; MITCHELL et al., 2010).

Os aspartatos que formam o sítio ativo são conservados na sequência de *L. major* (D931, D1115, D1116) e na estrutura terciária apresentam uma disposição similar à do sítio ativo de *T. castaneum* (Figura 49B). Estes dados sugerem que o sítio ativo é conservado em *L. major*.

A



B

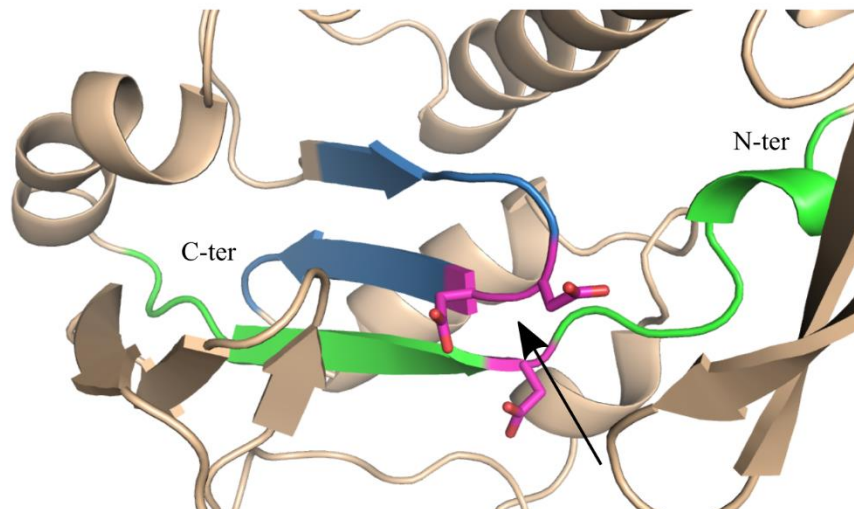


Figura 49. Conservação estrutural do sitio ativo do TERT em *L. major*

Representação em *cartoon* da estrutura do centro catalítico do TERT de *T. castaneum* (A), e *L. major* (B). Os aspartatos [D] que estabilizam o íon de magnésio estão destacados e coloridos em amarelo para *T. castaneum* e em magenta para *L. major*. A disposição dos glutamatos sugere que o sitio ativo do TERT é conservado em *L. major*. A seta mostra o local onde é estabilizado o íon permitindo a adição de repetições teloméricas. As representações das proteínas foram geradas com o programa PyMol.

Estas informações sugerem que o domínio RT de *Leishmania* sp. apresenta grande conservação de sequência primária e estrutural com telomerasas canônicas,

sugerindo que o mecanismo de catálise é conservado nas telomerasas de todos os eucariotos.

4.7 Obtenção do domínio RT recombinante

Para se obter informações experimentais, de forma a validar os resultados obtidos *in silico*, a expressão do domínio RT clonado no vetor de expressão bacteriano pET28a+ (Figura 14) foi testada em várias cepas bacterianas e condições variadas de temperatura e concentração IPTG (resultados não mostrados). Dentre todas as condições testadas, os melhores resultados foram obtidos com a linhagem BI21 Artic Express pRIL e indução da expressão da proteína recombinante a 25°C com 1mM de IPTG por 4 horas (Figura 50, linha 4h P). Os resultados dos testes de mini-indução mostraram que o domínio RT é expresso na fração insolúvel do extrato proteico bacteriano, provavelmente dentro de corpúsculos de inclusão, semelhante ao que ocorre com o domínio TEN (Figura 20) apresentado anteriormente. Por este motivo, as primeiras tentativas de obtenção do domínio RT recombinante foram utilizando o protocolo estabelecido para a obtenção do domínio TEN (Figura 21).

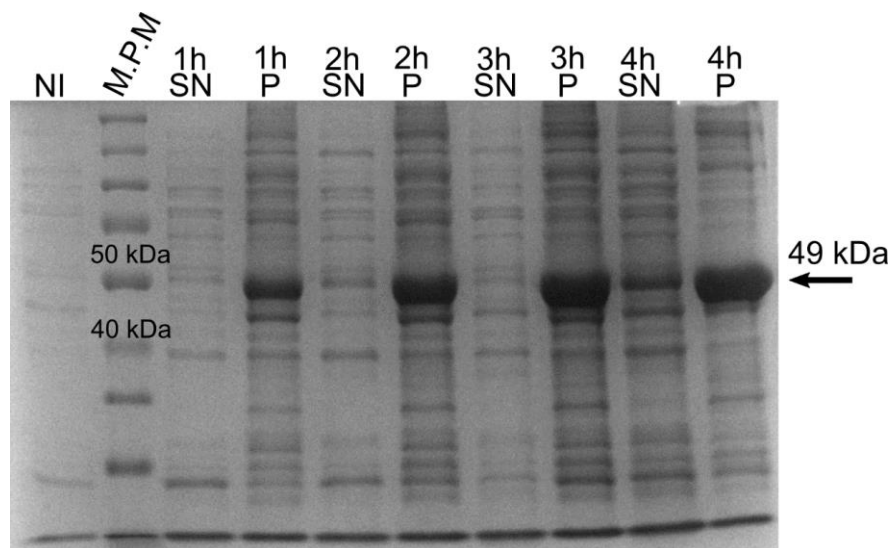


Figura 50. Teste de expressão e de solubilidade do domínio RT de *L. major* usando *E. coli* BL21 Artic Express pRIL

Extratos proteicos foram obtidos de *E. coli* BL21 Express pRil expressando o domínio RT, cultivadas a 25°C. A expressão da proteína recombinante foi induzida com 1mM de IPTG. As amostras foram analisadas a partir de 1, 2, 3 e 4 horas (h) após a indução com IPTG, utilizando-se fracionamento em SDS-PAGE 12% corado com azul de Coomassie. N.I. extrato de bactérias não induzidas; MPM, marcador de peso molecular (Spectra Multicolor Broad Range Protein Ladder (Thermo Fisher Scientific)); SN: Extrato solúvel de bactérias induzidas e P: extrato insolúvel de bactérias não induzidas. A proteína expressa em fusão com 6xHis, apresenta peso molecular estimado em 49kDa, apontado por uma seta.

As diversas tentativas de purificação do domínio RT usando cromatografia líquida não foram bem-sucedidas (dados não mostrados), devido ao fato que o extrato obtido era muito viscoso impedindo sua entrada na coluna e consequente purificação. Para resolver este problema, foi utilizado um protocolo de recuperação da proteína recombinante de corpúsculos de inclusão. Ele consiste em submeter a fração insolúvel do extrato bacteriano a várias lavagens com diferentes tampões, a fim de realizar uma purificação parcial da proteína. Após a recuperação da proteína a partir dos corpúsculos de inclusão, verificou-se que foram eliminados quase todos os contaminantes do extrato (Figura 51 linhas 1-4). Assim, essa amostra foi submetida ao processo de re-novelamento da proteína mediante diálise escalonada. Como após a diálise não houve precipitação da amostra, a purificação e re-novelamento do domínio RT heterólogo foi

verificada por fracionamento em SDS-PAGE, usando a fração solúvel obtida após a diálise. Como mostrado na Figura 51 pode se verificar que o resultado do isolamento resultou em uma única banda com o peso estimado (49kDa) do domínio RT recombinante (Figura 51 linhas 1-4). Este resultado sugere, que o protocolo utilizado para o isolamento de corpúsculos de inclusão foi suficientemente eficiente para purificar o domínio RT e que ele pode ser re-enovelado mediante diálise escalonada.

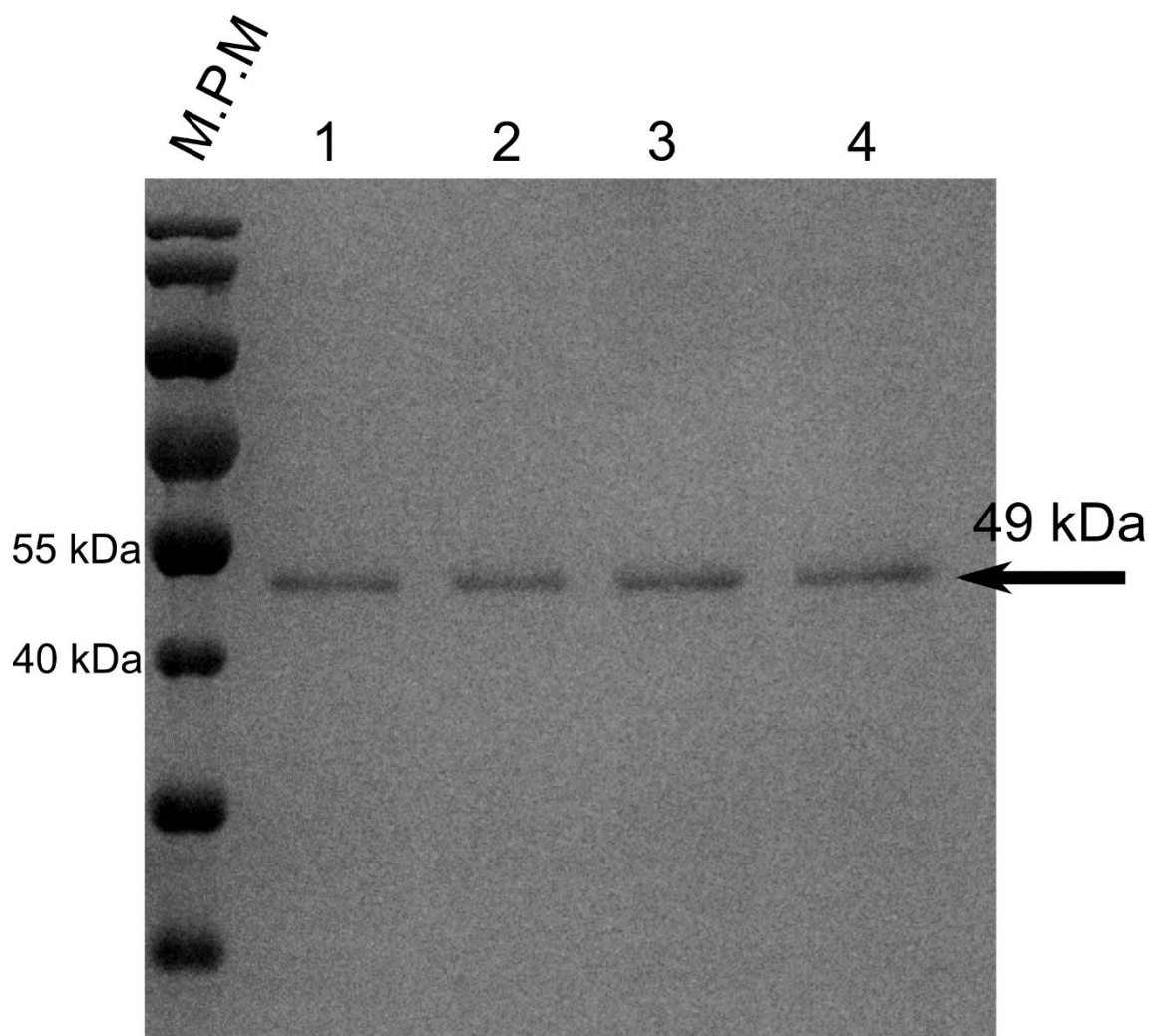


Figura 51. Purificação do domínio RT mediante isolamento de corpúsculos de inclusão
SDS-PAGE 12% corado com azul de Coomassie. M.P.M: Marcador de peso molecular: Page Ruler™ Prestained protein ladder (Thermo Fisher Scientific). As frações numeradas mostram os extratos proéticos após o isolamento de corpúsculos de inclusão e a diálise escalonada. Os extratos mostram uma única banda com peso molecular esperado para o domínio RT (49kDa).

A proteína purificada foi em seguida submetida a *western blot* revelado com soro anti-His, para se confirmar se a banda visualizada no gel correspondia ao domínio RT recombinante. O resultado mostra que a proteína recuperada dos corpúsculos de inclusão é o domínio RT (Figura 52).

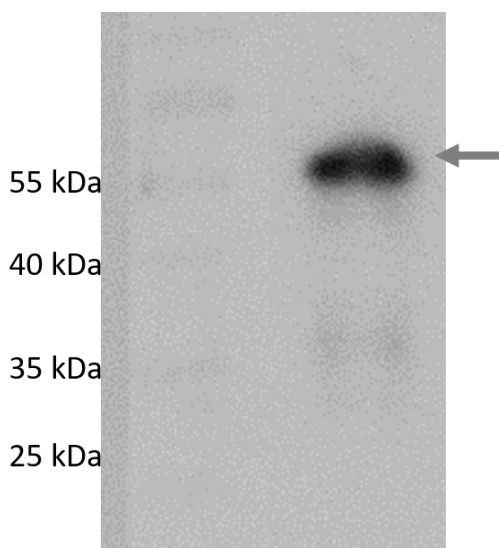


Figura 52. *Western blot* do domínio RT do TERT de *L. major* após purificação por isolamento de corpúsculos de inclusão

O domínio RT purificado foi fracionado em SDS-PAGE 12%, transferido para membrana PVDF e relevado utilizando um anticorpo primário comercial anti-His produzido em camundongo (Bio-Rad) e um anticorpo secundário anti-camundongo produzido em cabra, por quimioluminescência usando ECL plus (GE healthcare). A banda reconhecida tem peso molecular esperado para o domínio RT (49 kDa) o que permite concluir que se trata do domínio RT do TERT de *L. major*.

Análises preliminares usando espectrometria de fluorescência estática sugerem que o domínio RT purificado está enovelado (Figura 53), pois quando se compara a emissão de fluorescência da proteína re-enovelada com a da proteína desnaturada na presença de ureia, se obtêm dois espectros de emissão bastante diferentes. Adicionalmente, a proteína re-enovelada, excitada com um comprimento de onda de 295 nm atinge seu máximo de emissão em 335 nm, indicando que os triptofanos da proteína se encontram em ambiente hidrofóbico, característico de proteínas enoveladas. (MUNISHKINA; FINK, 2007) Testes experimentais adicionais visando a obtenção de

informação sobre as estruturas secundária e terciária deste domínio serão realizados futuramente.

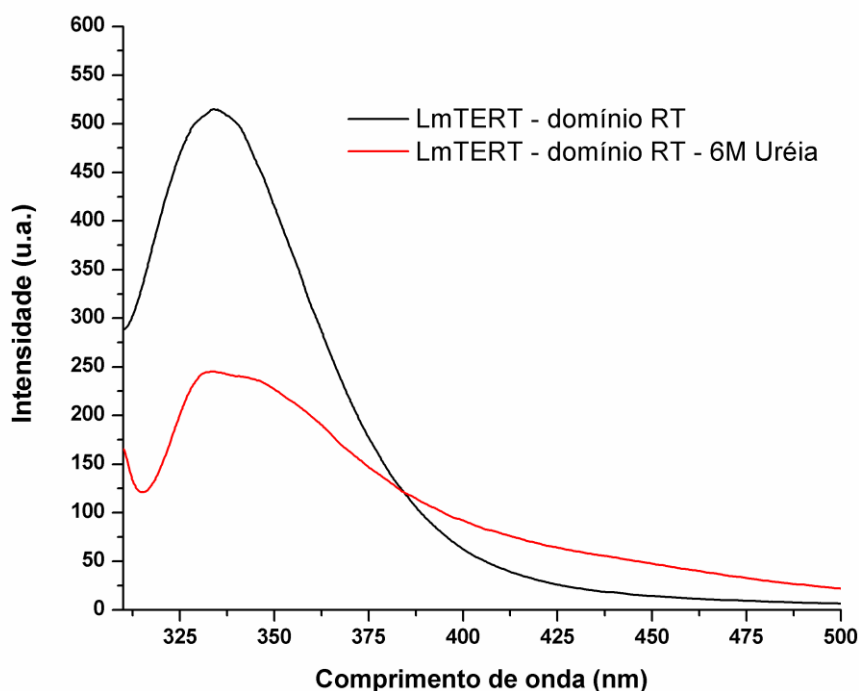


Figura 53. Fluorescência estática do domínio RT de *L. major* re-enovelado

Espectroscopia de fluorescência estática dos triptofanos (comprimento de onda de excitação 295 nm) do domínio RT da telomerase de *L. major* (LmTERT) na ausência (linha preta) e na presença de 6M uréia (linha vermelha). A emissão de fluorescência máxima do domínio RT se dá a 335 nm o que indica que os triptofanos presentes no domínio se encontram num ambiente hidrofóbico o que é característico de proteínas enoveladas.

5. CONCLUSÕES E PERSPECTIVAS

O presente trabalho apresenta o primeiro estudo de caracterização estrutural do componente TERT da telomerase de *Leishmania major*. Com os resultados apresentados, foi possível verificar que o componente TERT do gênero *Leishmania* é o mais basal dentre os TERTs descritos em eucariotos e que quando comparado com os

demais TERTs, a proteína de *Leishmania* apresenta conservação somente entre as espécies deste gênero, contendo substituições características dos TERT destes parasitos.

Usando diferentes ferramentas bioinformáticas foram delimitados no TERT de *L. major* os 4 domínios estruturais encontrados nos TERT canônicos, que permitiu a realização de estudos estruturais individuais de cada domínio.

O domínio TEN desenvolve a função de interagir com o DNA telomérico e com proteínas que se ligam ao DNA simples fita. A conservação da estrutura primária deste domínio em *L. major* é muito baixa, a ponto de que as comparações de sequência de aminoácidos entre os diferentes organismos não mostrarem identidade significativa. No entanto, foi possível se observar que a região carboxi-terminal deste domínio é mais conservada e que os resíduos invariáveis, que constituem o motivo GQ são conservados em *L. major*. A baixa conservação estrutural deste domínio não permitiu gerar uma predição de estrutura terciária, por isso, realizou-se a predição de estrutura secundária. Esta mostrou que o domínio TEN de *L. major* é formado principalmente por elementos desenovelados e hélices- α .

O domínio TRB que interage com regiões conservadas do componente TER e atua com importante destaque na definição do limite do molde, mostrou a presença dos dois seus três motivos, QFP e T. parcialmente conservados em *L. major*, pois eles apresentam substituições próprias do gênero. Com relação ao motivo CP, foi observada uma baixa conservação da sua estrutura primária e terciária, o que pode indicar que ele não está presente em *L. major*. O domínio TRB tem sido o mais estudado estruturalmente e várias estruturas cristalográficas de ciliados e vertebrados estão disponíveis. Estes organismos apresentam uma estrutura do TRB muito similar, com os motivos T e CP conservados e coordenando a interação do TERT com o componente

TBE do TER. A predição de estrutura terciária do domínio TRB de *L. major* mostra uma conservação geral de sua estrutura terciária. Entretanto, as substituições presentes no motivo T encontram-se em posições relevantes na estrutura, tanto para a formação do domínio quanto para a interação do domínio com o componente TER. Já o domínio CP do TRB de *L. major*, mostra uma conservação parcial de estrutura terciária, mas a região de interação com o domínio T e de formação da superfície de interação com o componente TBE do TER não é conservada e está contida em uma região desenovelada ao invés de uma hélice- α . Estes dados reunidos sugerem que o mecanismo de interação do TERT de *L. major* com o componente TER, especificamente a região do TBE, não é conservado. Estes dados são suportados por predições de *docking* molecular feitos entre o modelo tridimensional do domínio TRB de *L. major* e o elemento TBE de *T. thermophila* que predizem que o ácido nucléico interage com a proteína em uma região e com uma orientação diferente do observado em *T. thermophila*.

O domínio RT, por sua vez, forma o centro catalítico da enzima e mostra uma alta conservação com transcriptases canônicas. Este domínio é o mais conservado em *L. major* quando comparado com organismos eucariotos derivados. O RT de *L. major* apresenta todos os motivos de relevância funcional reportados em outros TERT. Estes motivos apresentam conservação tanto de estrutura primária como de estrutura terciária. Além disso, sabe-se que o sítio catalítico dos TERT é formado por três resíduos de aspartato que estabilizam um íon Mg^{2+} . Estes resíduos estão presentes no RT de *L. major* e sua orientação espacial sugere uma boa conservação do sítio ativo.

Em conjunto todos os resultados estruturais aqui mostrados permitem concluir que a maquinaria catalítica do TERT de *L. major* é mais conservada que a maquinaria telomérica que interage com ácidos nucléicos, quando se comparam as mesmas com eucariotos derivados.

Foi possível expressar dois domínios (TEN e RT) em modelo heterólogo e atualmente, os protocolos de purificação estão sendo padronizados. A purificação destes domínios permitirá a realização de experimentos de caracterização estrutural e bioquímica, como por exemplo, a análise da interação destes domínios entre si, com ácidos nucleicos e com outras proteínas, usando diferentes técnicas de biologia molecular estrutural (espectroscopia de fluorescência estática e dicroísmo circular, termoforese em microescala e ensaios de *pull-down*).

Os resultados já apresentados e os que podem ser obtidos no futuro a partir deste trabalho serão de grande importância para se entender o mecanismo de biogênese da telomerase em *Leishmania*, assim como para gerar novas informações acerca da história evolutiva desta família de transcriptase reversas especializadas.

6. BIBLIOGRAFIA

AKHOUNDI, M.; KUHL, K.; CANNET, A.; et al. A Historical Overview of the Classification, Evolution, and Dispersion of Leishmania Parasites and Sandflies. **PLoS neglected tropical diseases**, v. 10, n. 3, p. e0004349, 2016. Public Library of Science.

ALTSCHUL, S. F.; GISH, W.; MILLER, W.; MYERS, E. W.; LIPMAN, D. J. Basic local alignment search tool. **Journal of Molecular Biology**, v. 215, n. 3, p. 403–410, 1990.

ALVAR, J.; VÉLEZ, I. D.; BERN, C.; et al. Leishmaniasis Worldwide and Global Estimates of Its Incidence. (M. Kirk, Ed.) **PLoS ONE**, v. 7, n. 5, p. e35671, 2012.

ARMBRUSTER, B. N.; BANIK, S. S. R.; GUO, C.; SMITH, A. C.; COUNTER, C. M. N-terminal domains of the human telomerase catalytic subunit required for enzyme activity in vivo. **Mol Cell Biol**, v. 21, n. 22, p. 7775–7786, 2001.

AUTEXIER, C.; LUE, N. F. The Structure and Function of Telomerase Reverse Transcriptase. **Annual Review of Biochemistry**, v. 75, n. 1, p. 493–517, 2006.

BAJON, E.; LATERREUR, N.; WELLINGER, R. J. A Single Templating RNA in Yeast Telomerase. **Cell Reports**, v. 12, n. 3, p. 441–448, 2015.

BANIK, S. S. R.; GUO, C.; SMITH, A. C.; et al. C-terminal regions of the human telomerase catalytic subunit essential for in vivo enzyme activity. **Molecular and cellular biology**, v. 22, n. 17, p. 6234–46, 2002. American Society for Microbiology (ASM).

BARRY, J. D.; GINGER, M. L.; BURTON, P.; MCCULLOCH, R. Why are parasite contingency genes often associated with telomeres? **International journal for parasitology**, v. 33, n. 1, p. 29–45, 2003.

BEATTIE, T. L.; ZHOU, W.; ROBINSON, M. O.; HARRINGTON, L. Reconstitution of human telomerase activity in vitro. **Current Biology**, v. 8, n. 3, p. 177–180, 1998.

BERARDINELLI, F.; COLUZZI, E.; SGURA, A.; ANTOCCIA, A. Targeting telomerase and telomeres to enhance ionizing radiation effects in in vitro and in vivo cancer models. **Mutation Research/Reviews in Mutation Research**, v. 773, p. 204–219, 2017.

BERMAN, J. Recent Developments in Leishmaniasis: Epidemiology, Diagnosis, and Treatment. **Current infectious disease reports**, v. 7, n. 1, p. 33–38, 2005.

BLACKBURN, E. H. Telomeres: no end in sight. **Cell**, v. 77, n. 5, p. 621–3, 1994.

BLACKBURN, E. H.; COLLINS, K. Telomerase: An RNP Enzyme Synthesizes DNA. **Cold Spring Harbor Perspectives in Biology**, v. 3, n. 5, p. a003558–a003558, 2011.

BLACKBURN, E. H.; GALL, J. G. A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. **Journal of molecular biology**, v. 120, n. 1, p. 33–53, 1978.

BLASCO, M. A. Telomeres and human disease: ageing, cancer and beyond. **Nature Reviews Genetics**, v. 6, n. 8, p. 611–622, 2005.

BOSOY, D.; PENG, Y.; MIAN, I. S.; LUE, N. F. Conserved N-terminal motifs of telomerase reverse transcriptase required for ribonucleoprotein assembly in vivo. **Journal of Biological Chemistry**, v. 278, n. 6, p. 3882–3890, 2003.

BRYAN, C.; RICE, C.; HOFFMAN, H.; et al. Structural Basis of Telomerase Inhibition by the Highly Specific BIBR1532. **Structure**, v. 23, n. 10, p. 1934–1942, 2015.

BRYAN, T. M.; GOODRICH, K. J.; CECH, T. R. Telomerase RNA bound by protein motifs specific to telomerase reverse transcriptase. **Molecular cell**, v. 6, p. 493–499, 2000.

CANO, M. I. Telomere biology of trypanosomatids: more questions than answers. **Trends in**

parasitology, v. 17, n. 9, p. 425–9, 2001.

CANO, M. I.; DUNGAN, J. M.; AGABIAN, N.; BLACKBURN, E. H. Telomerase in kinetoplastid parasitic protozoa. **Proceedings of the National Academy of Sciences of the United States of America**, v. 96, n. 7, p. 3616–21, 1999.

CASH, D. D.; FEIGON, J. Structure and folding of the Tetrahymena telomerase RNA pseudoknot. **Nucleic Acids Research**, 2017.

CHAN, A.; BOULÉ, J.-B.; ZAKIAN, V. A.; MCELVER, J.; WEBER, S. Two Pathways Recruit Telomerase to *Saccharomyces cerevisiae* Telomeres. (S. Biggins, Ed.) **PLoS Genetics**, v. 4, n. 10, p. e1000236, 2008. Public Library of Science.

CHAN, H.; WANG, Y.; FEIGON, J. Progress in Human and Tetrahymena Telomerase Structure Determination. **Annual Review of Biophysics**, v. 46, n. 1, p. 199–225, 2017.

CHAPPUIS, F.; SUNDAR, S.; HAILU, A.; et al. Visceral leishmaniasis: what are the needs for diagnosis, treatment and control? **Nature Reviews Microbiology**, v. 5, n. 11, p. S7–S16, 2007. Nature Publishing Group.

CHEN, J.-L.; GREIDER, C. W. An emerging consensus for telomerase RNA structure. **Proceedings of the National Academy of Sciences of the United States of America**, v. 101, n. 41, p. 14683–4, 2004. National Academy of Sciences.

CHEN, J. L.; GREIDER, C. W. Template boundary definition in mammalian telomerase. **Genes and Development**, v. 17, n. 22, p. 2747–2752, 2003.

CHRISTEN, M.; HÜNENBERGER, P. H.; BAKOWIES, D.; et al. The GROMOS software for biomolecular simulation: GROMOS05. **Journal of Computational Chemistry**, v. 26, n. 16, p. 1719–1751, 2005.

COHEN, P.; BLACKBURN, E. H. Two types of telomeric chromatin in *Tetrahymena thermophila*. **Journal of Molecular Biology**, v. 280, n. 3, p. 327–344, 1998.

COLLINS, K. The biogenesis and regulation of telomerase holoenzymes. **Nature reviews. Molecular cell biology**, v. 7, n. 7, p. 484–94, 2006.

COLLINS, K.; GANDHI, L. The reverse transcriptase component of the *Tetrahymena* telomerase ribonucleoprotein complex. **Proceedings of the National Academy of Sciences of the United States of America**, v. 95, n. 15, p. 8485–90, 1998.

COLLINS, K.; RICHARDS, R. J.; WU, H.; et al. Structural study of elements of *Tetrahymena* telomerase RNA stem-loop IV domain important for function. **RNA (New York, N.Y.)**, v. 12, n. 8, p. 1475–85, 2006.

CONTE, F. F.; CANO, M. I. N. Genomic organization of telomeric and subtelomeric sequences of *Leishmania (Leishmania) amazonensis*. **International Journal for Parasitology**, v. 35, n. 13, p. 1435–1443, 2005.

CREGG, J. M. Introduction. **Methods in molecular biology (Clifton, N.J.)**. v. 389, p.1–10, 2007.

CROFT, S. L.; SUNDAR, S.; FAIRLAMB, A. H. Drug Resistance in Leishmaniasis. **Clinical Microbiology Reviews**, v. 19, n. 1, p. 111–126, 2006.

DEY, A.; CHAKRABARTI, K. Current perspectives of telomerase structure and function in eukaryotes with emerging views on telomerase in human parasites. **International Journal of Molecular Sciences**, 2018.

DIDWANIA, N.; SHADAB, M.; SABUR, A.; ALI, N. Alternative to chemotherapy-the Unmet demand against leishmaniasis. **Frontiers in Immunology**, v. 21, n. 12, 2017.

DMITRIEV, P. V; PETROV, A. V; DONTSOVA, O. A. Yeast telosome complex: components and their functions. **Biochemistry. Biokhimiia**, v. 68, n. 7, p. 718–34, 2003.

DOMINGUEZ, C.; BOELEN, R.; BONVIN, A. M. J. J. HADDOCK: A Protein–Protein Docking Approach Based on Biochemical or Biophysical Information. **Journal of the American Chemical Society**, v. 125, n. 7, p. 1731–1737, 2003.

DREESEN, O.; CROSS, G. A. M. Telomere length in *Trypanosoma brucei*. **Experimental parasitology**, v. 118, n. 1, p. 103–110, 2008. NIH Public Access.

EDGAR, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic acids research**, v. 32, n. 5, p. 1792–7, 2004.

FERNANDES, C. A. H.; PEREZ, A. M.; BARROS, A. C.; et al. Dual cellular localization of the *Leishmania amazonensis* Rbp38 (LaRbp38) explains its affinity for telomeric and mitochondrial DNA. **In preparation**.

FERNÁNDEZ, M. F.; CASTELLARI, R. R.; CONTE, F. F.; et al. Identification of three proteins that associate in vitro with the *Leishmania (Leishmania) amazonensis* G-rich telomeric strand. **European Journal of Biochemistry**, v. 271, n. 14, p. 3050–3063, 2004.

FIGUEIREDO, L. M.; FREITAS-JUNIOR, L. H.; BOTTIUS, E.; OLIVO-MARIN, J.-C.; SCHERF, A. A central role for *Plasmodium falciparum* subtelomeric regions in spatial positioning and telomere length regulation. **The EMBO Journal**, v. 21, n. 4, p. 815–824, 2002.

FRIEDMAN, K. L.; CECH, T. R. Essential functions of amino-terminal domains in the yeast telomerase catalytic subunit revealed by selection for viable mutants. **Genes and Development**, v. 13, n. 21, p. 2863–2874, 1999.

FU, G.; BARKER, D. C. Rapid cloning of telomere-associated sequence using primer-tagged amplification. **BioTechniques**, v. 24, n. 3, p. 386–90, 1998.

GHORBANI, M.; FARHOUDI, R. Leishmaniasis in humans: Drug or vaccine therapy? **Drug Design, Development and Therapy**, v. 22, n. 12, p. 25–40, 2018.

GIARDINI, M. A.; FERNÁNDEZ, M. F.; LIRA, C. B. B.; CANO, M. I. N. *Leishmania amazonensis*: Partial purification and study of the biochemical properties of the telomerase reverse transcriptase activity from promastigote-stage. **Experimental Parasitology**, v. 127, n. 1, p. 243–248, 2011.

GIARDINI, M. A.; LIRA, C. B. B.; CONTE, F. F.; et al. The putative telomerase reverse transcriptase component of *Leishmania amazonensis*: Gene cloning and characterization. **Parasitology Research**, v. 98, n. 5, p. 447–454, 2006.

GIARDINI, M. A.; SEGATTO, M.; SILVA, M. S. DA; NUNES, V. S.; CANO, M. I. N. Telomere and Telomerase Biology. **Progress in Molecular Biology and Translational Science**. v. 125, p.1–40, 2014.

GILLIS, A. J.; SCHULLER, A. P.; SKORDALAKES, E. Structure of the *Tribolium castaneum* telomerase catalytic subunit TERT. **Nature**, v. 455, n. October, p. 633–638, 2008.

GREIDER, C. W.; BLACKBURN, E. H. Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. **Cell**, v. 43, n. 2 Pt 1, p. 405–13, 1985.

GRIFFITH, J. D.; COMEAU, L.; ROSENFELD, S.; et al. Mammalian telomeres end in a large duplex loop. **Cell**, v. 97, n. 4, p. 503–14, 1999.

GUGLIELMI, B.; WERNER, M. The Yeast Homolog of Human PinX1 Is Involved in rRNA and Small Nucleolar RNA Maturation, Not in Telomere Elongation Inhibition*. **Published JBC Papers in Press**, 2002.

GUPTA, S. K.; KOLET, L.; DONIGER, T.; et al. The *Trypanosoma brucei* telomerase RNA (TER) homologue binds core proteins of the C/D snoRNA family. **FEBS Letters**, v. 587, n. 9, p. 1399–1404, 2013.

HARKISHEIMER, M.; MASON, M.; SHUVAEVA, E.; SKORDALAKES, E. A Motif in the Vertebrate Telomerase N-Terminal Linker of TERT Contributes to RNA Binding and Telomerase Activity and

- Processivity. **Structure**, v. 21, n. 10, p. 1870–1878, 2013.
- HOFFMAN, H.; RICE, C.; SKORDALAKES, E. Structural Analysis Reveals the Deleterious Effects of Telomerase Mutations in Bone Marrow Failure Syndromes. **Journal of Biological Chemistry**, v. 292, n. 11, p. 4593–4601, 2017.
- HOLT, S. E.; AISNER, D. L.; BAUR, J.; et al. Functional requirement of p23 and Hsp90 in telomerase complexes. **Genes and Development**, v. 13, n. 7, p. 817–826, 1999.
- HUANG, J.; BROWN, A. F.; WU, J.; et al. Structural basis for protein-RNA recognition in telomerase. **Nature Structural & Molecular Biology**, v. 21, n. 6, p. 507–512, 2014.
- HUARD, S.; MORIARTY, T. J.; AUTEXIER, C. The C terminus of the human telomerase reverse transcriptase is a determinant of enzyme processivity. **Nucleic acids research**, v. 31, n. 14, p. 4059–70, 2003. Oxford University Press.
- HUELSENBECK, J. P.; RONQUIST, F. MRBAYES: Bayesian inference of phylogenetic trees. **Bioinformatics (Oxford, England)**, v. 17, n. 8, p. 754–5, 2001.
- JACOBS, S. A.; PODELL, E. R.; CECH, T. R. Crystal structure of the essential N-terminal domain of telomerase reverse transcriptase. **Nature Structural & Molecular Biology**, v. 13, n. 3, p. 218–225, 2006.
- JANSSON, L. I.; AKIYAMA, B. M.; OOMS, A.; et al. Structural basis of template-boundary definition in Tetrahymena telomerase. **Nature Structural & Molecular Biology**, v. 22, n. 11, p. 883–8, 2015.
- JIANG, J.; CHAN, H.; CASH, D. D.; et al. Structure of Tetrahymena telomerase reveals previously unknown subunits, functions, and interactions. **Science**, v. 350, n. 6260, p. aab4070-aab4070, 2015.
- JIANG, J.; MIRACCO, E. J.; HONG, K.; et al. The architecture of Tetrahymena telomerase holoenzyme. **Nature**, v. 496, n. 7444, p. 187–192, 2013.
- KELLEY, L. A.; MEZULIS, S.; YATES, C. M.; WASS, M. N.; STERNBERG, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. **Nature Protocols**, v. 10, n. 6, p. 845–858, 2015. Nature Research.
- KIM, N.-K.; ZHANG, Q.; ZHOU, J.; et al. Solution Structure and Dynamics of the Wild-type Pseudoknot of Human Telomerase RNA. **Journal of Molecular Biology**, v. 384, n. 5, p. 1249–1261, 2008.
- LAI, C. K.; MITCHELL, J. R.; COLLINS, K. L. RNA Binding Domain of Telomerase Reverse Transcriptase. **Molecular and cellular biology**, v. 21, n. 4, p. 990–1000, 2001.
- LANGE, T. DE. Opinion: T-loops and the origin of telomeres. **Nature Reviews Molecular Cell Biology**, v. 5, n. 4, p. 323–329, 2004. Nature Publishing Group.
- LANGE, T. DE. Telomere-related Genome Instability in Cancer. **Cold Spring Harbor Symposia on Quantitative Biology**, v. 70, p. 197–204, 2005.
- LANGE, T. DE. How telomeres solve the end-protection problem. **Science (New York, N.Y.)**, v. 326, n. 5955, p. 948–52, 2009.
- LEEPER, T.; LEULLIOT, N.; VARANI, G. The solution structure of an essential stem-loop of human telomerase RNA. **Nucleic acids research**, v. 31, n. 10, p. 2614–21, 2003.
- LENK, E. J.; REDEKOP, W. K.; LUYENDIJK, M.; et al. Socioeconomic benefit to individuals of achieving 2020 targets for four neglected tropical diseases controlled/eliminated by innovative and intensified disease management: Human African trypanosomiasis, leprosy, visceral leishmaniasis, Chagas disease. (C. M. Budke, Ed.) **PLOS Neglected Tropical Diseases**, v. 12, n. 3, p. e0006250, 2018.
- LEWIS, K. A.; WUTTKE, D. S. Telomerase and Telomere-Associated Proteins: Structural Insights into Mechanism and Evolution. **Structure**, v. 20, n. 1, p. 28–39, 2012.
- LICHT, J. D.; COLLINS, K. Telomerase RNA function in recombinant Tetrahymena telomerase. **Genes and Development**, v. 13, n. 9, p. 1116–1125, 1999.

LIN, J.; BLACKBURN, E. H. Nucleolar protein PinX1p regulates telomerase by sequestering its protein catalytic subunit in an inactive complex lacking telomerase RNA. **Genes and Development**, v. 18, n. 4, p. 387–396, 2004.

LIN, J.; LY, H.; HUSSAIN, A.; et al. A universal telomerase RNA core structure includes structured motifs required for binding the telomerase reverse transcriptase protein. **Proceedings of the National Academy of Sciences of the United States of America**, v. 101, n. 41, p. 14713–8, 2004.

LINGNER, J.; CECH, T. R. Purification of telomerase from *Euplotes aediculatus*: requirement of a primer 3' overhang. **Proceedings of the National Academy of Sciences of the United States of America**, v. 93, n. 20, p. 10712–7, 1996.

LINGNER, J.; HUGHES, T. R.; SHEVCHENKO, A.; et al. Reverse transcriptase motifs in the catalytic subunit of telomerase. **Science (New York, N.Y.)**, v. 276, n. 5312, p. 561–7, 1997.

LIRA, C. B. B.; GIARDINI, M. A.; NETO, J. L. S.; CONTE, F. F.; CANO, M. I. N. Telomere biology of trypanosomatids: beginning to answer some questions. **Trends in Parasitology**, v. 23, n. 8, p. 357–362, 2007.

LIRA, C. B. B.; GUI, K. E.; PEREZ, A. M.; et al. DNA and heparin chaperone the refolding of purified recombinant replication protein A subunit 1 from *Leishmania amazonensis*. **Biochimica et Biophysica Acta (BBA) - General Subjects**, v. 1790, n. 2, p. 119–125, 2009.

LIRA, C. B. B.; SIQUEIRA NETO, J. L. DE; KHATER, L.; et al. LaTBP1: A *Leishmania amazonensis* DNA-binding protein that associates in vivo with telomeres and GT-rich DNA using a Myb-like domain. **Archives of Biochemistry and Biophysics**, v. 465, n. 2, p. 399–409, 2007.

LIRA, C. B. B.; SIQUEIRA NETO, J. L.; GIARDINI, M. A.; et al. LaRbp38: A *Leishmania amazonensis* protein that binds nuclear and kinetoplast DNAs. **Biochemical and Biophysical Research Communications**, v. 358, n. 3, p. 854–860, 2007.

LUE, N. F. Adding to the ends: What makes telomerase processive and how important is it? **BioEssays**, v. 26, n. 9, p. 955–962, 2004.

LUE, N. F.; LIN, Y.-C.; MIAN, I. S. A conserved telomerase motif within the catalytic domain of telomerase reverse transcriptase is specifically required for repeat addition processivity. **Molecular and cellular biology**, v. 23, n. 23, p. 8440–9, 2003.

MACNEIL, D.; BENSOUSSAN, H.; AUTEXIER, C. Telomerase Regulation from Beginning to the End. **Genes**, v. 7, n. 9, p. 64, 2016.

MADDISON, W. P. MADDISON, D. R. Mesquite: a modular system for evolutionary analysis. Version 2.72. , 2011.

MALYAVKO, A. N.; PARFENOVA, Y. Y.; ZVEREVA, M. I.; DONTSOVA, O. A. Telomere length regulation in budding yeasts. **FEBS Letters**, v. 588, n. 15, p. 2530–2536, 2014.

MASON, M.; SCHULLER, A.; SKORDALAKES, E. Telomerase structure function. **Current Opinion in Structural Biology**, v. 21, n. 1, p. 92–100, 2011.

MCWILLIAM, H.; LI, W.; ULUDAG, M.; et al. Analysis Tool Web Services from the EMBL-EBI. **Nucleic acids research**, v. 41, n. Web Server issue, p. W597-600, 2013.

MERGNY, J.-L.; RIOU, J.-F.; MAILLIET, P.; TEULADE-FICHO, M.-P.; GILSON, E. Natural and pharmacological regulation of telomerase. **Nucleic acids research**, v. 30, n. 4, p. 839–65, 2002. Oxford University Press.

MITCHELL, M.; GILLIS, A.; FUTAHASHI, M.; FUJIWARA, H.; SKORDALAKES, E. Structural basis for telomerase catalytic subunit TERT binding to RNA template and telomeric DNA. **Nature structural & molecular biology**, v. 17, n. 4, p. 513–518, 2010.

MOREA, E. G. O.; VIVIESCAS, M. A.; FERNANDES, C. A. H.; et al. A calmodulin-like protein

(LCALA) is a new *Leishmania amazonensis* candidate for telomere end-binding protein. **Biochimica et Biophysica Acta (BBA) - General Subjects**, v. 1861, n. 11, p. 2583–2597, 2017.

MORIARTY, T. J.; MARIE-EGYPTIENNE, D. T.; AUTEXIER, C. Functional Organization of Repeat Addition Processivity and DNA Synthesis Determinants in the Human Telomerase Multimer. **Molecular and Cellular Biology**, v. 24, n. 9, p. 3720–3733, 2004.

MUNISHKINA, L. A.; FINK, A. L. Fluorescence as a method to reveal structures and membrane-interactions of amyloidogenic proteins. **Biochimica et Biophysica Acta (BBA) - Biomembranes**, v. 1768, n. 8, p. 1862–1885, 2007.

MUSGROVE, C.; JANSSON, L. I.; STONE, M. D. New perspectives on telomerase RNA structure and function. **Wiley Interdisciplinary Reviews: RNA**, v. 9, n. 2, p. e1456, 2018.

NGUYEN, T. H. D.; TAM, J.; WU, R. A.; et al. Cryo-EM structure of substrate-bound human telomerase holoenzyme. **Nature**, v. 557, n. 7704, p. 190–195, 2018.

PATRA, A. K.; MUKHOPADHYAY, R.; MUKHIJA, R.; et al. Optimization of Inclusion Body Solubilization and Renaturation of Recombinant Human Growth Hormone from *Escherichia coli*. **Protein Expression and Purification**, v. 18, n. 2, p. 182–192, 2000.

PAVANI, R. S.; FERNANDES, C.; PEREZ, A. M.; et al. RPA-1 from *Leishmania amazonensis* (LaRPA-1) structurally differs from other eukaryote RPA-1 and interacts with telomeric DNA via its N-terminal OB-fold domain. **FEBS Letters**, v. 588, n. 24, p. 4740–4748, 2014.

PENG, Y.; MIAN, I. S.; LUE, N. F. Analysis of telomerase processivity: mechanistic similarity to HIV-1 reverse transcriptase and role in telomere maintenance. **Molecular cell**, v. 7, n. 6, p. 1201–11, 2001.

PHATAK, P.; BURGER, A. M. Telomerase and its potential for therapeutic intervention. **British Journal of Pharmacology**, v. 152, n. 7, p. 1003–1011, 2009.

POGACIĆ, V.; DRAGON, F.; FILIPOWICZ, W. Human H/ACA small nucleolar RNPs and telomerase share evolutionarily conserved proteins NHP2 and NOP10. **Molecular and cellular biology**, v. 20, n. 23, p. 9028–40, 2000.

POLSHAKOV, V. I.; PETROVA, O. A.; PARFENOVA, Y. Y.; et al. NMR assignments of the N-terminal domain of *Ogataea polymorpha* telomerase reverse transcriptase. **Biomolecular NMR Assignments**, v. 10, n. 1, p. 183–187, 2016.

RICE, C.; SKORDALAKES, E. Structure and function of the telomeric CST complex. **Computational and structural biotechnology journal**, v. 14, p. 161–7, 2016. Research Network of Computational and Structural Biotechnology.

RICHARDS, R. J.; THEIMER, C. A.; FINGER, L. D.; FEIGON, J. Structure of the *Tetrahymena thermophila* telomerase RNA helix II template boundary element. **Nucleic Acids Research**, v. 34, n. 3, p. 816–825, 2006. Oxford University Press.

ROBART, A. R.; COLLINS, K. Human Telomerase Domain Interactions Capture DNA for TEN Domain-Dependent Processive Elongation. **Molecular Cell**, v. 42, n. 3, p. 308–318, 2011.

ROUDA, S.; SKORDALAKES, E. Structure of the RNA-Binding Domain of Telomerase: Implications for RNA Recognition and Binding. **Structure**, v. 15, n. 11, p. 1403–1412, 2007.

SANDHU, R.; SANFORD, S.; BASU, S.; et al. A trans-spliced telomerase RNA dictates telomere synthesis in *Trypanosoma brucei*. **Cell Research**, v. 23, n. 4, p. 537–551, 2013.

SANDIN, S.; RHODES, D. Telomerase structure. **Current Opinion in Structural Biology**, v. 25, p. 104–110, 2014. Elsevier Ltd.

SARAFIANOS, S. G.; DAS, K.; TANTILLO, C.; et al. Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. **The EMBO journal**, v. 20, n. 6, p. 1449–61, 2001. European Molecular Biology Organization.

- SAUERWALD, A.; SANDIN, S.; CRISTOFARI, G.; et al. Structure of active dimeric human telomerase. **Nature Structural & Molecular Biology**, v. 20, n. 4, p. 454–460, 2013.
- SCHMIDT, J. C.; CECH, T. R. Human telomerase: biogenesis, trafficking, recruitment, and activation. **Genes & development**, v. 29, n. 11, p. 1095–105, 2015. Cold Spring Harbor Laboratory Press.
- SCHMUTZ, I.; LANGE, T. DE. Shelterin. **Current Biology**, v. 26, p. 387–407, 2016.
- SEKARAN, V.; SOARES, J.; JARSTFER, M. B. Telomere Maintenance as a Target for Drug Discovery. **Journal of Medicinal Chemistry**, v. 57, n. 3, p. 521–538, 2014.
- SEN, T. Z.; JERNIGAN, R. L.; GARNIER, J.; KLOCZKOWSKI, A. GOR V server for protein secondary structure prediction. **Bioinformatics (Oxford, England)**, v. 21, n. 11, p. 2787–8, 2005. NIH Public Access.
- SILVA, M. S. DA; PAVANI, R. S.; DAMASCENO, J. D.; et al. Nuclear DNA Replication in Trypanosomatids: There Are No Easy Methods for Solving Difficult Problems. **Trends in Parasitology**, v. 33, n. 11, p. 858–874, 2017.
- SILVA, M. S. DA; PEREZ, A. M.; SILVEIRA, R. D. C. V DA; et al. The Leishmania amazonensis TRF (TTAGGG repeat-binding factor) homologue binds and co-localizes with telomeres. **BMC microbiology**, v. 10, p. 136, 2010.
- SILVA, M. S. DA; SILVEIRA, R. C. V. DA; PEREZ, A. M.; et al. Nuclear DNA replication in trypanosomatid protozoa. In: L. L.P. (Ed.); **DNA Replication and Mutation**. p.123–177, 2013. New York: Nova Science Publishers INC.
- SILVEIRA, R. D. C. V. DA; SILVA, M. S. DA; NUNES, V. S.; PEREZ, A. M.; CANO, M. I. N. The natural absence of RPA1N domain did not impair Leishmania amazonensis RPA-1 participation in DNA damage response and telomere protection. **Parasitology**, v. 140, n. 04, p. 547–559, 2013.
- SIMPSON, A.; GILL, E. E.; CALLAHAN, H. A.; LITAKER, R. W.; ROGER, A. J. Early Evolution within Kinetoplastids (Euglenozoa), and the Late Emergence of Trypanosomatids. **Protist**, v. 155, n. 4, p. 407–422, 2004.
- SINGH, A.; UPADHYAY, V.; UPADHYAY, A. K.; SINGH, S. M.; PANDA, A. K. Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process. **Microbial cell factories**, v. 14, p. 41, 2015. BioMed Central.
- SINGH, S. M.; PANDA, A. K. Solubilization and refolding of bacterial inclusion body proteins. **Journal of Bioscience and Bioengineering**, v. 99, n. 4, p. 303–310, 2005.
- SIQUEIRA-NETO, J. L.; LIRA, C. B. B.; GIARDINI, M. A.; et al. Leishmania replication protein A-1 binds in vivo single-stranded telomeric DNA. **Biochemical and biophysical research communications**, v. 358, n. 2, p. 417–23, 2007.
- SODING, J.; BIEGERT, A.; LUPAS, A. N. The HHpred interactive server for protein homology detection and structure prediction. **Nucleic Acids Research**, v. 33, n. Web Server, p. W244–W248, 2005.
- SPOEL, D. VAN DER; LINDAHL, E.; HESS, B.; et al. GROMACS: Fast, flexible, and free. **Journal of Computational Chemistry**, v. 26, n. 16, p. 1701–1718, 2005.
- SREERAMA, N.; WOODY, R. W. Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set. **Analytical Biochemistry**, v. 287, n. 2, p. 252–260, 2000.
- TAGGART, A. K. P.; TENG, S.-C.; ZAKIAN, V. A. Est1p As a Cell Cycle-Regulated Activator of Telomere-Bound Telomerase. **Science**, v. 297, n. 5583, p. 1023–1026, 2002.
- THEIMER, C. A.; BLOIS, C. A.; FEIGON, J. Structure of the Human Telomerase RNA Pseudoknot Reveals Conserved Tertiary Interactions Essential for Function. **Molecular Cell**, v. 17, n. 5, p. 671–682, 2005.

- VALLEJO, L. F.; RINAS, U. Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. **Microbial cell factories**, v. 3, n. 11, 2004.
- VASCONCELOS, E. J. R.; NUNES, V. S.; SILVA, M. S. DA; et al. The putative Leishmania telomerase RNA (LeishTER) undergoes trans-splicing and contains a conserved template sequence. **PloS one**, v. 9, n. 11, p. e112061, 2014.
- VIJAYAKUMAR, S.; DAS, P. Recent progress in drug targets and inhibitors towards combating leishmaniasis. **Acta Tropica**, 2018.
- VIVIESCAS, M. A.; SEGATTO, M.; CANO, M. I. N. Chaperones and Their Role in Telomerase Ribonucleoprotein Biogenesis and Telomere Maintenance. **Current Proteomics**, v. In press, 2018.
- WALLACE, A. C.; LASKOWSKI, R. A.; THORNTON, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. **Protein engineering**, v. 8, n. 2, p. 127–34, 1995.
- WANG, Y.; FEIGON, J. Structural biology of telomerase and its interaction at telomeres. **Current opinion in structural biology**, v. 47, p. 77–87, 2017.
- WATSON, J. D. Origin of concatemeric T7 DNA. **Nature: New biology**, v. 239, n. 94, p. 197–201, 1972.
- WEINRICH, S. L.; PRUZAN, R.; MA, L.; et al. Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT. **Nature Genetics**, v. 17, n. 4, p. 498–502, 1997.
- WHITMORE, L.; WALLACE, B. A. Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. **Biopolymers**, v. 89, n. 5, p. 392–400, 2008.
- WOJTYLA, A.; GLADYCH, M.; RUBIS, B. Human telomerase activity regulation. **Molecular Biology Reports**, v. 38, n. 5, p. 3339–3349, 2011.
- WORLD HEALTH ORGANIZATION. Research priorities for Chagas disease, human African trypanosomiasis and leishmaniasis. **World Health Organization technical report series**, , n. 975, p. v–xii, 1-100, 2012.
- WU, R. A.; DAGDAS, Y. S.; YILMAZ, S. T.; YILDIZ, A.; COLLINS, K. Single-molecule imaging of telomerase reverse transcriptase in human telomerase holoenzyme and minimal RNP complexes. **eLife**, v. 4, 2015.
- WU, R. A.; UPTON, H. E.; VOGAN, J. M.; COLLINS, K. Telomerase Mechanism of Telomere Synthesis. **Annu. Rev. Biochem**, v. 86, p. 439–60, 2017.
- WYATT, H. D. M.; WEST, S. C.; BEATTIE, T. L. InTERTpreting telomerase structure and function. **Nucleic Acids Research**, v. 38, n. 17, p. 5609–5622, 2010.
- XIA, J.; PENG, Y.; MIAN, I. S.; LUE, N. F. Identification of functionally important domains in the N-terminal region of telomerase reverse transcriptase. **Molecular and cellular biology**, v. 20, n. 14, p. 5196–207, 2000.
- XIE, M.; PODLEVSKY, J. D.; QI, X.; BLEY, C. J.; CHEN, J. J.-L. A novel motif in telomerase reverse transcriptase regulates telomere repeat addition rate and processivity. **Nucleic acids research**, v. 38, n. 6, p. 1982–96, 2010. Oxford University Press.
- ZHANG, H. Molecular signaling and genetic pathways of senescence: Its role in tumorigenesis and aging. **Journal of Cellular Physiology**, v. 210, n. 3, p. 567–574, 2007.

ANEXOS

Anexo 1: Fragmentos clonados nos vetores de expressão

Sequência de aminoácidos da telomerase de *L. major*. Destacados em cores estão os três domínios clonados nos vetores de expressão.

MSASFPSIPGFAGPLSLKAFLEEFGLHLTFVETASPSPRAAATAATPSAAEFRALRDVVLPPNQSFLLVVYVALHASSPPPTTAH
ASPTPPTPALGRAASATGFERLRQPLTHQTVASSAHDTCMTRCNPADRQNSWSSSSSFAGKHGSTRDPWRASAFRLLYTNTSHRPLTD
ALLRHPWASFAACLGPAAMGFIEMYCPIVLQLEAMAGGVQLGPALKHTAFERSFSAAAPSEMKGPLPQSGCISSRGVKRALDTR
ECSAPLPQKQRRVEAPPKAKAGRRLQRGVAPDSRPCRDDPFSSSVNRAAMAATWAAALTRTDVPRTRLYSARISDGDGAGDGSAL
PLPAGFLEAQWLRRHPSLHRALQAALPKRAAYGASTRRYSVGASERGTGCTSVEDI PMWHVTHVFRWLVLQPPQSSADASPTTPPKF
DLPShLRLRLSTVVDQCSRLDLRGAALKHTGYLEAFRRQQQGVPEWVDVQRLSTPVDVVVSYLRTLSTLRWAPLKETDNGSFWGRDA
AGSERVLDALMRAVRGWLIIAGRHAVFPVSRFLDGVVPAQVPLNGFYTTAPSLPFPATSSAARHARRERSHVQQRVWLQFALFLTQD
ILPFLLRASFTITWSKNTHKLLFFPAVWVRLVREVRRTSCGAPRSQMPLAGESARRIGAEPALSSAPSACGGVTAALLTPHS
AVSKASAAIAAPRDVWRAVRTGGALTHR GARATLATRGGGASCLYAGVFRFPDRRKL RPIAVVRSASLSL KEMARGSPSPYSHASAI
ARLLRRLGCSDTDGQLPAATATLLRRVQARSRHNRAGVHRFPHPLPHLPHKAALRDALRCLVSGVEEQVRDGLPRLSNLSHQDEY
AELRSFCEEVRGRHALLCEGPAGTAPAVSSAPPPGSAACCFAPYVTLVRS DASRCYDNLPOERVLA AVASLVKHDA YRVLRF TAIHAL
DGEACRGGCLLRRFTTTRTIPCAEAE CGVLARI PRGHI YWEEEEE GHTQGG LHTTA AVSR TTDANRR CGANLISGA AVRALLSEHIR
HHLVVVSGGSLFEQRVGI LQGS PVAMLLCDRLFS DVVDTALSGILSEHAERS LLLRRVDDVLVATTS PAAAERCLRAMQCGWPSVGYV
SNPSKLTLS TACGR LVPWCGLLLHDTTLEVSVEWRRIGVLLASLRVGDPHYVHRGDHDPLYLTQRFLAVLQLRVAPTALCGRMNSKTR
QLQTFYEVGLLWSRVVLEK VQEALPVARNRCVAVLLLRPLAVCVGR LCRLLSGHQRF LAARQSACDVSAAEVRACVLTALHRTMQAKL
RVLQARTVCVVTAQRGRTPRGTS LKGRNHFC SHEALSAGRKCCERRRGRNRRRNTRVCLRSFWWLAAA EVESQWRRSLSALHRAAP
RLGEARGPTALSPSPGASASLLMDDGPLSMHARALSATRLSQT

Domínio TEN

Domínio TRB

Domínio RT

Anexo 2: Sequências utilizadas para se fazer o cladograma

>AAO67516.1 telomerase reverse transcriptase [*Leishmania amazonensis*]

MSASFPSIPGFAGPLSLKAFLEEYFGLHLTFAAETASPSPRAAATAETPSAAGFRALRDVVLPPNQSFVLLVVALHASSSPPTTAAH
ASPTPRNTGSRVALRPPRGSAGFRQAAPWRIRRRYRQAHDTCTVTRCNPSDRQNPFCLSSTGKNGNARSFPCASAWLTYTNTSHRPF
TDALLRHPWRASFSACLGPAAAMGFIEMYCPIVLQLEAMAGGVQVIGPALKHVALETSSFSATAQLSEMKGPPRGSASVSSRGVAKRAVD
TGECSAPLPLQKQRREAVASP PKAARRLQREVA PSHRRCRDDSFSSPVNRAAMPATWAAALTRTDDPRTRLYSVRVSDDCTGDGGG
ALPLPAGSLEAHLVLRHRSRSHRVLQAALPKRRAAYGASTRHYCAGTGERGTGCTSVKNISMWHVAHVFRWLVLQPSQSSEVAFQTTSF
KFDPSPYLRRFLSTDVDQCSRLDLRGAALRHTGYLEEAFRRQQQVGEVPDQVQLRSTPVDVVVSYLRTLPTLRWAPLKEANNGLFWGR
DAAGSERVLDALMRAVRGWLIAQRQAVVPSRFLDGVVPAQVPWLNGFYTTTSPFLPFS DATSSASRHARRERSQVQRVWVLFVLTFL
QDILPFLLRASFTITWSSKNTHLFFPAFVWRRLVREVRRTSRCHAPRSQMSLAEERTRMSGADHAALVSA PNACCGVAAAFFSP
HSAASNASAAI SAPRDEWRAVRTGGALAHWSARATLATRGGGASCYAGVFRPDRKRLRPIAVVRSASLRLSKEMARGSPSPYSHAS
AIVRLRLRLGCSADADQLPAATATLLRQVQARSRHNRRTGVHRSAHPLPHPLPKAALQDALRCLVSGVEEQVRVLDGLPRLNLNSHQD
EYAE LRSFC EEVGRHAIPCEGPAKTT PAVSSASPPGATACFAPYVTLRLSDASRCYDNLPPQGRVLA AVRSVLVKHDAYRVLRFVTHA
VDS EATCKGGLLRRTFTTTRIPCAEAECGFLARIPRGHIYWE EEEEGHTQGGLHTTAAVSRRTDANRRCGANLISGA AVRALLS EHIRH
HLVSVSGS LFEQRVGLIQGSPVAMLLCDRLFSNVVD TALSSILSEHAERSLLLRRVDDVLTATTSPAAAE RCLRAMQRGWPSVGYLS
NPSKLTLSKACGSLV PWCGLLLHDTTLEVSVWRRIGVLLGSLRVGDPHYVHRGDYEP LYLTRQFLAVLQLRVAPTALCGRMNSKTRQ
LQTFYRVGLLWRVVLKQVEALPVARNCVTVLLRRLPLAVCVGRLCRLLTRHQRF LAARQ S ADCVSAAEVRACVLTALHRTVQA KLR
VLQARTVRAMTVQSGRTQRGLSGLKGRNNFCSTKASSVGRKCEQRRRGRNRNRTRVCLRSFWWLTA AEVESQWRRSLGALYRAAPR
GEACGSTASSPSPAASLMLMEDGPLSMHARALSATRLSQT

>AAO67514.1 telomerase reverse transcriptase, partial [*Leishmania major*]

MSASFPSIPGFAGPLSLKAFLEEYFGLHLTFAVETASPSPRAAATAETPSA AEFRALRDVVLPPNQSFVLLVVALHASSSPPTTAAH
ASPTPPTPALGRAASATGFERLRQPLTHQTVASSAHDTMTRCNPADRQNSSWSSSSS FAGKHGSTRDPWRASAFRLTYTNTSHRPLTD
ALLRHPWWASFAACLGPAAMGFIEMYCPIVLQLEAMAGGVQVLPALKHTAFERSFSAAPSSSEMKG PLPQSGS GCISSRGVKRALDTR
ECSAPLPLQKQRREAPPKAKAGRRLQGVAPDSRPCRDDP FSSSVNRAAMAATWAAALTRTDVPRTRLYSARISDGDGAGDGSAL
PLPAGFLEAQLRRHPHSLHRALQAALPKRAAYGASTRRY SVGASERGTGCTSVEDI PMWVHTVFRWLVLQPPQS ADASPTTPPKF
DLPShLRRLLSTVVDQCSRLDLRGAALKHTGYLEEAFRRQQQVGEVPDQVQLRSTPVDVVVSYLRTLSTLRWAPLKEANNGLFWGRD
AGSERVLDALMRAVRGWLIAGRHAVFVPSRFLDGVVPAQVPWLNGFYTTAPSLPFP AATSSAARHARRERSHVQQRVWVLFALFLTQD
ILPFLLRASFTITWSSKNTHLFFPAFVWRRLVREVRRTSRSCGAPRSQMP LAGESARRIGAEPALSSAPSACGGVVTALLT PHS
AVKASAAA IAPRDVWRVAVRTGGALTRHGARATLATRGGGASCYAGVFRPDRKRLRPIAVVRSASLRLSKEMARGSPSPYSHASAI
ARLLRRLGCSDTDQG LPAATATLLRRLVQARSRHNRAGVHRFPHPHPLPHKAAALRDALRCLVSGVEEQVRVLDGLPRLNLNSHQDEY
AELRSFC EEVGRHALLCEGPAGTAPAVSS APPGSAACCFAPYVTLVRS DASRCYDNLPPQERVLA AVASLVKHDAYRVLRF TA IHAL
DGEACRGGCLLRRFTTTRIPCAEAECGFLARI PRGHI YWEEEEE GHTQGLHTTAAVSRRTDANRRCGANLISGA AVRALLS EHIRH
HNLVSVSGS LFEQRVGLIQGSPVAMLLCDRLFSNVVD TALSSILSEHAERSLLLRRVDDVLTATTSPAAAE RCLRAMQRGWPSVGYV
SNPSKLTLS TACGRVLPWCGLLLHDTTLEVSVWRRIGVLLASLRVGD PHYVHRGDHDP LYLTRQFLAVLQLRVAPTALCGRMNSKTR
QLQTFYEVGLLWSRVVLEKQVEALPVARNCVAVLLRRLPLAVCVGRLCRLLSGHQRF LAARQS ADCVSAAEVRACVLTALHRTMQA KLR
RVLQARTVVCVTAQRGRTPRGTSLKGRNNHFC SHEALSAGRKCCE RRKRGRNRNRTRVCLRSFWWLAAA E VESQWRRSLSALHRAAP
RLGEARPTALSPSPGASASLLMDDGPLSMHARALSATRLSQT

>XP 819198.1 telomerase reverse transcriptase [*Trypanosoma cruzi* strain CL Brener]

MSSTFPIVPGYDGPWSLKDFLQRYFFVRLRFEAS PPTCFVPPNERNLVFRPTGGFVAVMYASLSAPPPPPP PVLVVTTGEGEFAS
KGHLLPRATSLT PHGRPLTSSLLRFPFWNSLFTQLGPAAVSFI LAWCPVVVQFEPNGGGQLQVLGPPLKNAVPRVAVNRQRGPGWTSGR
KMGYERNAQHRI LQGPSPELSISSLYRLNVSRLPLIGEAPLCTFAE IDRVRNPRLSKLWQEHHGGMGAA P SSSC RCS SVTAGAVPMPCA
DACSVLAVSRELVH LVPFPTLS INV DGDNTVMKASVERHRLTHVLHCVLSGVQMNFVREGAAIKHTGFLEERSLSFNHVRPRGEAGACAA
FPVSELTVPETIMT SYLHSLG IWWRAPQATERVNFWGEALVLERLCAITLSW LQCGRHENFPLSHFLH QIPVAKL PWLRGFYTRGS
GRRRSMIQQRVLLQ FFFFVQHVVPFLVRSSFHVTWSSKKNAPFLFIPKSVWIRLVSRELRQVCLRRRRTRKRERDGDGAPQPMALP
PMALERLSDNIVGLKI PPGKDAVPAGTAA P LLYSSIRFLDQRKLRPIARVRFANARWLK MADGLCS PATPVSCNFSGRAVASTT
QTKR SAVSNATVLRGALRCLLAGLAEQVSGSSLVKFTNI SHQDEYAEVRSFVEGARNFHHPLASTPVPSAA PVSSAVVMTVRG
DAARCYDCLPQEA VMNTVRRLLPHEFYHTLSFTAVAPLAGGRGTANVEEVGGQGWKRAEEDAPSVGVLLHRRVKLQTVSGKCVQNGTL
PGIPKGWIMYEEPTP PSESVMRGEEMRTVLGQHLQKHLVLEGRYSQGVITQGS AVAMLLCDI LLETVDHSLSNL S QHEE PALL
RRVDDVLTVPVAAAARCDVALRHGWPEVGFCCQDEKL RCTVALIPWCGLLWNPQTLEFSVEWRRLLSLLPHIASRPMTGNEP LHC
SLRLMSILCLRTPLTRVIVLCPRTLVVQTLCEI GFLWSRFLPHLKLQNAMFIRPHVRLRPLALAIASLRLLIRRQKYGLNERLGSF
CDVTDEIRL CVLFTLHHTLQEWLPRMISQIRRRREGLKQFLLI SALVRKRMRAALLMDGASC EGDDEHIAASRGSSHGVA AAVVN
EDDDVGKLLAEGDDTVIVRALAAVRFPPSRAFTHEGNLWRLEDPTGGRLKKNK KLLPVLK

>XP 829083.1 telomerase reverse transcriptase [*Trypanosoma brucei* TREU927]

MAQMFSPVPGYDGPWSLKDFLHNYIGLQLRWSTPQENSTLTPPNSNSNVVVVPTGNFHVVVVYSRSAPPPQSVRSVAVLS SCSNKMD
GDHPQPSVPTHFHGRPLTSSVLKHPFDALFSQIGPTAVSFIVLWAPIVVQFEASGGQLVLPPLKRSASCAVKRQPAGAGWEGK LK
VSRTELTQNDPHFCIEGTS SPVALYRLNVPRLLHVEQSP LRSTADARHAVTLILEQLWIEHHKSSCSPS SDRDI SQRLCGVATGSA
TNCCGDSCVVLQAQRNQNVA VFPYTL PVDHDAVSGNGT VKEYLIHVLHLCALGSVCRMNIRGAAIKHTEFLEKRGKSVTSISQKAQTA
GQENKMAIPV GELTVP E P VASYLETLFNMMWWRVPRFVGGAVTADRKFWGEEGTVLTRLRETVD WLVCCGRHEVFLSHFLDGV
PVSRI PWLRGFYTK EPARKRSMIQQRVFLQVLFVFLYQCVF PFLIRRSFHVACTSKSPYIFFFI PRVAVWRLTRREMRRVCIRRVKRM
RSDTSLSDPQMRHAGLP PMALE RVTSEGLMQ LIGPSSQPVGDDSGGVSDVCKLRDSSS PLYSDVRFVLDGSKLRPIARPRFGWQRS
LKAADGVSSTLASALAKAGKYLATD GKAMWPSPNSAVLRDALRCLDAGVEERRITTGARQRAIGSHNDEYIEVRSFIQCARQFC T I P
TLVQGETSAALLADATAVTMRVGDAA RCTDHLPQEA VVNIMRGLVTHEGYEYS LKLT VTTVTKVNSKKTSSHAESQLSSMETNGS
TVKHRKLQLSKRKVLVPGRVYQEGVLYGIARGCIA YEEQSVGSDSLVSGEEIRS VLQKHLQSHLVMLNGMYIORMGINQGS AVAMLL
CDQLLERVDAALSSILSEHEEPALLRRVDDVLT VLTLSRAASRCEDALRSGWSEIGFFCQEEKLRVTCGQPVRVWCGLLWDPVTLEF
AVDWARLAKMMPYLAVRPR TGCEP LLSLRFIRILRLRTPM TALCRQ INSKSRVVQTYEIGLLWSRFFL DKLKANA A FF RHV R T I L
QPLALATALRRLVRKHSF DLQRLGSFCVDTV E VRLCIAALYLR LQHRALF M IGRMSVSGKKKFLI LMTAVLKRKMGALYGRINRN
NQMESSLPSAADLLLVEGGDGVVSRLAAVRFRP VETVRFQGCATRSA

>AAK60396.1 telomerase reverse transcriptase [*Cryptosporidium parvum*]

MSVLK GALAVLMKQEVTLTGEYLEQKRKIRVFKEEGLYNDSSGELRRLEETLILEDP S I P KALKEGFLGLRELTFNSL D N S L D I N G D
RRKLLGINIEKLLLEWSSILGQNYFSSPRNQSSMS S GKR S RNTEYNETLSP SFITDQIKNI LPDSEKDSLTYQVQDFELLRRKHFPRI R
IFHS GYSYGLRTSFDTIKPKKYSFCNSTLYKFQNLHAWNTLLSKIGPIEILFLFVCCIFRILGDHSEILIQAGRM L TNDFFLEELAR
LYETGPKKTNSFVSSSLTPILTIEEKEEVKPLDETFRQVNRSGGI IKDSRSLSKIYKIDIPYRSTFLYCDHFSKRGGPLCLSVLRL

RWIFEDLVVSLIRCFYVTEQOKSYSKTYRKNIWDVIMKMSIADLKKETLAEVQEKEVEEWKSLGFAPGKLRLLIPKKTFRPIMT
FNKKIIVNSDRKTKTLTKLLNSHMLMLTKLNRMFKDPFGFAVFNYDDVMKKYEEFVCKWKQVGPQKLFATMDIEKCYDSVNREKL
STFLKTKLLSSDWIMTQILKRRNNIVLDSKNFRKKEKMYFRQKFKIALEGGQYPTLFSVLENEQNDLNAKKTLLIVEAKQRNYF
KKNLNLQPVINICQYNYINFNGKFKYKQTKGIPQGLCVSSILSSFYATLEESSLGLFRDESMNPNPNVLLMRLTDDYLLITQENN
AVLFLIEKLNVSRENGFKFNMKKLTQTSFPLSPSKFAKYGMDVVEQNIVQDYCDWIGISIDMKTLALMPNINLRIEGILCTNLNMQT
KKASMWLVKKLKSFLMNNITHYFRKTTITTEDFANKTLNKLFISSGKYMQCAKEYKDHFKKLNAMSSMIDLEVSKIIVSIVTRAFFKYL
VCNFKDTIFGEEHYPDFFLSTLKHFIIEFSTKTYIFNRVCMILKAKEAKLKSQCCQLIQYDA

>AAP42278.1 telomerase reverse transcriptase [*Moneuplotes crassus*]
MNPSPFAGKDNNGSNEIHSALKTSEEIPWCKSLDLFLKEMAENNHLEEKWNNFLKSDKIFNKFASNIILAFDNNYVQELKDISFREKR
RSSEQLDELILKCCIQLSLESSKSDSRKKLQCFGYQLTRHRGETTMKALNLSTQKQYFLRKEWADVRLMIGDELFCHYKNYLIFERTS
EGSLVQFAGKNIFEFLGVSQYQAAADKSNLKVSKCKYKINKNKDDHYINNINDVMWNNMKSRTTRIFCYQYRNRTQFFKHLMLMKDHN
TLSAVDRAEKLFQDIFRFRNIRKDLKTNVNILQMVVEKIKKFRFSYLLSKSCPLPENWVERKQVILSHLNDDRQAKAKYEEELFSYT
IENKCVTQFLNEFFYHTLPKDFMTGKNRKSFSQSKIKKYVGLNKHელიHKNLLLGGKLDIKSIKWLGFKTSKKNYHYFDKENRFLWRIL
RWIFEDVVVSLIRCFYVTEQOKSYSETYYRKNIWDIVMKYSIADLNRRETLLEEVPKEDMVKWKNEFLRFAPGKLRLLIPKKTFRPIMT
FNKKILDPEKGTSMKTNTKLSYSHLMLTKLNRMFKDPFGFAVFNYDDVMRKYEEFVLKWRQVGRPKLYFVMDIEKCYDSVDREKL
SQLLGTTRLLSSEFRIMTVQAMKRNAEVVDLKDCAKNMKCECFRQKWKQIALEGGQYPSLLNVLEDDQNDLNAKKTLLVENQKRDPY
KKKALLDPVIAICRHHYIEFNRYKYKQTKGIPQGLCVSSILSSFYASLEENALGYLRKESMDANDPNTLLMRLTDDYLLITTKENN
AFLFIEKLNVSQNKFKFNMKKLTQTFPLDPFSLNKYKMASVEDQNALAHYIDWIGISIDMSTLALMPNINLRIEGLLCTNLNMQT
KKASMWLVKKLKSFLMNNITHYFRKTTITNREFSNKTLNKLVIAGAYKMQCCIEYKDHFKNTNTEIHPQLDKIICAIIVSIVTRAFFKYL
VCNFKDSIYDREHSRDFSIATLRHFIEIFSLKKNHFSGVVILVAKEQKMLEEKDNACMTAIEN

>BAB18587.1 telomerase reverse transcriptase [*Paramecium caudatum*]
MLITLQQAIAQLYNVITLQDYLYKYLDVDPKNIFFQNTIICSETQNLQKFPKIKPLNLFYQKLSDDQRRDEIVTDSRNDFFQKQKQKSF
GCQKMLMSDNMEHLYREIGEQNFYLLKKNYKILIKDDNDSETYIQSGDPMKIIPKQKQQAIAEKIQRQQLLYCNCNMSRQVGMFKKNE
ISYAKKQIIKSNYKMQQITPEIIAIAKSTSNILDYFNKQLSPIEISHLRPILESCMMLKQKFNHGHGFVFPVLSNYKQLKQKETE
RHLKGLSNLSDHLNHIHQQTITSLVNRNLEIIPIDFLFQGNLFTLTDISDFITLRFEDQSYLDYIKKMNVPWLNKTYF
NKKQRLLIHKRNRQMIPLMKFMQEIIIPFLHRNFYITERMKDWWKLFYRKEIWHLILQLSLTSLTKNNFEEISVNSIKTPYVQGL
RIVPKPGTFRPIVITNRRKAKTSKISLNRKLVDIKYILRNLRIOPLGYSVFNPEVSRLEEFKKLWIKLQPKTYITMDIQKCYDTI
ILQKLQFIEESKQFSSIIYANKYHIVSRNNRMLKPSFSMCKDLFNILDRCAIPFNKPQLLQKGYIEHQNKQTIIINQGFQNSVTF
SEFLNSIKNICQNNIVQFENRYFRQTLGIPQGLNISGILCSFYLANLEQKLTNKLIGDITLMLRDTDDYCCCLAFSQQSSIKILNNFTQI
EKQYQIRLNDHKTQHNERSDRYFKWIGKIIDIETLTLKPAFYLESDTKFYQINVNLPKTKTYIKSKLSLILNQKFFFNKLN
DKPTMIKVLKIFVHSGLVKLSFGLKRLKYNYGASKLRNQGKSIKLIKIKVIEAAYCVQSSRESNMEIQEYLMRIIQKQLKKT
RNSYLKRNKINYKQI

>AAF26732.1 telomerase reverse transcriptase 1 [*Candida albicans*]
MTVKVNEKKTLLQYVDLNTSNDVPLPSLKEYMETVLVYKSIKRLPAIRPQESFDEFMKELVTRIVMEKSNNVIAVYGYKTSAMESRS
IFTFHSSGNFLTHITSHNWTIFSLGPKKFLELLVNNGFVSKVNGESVQIFGDVNSHRKAVVVSKEYITKFNLYNSYSRDFSRF
EMIRPSIQITLQDILSFGSLNPGRSSKRYRGFKSLRSRIANDKKCRYDILYAKFIGTSCNFANVSNKTEISQVIFVLLVLGKLL
PLDAWGGVSNKIKDRVDFLLGANEKIHMDLFRGIRLKDQKWLGRAHQISSKQDFELRTAFLKGYLWVFEHLNKLILRSFWYI
TETSSIVSSELNYPQYLWLKELYESWVSKYAKNNLVKMPKSKIQRQLPCGKIKLIPKRSSFRVICVPIKRSKLLLNKLELDTLEK
REFERYRKEVLSVPGYIILRLKLSKLRDYESYRAVSHSSSDVAEKISDYRDSLTRFGEIPKLFILKFDMKCEYDRLSQPVLMKLEE
LFENQDNKTSYYVRYQAQLDASHLKKVKTITIDTQYHNLNLSRRHLSNCKSLVDKTKTIALQKGNILEVCRSQIYDVVSGVNDRDK
NLHLYKRRGVQGFSLSSIFCDILYSAMVHDCQFLWKSQDFLVRVLDVDFLLVTPDSNIYDQVHNILSGKILESYGAFVNDKTV
VVNQTTTETSIFDVLGVNTDLSIKRNSGSISLVTNFRFKTLVKYLKTFYQLNLEGLDLSFGVLENVLENNMSGLLRVLVREFK
TKFTSIVKYDTFHCKYKIFLYDISNYTIVKYVETNSDWEAPELNCIKQIIVKEFSSFESYSEIVEVWQTLNI

>NP_013422.2 Est2p [*Saccharomyces cerevisiae* S288C]
MKILFEFIQDKLDIDLQNTSTYKENLKCGHFNGLEILTTCFALPNSRKIALPCLPGDLSHKAVIDHCIIYLLTGELYNNVLTFFGYKI
ARNEDVNNLSLFCNSANVNVLLKGAAWKMFHSLVGYTAVFDLLINYTVIQFNGQFFTQIVGNRCNEPHLPKQWAQRSSSSSATAAQIK
QLTEPVNTKQFLHKLNNSSFFPYSKILPSSSIIKLLDLREAIFFTNLVKIPQRKVRINLTQLKLRHKLNNVLSILNSICPPL
EGTVLDSHLSRKNKFNKFIIVILQKLLPQEMFGSKNKGIKINLNLLSLPLNGLYLPFDSLLKLLRDLKDRWLFISDIWFTKH
NFENLNQLAICFISWLFRQLIPKIIQTFYCYTEISSTVTIVYFRHDTWNKLTIPFIVEYFKTYLVENNVCRNHNSYTLSNFNHSMKRI
IPKSNNEFRIIAIPCRGADEEFTIYKENHNKNAIQPTQKILEYLRNKRPSTFTKIYSPQTADRIKFKQRLKLFNNVLPPELYFMK
FDVKSICYDISIPRMCMRILKDALKNENGFFVRSQYFNTNTGVLKLFVNVNASRVPKPYELYIDNVRTVHLSNQDVINVVEMEIFKTA
LWVBDKCYIREDGLFQSSLSAPIVDLVYDLDLEFYSEPSQDTLILKLADDFLIISTDQQQVINIKKAMGGFKYNAKANRDK
ILAVSQSDDTVIQFCAMHIFVKELEVVKHSSTMNFNHRSKSSKGFISRLIALFNTRISYKTIDTNLNTNTVLMQIDHVVKNISE
CYKSFAFKDLSINVTQNMQFHSFLQRIIEMTVSGCPTKCDPLIEYEVRFTILNGFLESLSNNTSKPKDNIIILLRKEIQHLQAYIYIYI
HIVN

>NP_001035796.1 telomerase reverse transcriptase [*Tribolium castaneum*]
MVHYRSLKSRQKAPKIVNSKYNSILNIALKNLFLCKKHHTKPKVPQILALLQEIIPKSYFQTTLNKRFRYKVEKILTQSSFECIHL
SVLHKCYDYDAIPWLQVNPENLRPKLLKHNLFLLDNIVKPIIAFYKPIKTLNGHEIKFRKEEYISFESKVFHKLKMKYLVVEVQD
EVKPRGVNLIIPKQDNFRAIVSIFPDSARPPFKLLTSKIYKVLEEKYKTSGLSLYTCSEFTQKTQQQIYGIKVDIRDAYGNVKIIPV
CKLIQSIPHTLLDSEKKNFIVDHISNQAFAFRRIYKWNHGLQGDPLSGCLCELAFMDRLYFSNLKDAFIRHRTVDDYFPCSPHP
HKVYDFELLIKGVYQVNPPTTRTNPTRHPQDEIPYCGKIFNLTRQVTRTYLKLPPNYEIRHFKLWVFNQISDDNPARFLQKAMD
FPFICNSFTKFEFNTVFNDQRTVFANFYDAMICVAYKFDAAAMLRTSFLVNDGGFIWLVLSSTVRAVASRAFKKIVTYKGGKYRKT
FQCLKSIAWRAFLAVLKRREIYKGLIDIRIKSREKLTMKFHDSEVDAKYFCKLPEKFRFVKINRKAS

>AGH33838.1 telomerase reverse transcriptase [*Spodoptera exigua*]
MESPVCLSQYFVNQNVSFNNILKSNILKNESVESFAEFIGDDTSVSDLPDFTDMNILSIKINLRNKSCLDFKDYFPVNNIPQLCP
VDKEYILTNCWKALSELIPRNFGSDKNKREKFDYLRVTVYSMKRQHFIFQRLTVGWDLDVLPWKNLITINGNLIFKILFLWLTXYV
SPMICLVYTTCKLDTNENKLNFWKSQWQCFYDKQVSKMYVFRWVKKCBETSLGKLKKNHNSDLRLKLVAKKSIIPKLYLTKSN
NDCRPIVCYKNESVSQTEKYKIKDRLLKFLRLTLGKPLKLESQYKTLHAKWVAANKPKLYFIKTDLSNAFGCINRDKLAKILSEKHIN
IQKTEKCMVMKKKIAQYRDLVTELKRPILICAGSTVYEWKEGLVQGYKYSFALSELVYSDLEIYFCDHMKLRDQVQKLFIRVVDDY
LYITDNLADATSLFNLSNYRNVNVEKTIIVNFPHENIKQSEDIFFLGVCYNTSTMQVSRADNIFAGQMSYKIAFTSRFSEMFKFIESR

IGQSGIPINSHIFNLNYNNEELIWRHIYTTFFCLSANKLCTIMAILCNEQEMKFLWLYKKRVAVKLSNSMIEVLIRNKPADFIIFYCYI
NHRFLAWKALYLCAKQTPKCTGLVVFINDELTSNCIFGKWEHARRIDTNGECERKSIREVCCRPLRKRIFRDFDVLPGKGFECYHH
KRLI

>XP_003969013.1 Telomerase reverse transcriptase isoform X2 [Takifugu rubripes]
MSITDLSPTLGLIRSLYPHVQVLFVDFADDIVFREGHKATLIEESDTSHFKSFVVRGIVFCFHKELQQVPSNCQICTLPPELLAFVLNSVK
RKRKRNVLAHGYNFQSLAQEERDADQFKLQGDVTSAAAYVHGSDLWRKVMRLGTDITRYLFESCSVFAVAPPSCFLPVCQGIPIYDCF
SLATASLGLFSLQSRGCRCRCLGVNSMKRRAFNVKRYLRKRKTTETDQKDEARVCSGKRRRMEEDKVSCEMTQDGESGKTTLVQKQPGS
KRRSEMEATLLPLEGGPSWRSGTFPPLPPSQSFMRTLGLFYGGRMRSFLLNRRKKTAEGRFKIQRDLIRIVFFEGVLYLNGLERKF
KKLPRFFNMPVPLFSQLLRQHRRCPYSRLLQKTCPLVGIKDAGQAELSSFLPQHCGSHRVLVFRVRELLAVIQELWGSEHNRLLYFA
RVRFFLRSGKFERLSVAELMWKIKVNNCDWLKISKTRVPPSELVSYRTQILGQFLAWLLDGFVVGLVTRACFYATESMGQKNAIRFYRQ
EVMWAKLQDLAFRSHISKQMVELTPDQVAALPKSTIISRLRFIPKTGMRPITRVIIGADAKTRLYQSHVRLDMLRACVCSSTPSSLG
STVWGMTDHVKVLSIAPAQKEKQPPLYFVKMDVSGAYESLPHNKLEIVINQVLTLPVLFNEVFTIRRFACKIADWSHEGLKKAIFRQADF
LEANMGSINMKQFLTSLQKGGKLLHHSVLVEQIFSSDLEGKDALQFFFTQILKGSVIQFGKKTYYRQCQVQPGSAVSVLCCCYGHMEN
VLFKDIINKKSLDMLITPNLHDAQTFLLKILLAGVPQYGLVNVNPKQVVVNQKVVVNFEDYGSTDPCPLRSLTQVFNKDFLIDHDT
TLDIYKDYSSYADLSRYSLTLGSCHSAGHQMKRKLMLGILRLKCHALFLDLKTNLSLEATYKNIYKLLLLHALRHFVCAQSLPFGQSV
KNPAYFLLMWDVMEYNTNYLIRLSNGLISGSTSQTSVQYEAVELLFLCSFLLVLSKHRRLYKDLLLHLLHKKRRRLQCLGLDLRLAR
VRQAANPRNPLDFLAIKT

>AAI63300.1 Telomerase reverse transcriptase [Danio rerio]
MSGQYSTDGGFRPVLEILRSLYPVVQTLLEEFTDGLQFPDGRKPVLEETDGARFKLLSGLIVCAYTPPQLRVPALQSTLPEVLAFTL
NHKRRKLRNVLGFYQCSQSDVTTSSDPFRHGDVSTQTAASISTSEVWKINRQLGTEVTRYLQDCAVFTTVPSCVQLVCCGEPVYDL
LRMSRWSNDRFISGAMRKFPAVQKTVAISKRRFDNEKYISVKRRRVRVETVNNNNNGNYSRSLCLFAISKRRAIDNEENISLKR
RRMEETDQVAKIRNENHESQSAFISKRRARDNEENISLKRQRMEEIDQVAKIRNENHGSQSWKPADQRPPRPSQCSIRVLSMLYNGRG
MKNFLLNRKLGKGGARRMQGEDLVRMIFLQSESNDKPKKLPKRFAMVPLFSRLLRQRHRKCPYRFLRQRKAGNPDVKDMESLKS
HSSPYRVYLVFRVRECLRHIIPEHLWCQENLHFLSNVKNFLLGKFERLTLVQLMWRMKVQACHWLGPKKRQCASEHRYREWMLGQCM
GWWLSGFVGLVRAQFYITESTSMGHKHTLRFYRGDVNSRFLQAFRAHLQVWRPLSPSQALKVPSAVTSLRIFHPKTSMSGLLDTL
SGSRDITLQYFQSCVRVLQNVLSVVCVREAPGPMGSTVWGWQDIHRRQLQDFSPQKSSPRPLYFVKVDVSGAYDSLPHLKLVEVLEKVLG
PFAEQSFFLRQYSSVWSDPTRGLRKRKREKAEEMSEPLNMGKGFVVDEQVSGRLHDAIIVERHSSEVRGGDVQFFQKMLCSYVIHYDQ
MFRVQCGIQGSSVSSLLNLCYGHMEKALLKDIAGGCLMRLIDDPELLTIPHLSKATEFLLTLLSGVPDYGCQINPQKAVANFPVVCV
SWNSGVSVLPSCLPFCWGLMIHTHTLDVYKDYSDYDGLSLRYSLTLGSAHSPTVMKKLLSVLSIKSTDFLDRRLNSVEAVYRSL
YKILLQALRFHACVRSPLPQSVNRNPSFLLKMIWTRVNTNKLTHINKGLPVCVSDSGGVQLQSEAVQLLFLCLAFETLFRFRSRY
HCLIPALHKKRRLQRELCGILTARVVRQASPRIPLDFSMRV

>NP_001079102.1 telomerase reverse transcriptase [Xenopus laevis]
MPLRTGGATLLSILQRLYGQVGLGIVEYTDLTQVPPGGIKVPLVLEGDSEKFRSFVAELVVICIPRGTKPLSPVSVLQSTQREVVARVI
QRIKCKKRNVLAFYGLVDEKNSLNIRLTPNICNYFPNPTTTTISTSILWETLLTRVGDVMMYWLEQCSIFVFVPPRCCYQITGQV
IYTLPSDDVFLFQSQSFTQSNVLLRYIKRNVFHLRKKYLPKPKHSMTRSMRLLTWRNRKSPSGLLIRSKTSMAVTTEIHSKRKLSKDICV
IPDKRRDNLDDKDDTDFHDLPMCRSVSYSLSNMYPKTNVQVTGLITSYGGYKTKTFQCQKPVSECRKKTAFYSVAGDNLSLKDNVN
LITNASVPTAQSRLSFSNIFIDFGRTLYLSISYKKGFSSEFLLNSLSDSTPSGSQKLVETIFLSNFLAEQNFQPKRDENCYKLPKRY
WKMKPHFQELIQNHKFPYLYVNLKHCPVRSSMACSEKRSQKRNRENDGQLKHFPTKANLLSLKQHSIWQVYMFVRECLNNSVP
DIMWSSHNKCRFFRNKVFLLFSGKFKISLSELMWSMRVEDCSWIIRLQKSDHFVPASEHLLREKILAKFVFWMLTDTYVQLLKSFF
YVTEMTFQKRLIFRKSVMKKLQNLIGLRKHLEKVKLRSDELSDELNMQWKNVPLVSLRIFIPKTNGLRPIKISSTLSVQSKENF
EKKIHHFSSQIRNLFVSNLYEWNRRNCCLISSVFGMDDIYKWKKFVLDPEKQVEKLQFYFVKTDVKGADTIIPHKLDEVIKVIN
PNANEVYCIIRYATVSDPTGRIIKSFKRHVSELDVLPNMKQVFSNQQEKNLRLNITLVEQNLNNESSVKLLAVFQQIIRSHILRI
KDRYIMQCCGIPQGSMLSTILCSLCYGDMEAMLGGIQKNGVLMRLIDDPELLTIPHLQAKTFLRLTAEIPIYQYGSISIPQKTVNFP
VDDIPECSEVEQLPSHCFLFRWCGLLDQTLDVYDYSSYACTSIRSMFTFCHSSAAGKYMQLKIRVLRKCHSLFLDLKVLNLSRTV
CINTYKIFLLQAYRFHACVVLQPFQQRVMNNPFFLTVISDMAPCFYTTFKAKNKDLTRGYKDVSCQFNFEAVQLWSYQAFLLKHHN
KVLYKCLIGPLQNCKMQLSRRLSQDTIELKSVTDSLSLHKDFSCIMD

>NP_001026178.1 telomerase reverse transcriptase [Gallus gallus]
MERAQPQGVVRRRLRNVAEPPFAAVLGLALRGYAEATPLAEAFVRLQEGGTGEVEVLRGDDAQCRTFVSQCVCVPRGARAIIPRI
CFQQLSSQSEVITRIVQRLCEKKNLILAYGYSLLDENSCHFRVLPSSCIYSYLSNTVTTETIRISGLWEIILSRIGDDVMMYILLEHCA
LFMLVPPSNCYQVCGQPIYELISRVNGPSPGFVRRRYSRPFKNSLDDYVRRKLVFHRHYSKSSQWVKCRPRRRRGSRRRKRSHRIQ
SLRSGYQPSAKVNFQAGRQISVTVARLEKQSCSSCLPARAPSLRKRKRDEQEVIITAKRVKMEKEIEEQACSIVPDVNSQQSRHGT
SWHYAPRAVGLIKHEHIISERNSEMSPVVRRSHPGKRVADKSSFPQGVQGNKRIKGTGAEKRAESNRREMEYINPIHKNRRGIE
RRINPHTKPELNSVQTEPEMAGASSGDRKQENPPAHLAKQLPNTLSRSTVYFEKFKLILYSRSYQEFYFKSFILSRQLQCAGGRRLIET
IFLSQNPKEQQNSLPQQKWRKRLPKRYWQMREIFQKLKVNHEKCPYLVFLRNKCPVLLSEACLKTELTLQAALPGEARVHKHTE
HGKESTEGETAPNSFLAPPSVLACGQPERGQHPAEGSDPDLRELLRQLHSSHWQYVGFVRECLERVI PAELWGS SHNKRFFKNVKAIF
SMGKYAKLLSQQLMWKMVRNDCVWLRLAKGNHNSVPAYEHCYREEI LAKFLYLWMLDSYVIEILLKSFYITCFMFOKNNMFLYRKF
LQNLGIRDHFAKVLRLALSSEMEVIRQKKYFPPIASRLRFIPKMNGLRVPVRLSRVVEGQKLSKESREKKIQRYNTQLKNLFSVLNYE
RTVNTSIIIGSSVFGRDDIYRKKWKEFVTKVFESGGEMPHFYFVKGVDVSRADFTIPHKLVLEVISQVLKPESQTYVYIRWYAVIMTPTG
KARKLYKRVHSTFEDFIPDMKQFVSKLQERTSLRNAIVVEQCLTFNENSTLFTFFLQMLHNNILEIGHYYICQSGIPQGSILSTLL
CSLCYGDMEKLLCGIQKDDVLIIRLIDDFLLVTPHLMQARTFLRTIAAGIPEYGLFINAKTIVVNFVDDIPGCSKFKHLPCDLRLISW
CGLLLDVQTLEVYCDYSSYAFTSIRSSLSFNSSRIAGKNMKKLTAVLKLKCHPLLLDLKINSLOTLVINIYKIFLLQAYRFHACVQL
LPFNQKVRNPFDFLRIISDTASCCYFILLKAKNPGVLSGSKDASGMFPFEAAEWLCYHAFVVKLSNHKVIYKCLLPLKLVYRMHLFGK
IPRDTMELLKTVTEPSCQDFKTIID

>NP_033380.1 telomerase reverse transcriptase [Mus musculus]
MTRAPRCPAVRSRLRSRYREVWPLATFVRRLGPEGRRLLVQPGDPIYRTLVVAQCLVCMHWGSPPPPADLSFHVQVSSKELVARVVQRL
CERNERNVLAVFGFELLNEARGGPPMAFTSSVRSYLNPVTIETLRVSGAWMLLRSVGDLLVYLLAHCALYLLVSSPCAYQVCGSPLY
QICATTDIWPVSASRYRPTRPVGRNFTNLRFLLQIKSSSRSEAPKPLALPSRGTKRHLSTSTSPSAKAKRCYPPVPRVEEGPQRQVL
PTPSGKSVVPSPARSPEVPTAEKDLSSKGVSDLSLGSVCCKHKPSSTLSSPPRQNAFLRPFIETRHFLYSRGDGQERLNPSFLL
SNLQPNLTGARRLVBEIFLGRSPRTSGPLCRTHRLSRRYWQMRPLFQQLLVNHAECQYVRLRSHCRFRTANQQVTDALNTPSPHLM
LLRHLSSPQVYGFRLACLCKVVSALWGTTRHNERRFFKNLKKFISLGYKLSLQELMWMKVEDCHWLRSSPGKDRVPAEAEHLRE
RILATFLWMLDDTYVQVLLRSLFFYIESTFTQKNRFLFYRKSVWSKLSQIGVRQHLEVRVLRRELSQEEVRRHHQDTWLPALRAIFIPK
PNGLRPIVNMSSYMGTRALGRRKQAQHFQTQLKTLFSLNYERTKHPHMLGSSVGLMNDIYRTWRAFLVLRVRLDQTPRMVYFKADVT

GAYDAIPQGGKLVVAVMIRHSESTYCIQYAVVRRDSQGGVHKSFRQVTTLSDLQPYMGQFLKHLQDSDASALRNSVVEEQSI SMN
ESSSSLFDFLHFLRHSVVKIGDRCYTQCQGI PQGSSLSLTLCSLCFGDMENKLF AEVQRDGLLRFVDDFLVTPHLDQAKTFLSTL
VHGVPYEGCMINLQKTVVNFVPEPGLGGAAPIQLPAHCLFPWCGLLDQTLEVFCDYSYAQTSIKTSLTFQSVFKAGKTMRNKLL
SVLRLKCHGLFLDLQVNSLQTVCTINIKIFLLQAYRFHACVQLPFDQVRKRNLTFFLGI ISSQASCCYAILKVKNPGMTLKASGSFP
PEAAHWLCYQAFLLKLAHSVYIKCLLGLPRTAQKLLCRKLEPATMTILKAAADPALSTDFQTILD

>NP_001268556.1 telomerase reverse transcriptase [Mesocricetus auratus]
MPRAPRCRAVRALLRSQYRQVWPLATFVRRRLGPEGRQLVQPGDKVFRVTLVARCLVCPVWDSQPPPADLSFHQVSSLKELVARVVQRL
CERGERNVLTFFGALLNGAQGGPMTFTTTSVRSYLPNSVTESLRVS GAWMLLNVRVGDLLVYLLARCALYLLVPPSCAYQVCGSPLY
QICATAETWPSVSRIYRPTRPVGRNFTHLGS THRVNRSSHQEAWKPPPLPSREAKRSLSI TNRSVPPSKKARCDLAPRLEKGFYRQAV
PTPSDKTWPVNPASHAVPI SRITTKEDLSSGVKAPGLSRSGSVCYKHKPSSTSLQSPLCQNAFQLRPHYTETKRFYLYSREGGRERLNSP
FLNNLQPSLGTARRLVEILFLGMRPRTSGPLCGRRRLSKRYWQMRPLFQQLLVNHARCPYVRLLRSHCFRTAAHQVAGALNTTSPQ
RLMNNLRLHSSPWQYVGFLLQACVGLVPPGLWGSRHNQRFFRYKRRFVLSLGGYDKLSLQELTWMKVKQDCRWRSSFGNCCVPAAEH
RTRERILAVFLFWLMDAYVVELLRSFFYVTETTTFQKNRLLFFYRKSMMWRLQSIGVRHHLERVRLQELSQQEEVVRQREAWPAMPICRLR
FIPKPSGLRPIVNMYSYMGTRAFDQKQAQHFQCLKTLFVSVLNYELTKHTNLLGASVGLNDIYRTWRTFVLRVRLDPAAPRMYPVFKV
DVTGAYDAIPQDKLVEVIANMIRHPDNSYCIHQYAVVQRDRQGIHKSFRQVSTLSDLQPHMGQFLKHLQDSDTSLRNSVVEEQSL
SLNEASSLFFDFLRFVNRVSVKIGRCYVQCQGI PQGSSLSLTLCSLCFGDMENKLF AEVQRDGLLRFVDDFLVTPHLDQAKTFLSTL
RALVRGIPEYGCMINLQKTVVNFVVDAGTLDGTAPHQLPAHCLFPWCGLLDQTLEVLCDYTYGARTSIKASLTFQRTFKAGRNMRQ
KLLAVLRLKCHSLFLDLQVNSLQTVCTINVIKIFLLQAYRFHACVQLPFDQVHRKNPAFFLSI ISNIASCCYSILKVKNAGMTLKAKG
ASGSFPPEAARWLCYQAFLLKLAGHSVYIKCLLGLPRTAQKQLCRKLEPATMAILETAADPALSTDFQTILD

>NP_001039707.1 telomerase reverse transcriptase [Bos taurus]
MPRAPRCRAVRALLRSYRQVPLAAAFVRRRLRQGHRLVRRGDPAAFRALVAQCLVCPVWDAQPPPAAPSFRQVSCLELVARVVQRL
CERGARNVLAFGFTLLAGARGGPPVAFTTTSVRSYLPNTVTDTLRSGGAWGLLHRRVGDVLLHLLSRCALYLLVPPTCAYQVCGPPLY
DLRAAAAARRPRTQVGGTQAGFGLRPPASSNGGHEAGLEEARAQGARRRRSSARGRLPAPKRPRRGLPEGRDLEGQVARSPPRVV
TPTRDAAEAKSRKGDVPGPCRLFPGGGERGVSASWRLSPSEGEPEGAGACAETKRFYLYCSGGGEQLRRSFLLCSLPPSLAGARTLVETI
FLDSKPGPPGAPRRRRLPARYWQMRPLFRKLLGNHARSY GALLRAHCLPASPAPRAGPDHQKCPGVGGCPSERPAAAPEGEANSGR
LVQLLRQHSSPWQYVGLLRACLRRLVPAGLWGSRHNERRFLRNKVKLLSLGKHGRLSQQELTWMKVKVQDCAWLRASPGARCVPAAEHR
QREAVLGRFLHLMGAYVVELLRSFFYVTETTTFQKNRLLFFRKRKRWSQLRQLGVRQHLDRVRLRESEAEVRQHQEARPALLTSRLRF
VPKPGGLRPIVNVGCVGAPAPPRDKKVVHLSRVKTLFAVLNFERARRPGLLGASVGLMDDIHRARAFVLPRLRAGPAPPLYFVKV
DVVGAYDALPQDKLAEVIANVLPQENTYCVRHCAVMTARGMRKSFKRHVSTFSDFPYLRQLVEHLQAMGSLRDAVVEIQSCSLN
EPGSSLFNLFHLVRSVIRIGGRSYIQCGI PQGSSLSLTLCSLCFGDMENKLF PGVQDGVLLRVDVDFLLVTPHLTRARDLRTL
VRGVPEYGCQVNLRRKTVVNFVPEPGLGGAAPLQPAHCLFPWCGLLDTRTLEVHGHDHSSYARTSIRASLTFQGFKPRGNMRRKLL
AVLQKCHGLFLDLQVNSLQTVFTNVIKIFLLQAYRFHACVQLPFSQVRRSPPAFFLQVIADTASRGYALLKARNAGASLGARGAAG
LFPSEAAQWLCLHAFLLKLAHRVTVYSRLLGALRTARARLHRQLPGFTRAALAAADPALSTADFKTILD

>NP_001177896.1 telomerase reverse transcriptase [Macaca mulatta]
MPRAPRCRAVRALLRSYREVLPLATFVQRLGPEGWRLVQRGDPAAFRALVAQCLVCPVWDARPPPAAPSFRQVSCLELVARVLQRL
CERGARNVLAFGFALLDARGGPPPEAFTTTSVRSYLPNTVTDALRSGGAWGLLHRRVGDVLLVHLLARCALFVLVAPSCAYQVCGPPLY
ELGAATQGRPAHAASGTRRGLGCELANRNSVREAGVPMGLPAPGARRRRGSANRSLPLPKRPRRGAAPERTPVGGQSWAHPDRTRG
PSDRGFVCSVPARPAEAEATSLEALSGTRHSHPSVGRQHAGPPSTSRPPSPWDTRCPLVYAETKHFLYSSGDKEQLRPSFLNLSLRF
SLTGARRLVETIFLGRSPWMPGTPTRRPRLPQRYWQMRPLFLELLGNHAQCYPY GALLKTHCPLRAAVTPAAGVCAREKPPQGSVAPEE
EDTDRRRLVQLLRQHSSPWQYVGFVRACLRRLVPPGLWGSRHNRRFLRNKVKLLSLGKHAKLSLRELTWKMVSVDCAWLRRSPGVGC
VPAEHRRLREELAKFLHLMVSVYVVELLRSFFYVTETTTFQKNRLLFFYRKSVMWKLQSIGIRQHLKRVQLRELSAEARQHREARPTL
LASRLRFLPKPDGLRPIVNMVYVVGARTFRREKRAERLASRVKALFVSVLNYERARRPGLLGASVGLDDIHRARNFVLRVRAQDPPP
KLYFVKVDVMGAYDTIPQDRLEVIASIIKQPNTYCVRRYAVVQKAAHGHVRRKTFKSHVSTLTDLQPYMRQFVAHLQETSPLRDAVVI
EQSSSLNETSSGLFDVFLRFVCHHAVIRIGKSYVQCQGI PQGSSLSLTLCSLCYGD MENKLFAGMRRDGLLRLVDDFLVTPHLTHA
KAFRLRTLVRGVPEYGCVNVNLRKTVVNFVPEDEALGGAAFVQLPAHGLFPWCGLLDTRTLEVQSDYSSYARTSIRASLTFNRGFKAGR
NMRRKLFVGLRKLKCHSLFLDLQVNSLQTVCTNVIKILLQAYRFHACVQLPFFHQVWKNPAFFLRLVSDTASLCSYILKAKNAGMSL
GAKGAAGPLPSEAVQWLCHQAFLLKLTQHRVTVYVPLLGLSRLTAQTQLSRKLPGTTTAALEAAANPALPSDFKTILD

>NP_937983.2 telomerase reverse transcriptase isoform 1 [Homo sapiens]
MPRAPRCRAVRALLRSYREVLPLATFVRRRLGPEGWRLVQRGDPAAFRALVAQCLVCPVWDARPPPAAPSFRQVSCLELVARVLQRL
CERGAKNVLAFGFALLDARGGPPPEAFTTTSVRSYLPNTVTDALRSGGAWGLLHRRVGDVLLVHLLARCALFVLVAPSCAYQVCGPPLY
QLGAATQARPPPHASGPRRRLGCERAWNHSVREAGVPLGLPAPGARRRRGASRSLPLPKRPRRGAAPERTPVGGQSWAHPGRTRG
PSDRGFVCSVPARPAEAEATSLEALSGTRHSHPSVGRQHAGPPSTSRPPSPWDTPCPPVYAETKHFLYSSGDKEQLRPSFLNLSLRF
SLTGARRLVETIFLGRSPWMPGTPTRRPRLPQRYWQMRPLFLELLGNHAQCYPY GALLKTHCPLRAAVTPAAGVCAREKPPQGSVAPEE
EDTDRRRLVQLLRQHSSPWQYVGFVRACLRRLVPPGLWGSRHNRRFLRNKVKLLSLGKHAKLSLRELTWKMVSVDCAWLRRSPGVGC
VPAEHRRLREELAKFLHLMVSVYVVELLRSFFYVTETTTFQKNRLLFFYRKSVMWKLQSIGIRQHLKRVQLRELSAEARQHREARPAL
LTSRLRFIPKPDGLRPIVNMVYVVGARTFRREKRAERLTSRVKALFVSVLNYERARRPGLLGASVGLDDIHRARVTVLRVRAQDPPP
ELYFVKVDVTDGAYDTIPQDRLEVIASIIKQPNTYCVRRYAVVQKAAHGHVRRKAFKSHVSTLTDLQPYMRQFVAHLQETSPLRDAVVI
EQSSSLNEASSGLFDVFLRFMCHHAVIRIGKSYVQCQGI PQGSSLSLTLCSLCYGD MENKLFAGIRRDGLLRLVDDFLVTPHLTHA
KTFRLRTLVRGVPEYGCVNVNLRKTVVNFVPEDEALGGAAFVQMPAHGLFPWCGLLDTRTLEVQSDYSSYARTSIRASLTFNRGFKAGR
NMRRKLFVGLRKLKCHSLFLDLQVNSLQTVCTNVIKILLQAYRFHACVQLPFFHQVWKNPTFFLRLVSDTASLCSYILKAKNAGMSL
GAKGAAGPLPSEAVQWLCHQAFLLKLTQHRVTVYVPLLGLSRLTAQTQLSRKLPGTTTAALEAAANPALPSDFKTILD

Anexo 3: Alinhamento múltiplo de sequência entre diferentes espécies de *Leishmania sp.*

Domínio TEN	
Lpan :	MSSSFPSITGFGPLSLKAFLEEQYFGLRLTFSEVERASFP--PRATTEAETPSAARFRALRDVVLPNQSFLVVVYVALHASSSPP : 82
Lama :	MSSSFPSIPGFGPLSLKAFLEEEYFGLHLTFPAETASPSPRAAATAETPSAAGFRALRDVVLPNQSFLVVVYVALHASSSPP : 83
Imaj :	MSSSFPSIPGFGPLSLKAFLEEEYFGLHLTFPAETASPSPRAAATAETPSAAGFRALRDVVLPNQSFLVVVYVALHASSSPP : 83
Ldon :	MSSSFPSIPGFGPLSLKAFLEEEYFGLHLTFPAETASPSPRAAATAETPSAAGFRALRDVVLPNQSFLVVVYVALHASSSPP : 83
Domínio TEN	
Lpan :	PIITAPGCTTPATLASGCG----ASATGLGWLRCPLAHCPPVATAHDTGVKRGSLTDRENSSVSSLSTGKSGSARCPWRASVP : 161
Lama :	PIITAHASPTFRNTGSRVLRPPRGSAGFRQAAWPIRRRYRQA--HDTGVTRCNFSDRCNFFCLSSSSIGKNGNARSPPWCSAS : 164
Imaj :	PIITAHASPTPTPALGRA----ASATGFERLRCPLHCTVASSAHDTCMTRCNFADRCNSSWSSSSFAGKHGSTRDLWRASAF : 162
Ldon :	PIITAHASPTPTPALGCA----ASATGEGRLRCPLHCTAASFAHDTVVARSEADRENSSCSSSSSSTGKNGNARGIWRSSAF : 162
Domínio TEN	
Lpan :	WLLYTNTSHRPEITVLLRHPWASFAACLGAAMGFITEMYCPIVLQLEAMAGGVQVIGPALKHVALETSFSVPAQSSSEMKGLE : 244
Lama :	WLLYTNTSHRPEITDALLRHPWRASFSACLGAAMGFITEMYCPIVLQLEAMAGGVQVIGPALKHVALETSFSATAQISEMKGLE : 247
Imaj :	RLLYTNTSHRPEITDALLRHPWASFAACLGAAMGFITEMYCPIVLQLEAMAGGVQVIGPALKHVALETSFSAPAPSSSEMKGLE : 245
Ldon :	CLLYTNTSHRPEITDALLRHPWASFAACLGAAMGFITEMYCPIVLQLEAMAGGVQVIGPALKHVALETSFSATAQSSSEMKGLE : 245
Linker	
Lpan :	PHRGANVSSRGVKRAAHAQVCTAPLPIPKQRRVETVEELLKAKAGGGLCRGVLLSSHPKIDGESATFVRSAAIPATWAAALIR : 327
Lama :	PRGSASVSSRGVKRAVDITGCSAPLPLQKQRRVEAVAS--PKKAGRRLQREVAFSHRPPORDLSSSSEPNRAAMPATWAAALIR : 328
Imaj :	PCGSGCISRRGVKRALDTRCSAPLPLQKQRRVEALP---KAKAGRRLQRGVADSRPPORDPSSSVNRAAMATWAAALIR : 325
Ldon :	PCGSGYVSSRGVKRAVDITGCSGFLPLQKQRRVEALP---KAKAGRRLQRGVVECSRPPRDDPSSSEPNRAAMEVTWAAALIR : 325
Linker	
Lpan :	TNPVPRIRLYSVRISDSYVDG--DGALPLPAGELEAGWLRHRHPSLHSLVLCIALPKRDAYGASRRRCSTCTSERAGYTSVED : 409
Lama :	TNPVPRIRLYSVRISDSYVDG--DGALPLPAGSLEAHLWLRHRHPSLHSLVLCIALPKRDAYGASRRHYCACTGERSTGCTSVKN : 410
Imaj :	TNPVPRIRLYSARISDGDGADGSGSALPLPAGELEAGWLRHRHPSLHSLVLCIALPKRDAYGASRRYSVGCASERSTGCTSVED : 408
Ldon :	TNPVPRIRLYSVRISDSYVDGADGSGGALPLPAGELEAGWLRHRHPSLHSLVLCIALPKRDAYGASRRYSVGCASERSTGCTSVED : 408
Domínio TRBD	
Lpan :	IPMWHVAVHVRWLVLCPECSGADASEPTTIRFDLPSILRRRLSTVVDQCSRLLDLRGAALKHAGYVEEAFRRQQQGVPEPVDVQR : 492
Lama :	IPMWHVAVHVRWLVLCPECSGGEVAFCTTSPKFDLPSILRRRLSTVVDQCSRLLDLRGAALRHGYLEEAFFRRQQQGVPEPVDVQR : 493
Imaj :	IPMWHVAVHVRWLVLCPECSADASEPTTTPKFDLPSILRRRLSTVVDQCSRLLDLRGAALKHGYLEEAFFRRQQQGVPEPVDVQR : 491
Ldon :	IPMWHVAVHVRWLVLCPECSGADASEPTTTPKFDLPSILRRRLSTVVDQCSRLLDLRGAALKHGYLEEAFFRRQQQGVPEPVDVQR : 491
Domínio TRBD	
Lpan :	LSTPVDVVVSYLRTLLSTLRWAPLKEAKESFWGRTAAGSERVLDALMRAVQWLIAGRHAVPVSFRFDGVPVAQVPWLNPF : 575
Lama :	LSTPVDVVVSYLRTLLSTLRWAPLKEANNCLFWGRTAAGSERVLDALMRAVQWLIAGRCAVVPVSFRFDGVPVAQVPWLNPF : 576
Imaj :	LSTPVDVVVSYLRTLLSTLRWAPLKEADNGSFWGRTAAGSERVLDALMRAVQWLIAGRHAVPVSFRFDGVPVAQVPWLNPF : 574
Ldon :	LSTPVDVVVSYLRTLLSTLRWAPLKEAKNRSFWGRTAAGSERVLDALMRAVQWLIAGRHAIVPVSFRFDGVPVAQVPWLNPF : 574
Domínio TRBD	
Lpan :	YTS--PSTSSAATSSAARHARRHRSRVQQRVWLQFVLFLLTQDILPFLLRASFTITWSSKNTHKLLFFPAVWRRLLRREVRRRT : 657
Lama :	YTTTPSLPFSLATSSAARHARRERSVQQRVWLQFVLFLLTQDILPFLLRASFTITWSSKNTHKLLFFPAVWRRLLRREVRRRT : 659
Imaj :	YTTTPSLPFAATSSAARHARRERSVQQRVWLQFVLFLLTQDILPFLLRASFTITWSSKNTHKLLFFPAVWRRLLRREVRRRT : 657
Ldon :	YTTTPSLPYSAATSSAARHARRERSVQQRVWLQFVLFLLTQDILPFLLRASFTITWSSKNTHKLLFFPAVWRRLLRREVRRRT : 657
Domínio TRBD	
Lpan :	RSCGAPRSCMELACESARRTGAEPALSSAFSACCGVVTAAILTEHSAVSKASAAIAPRDVWRVAVRTGCVLAHRGARATLAT : 740
Lama :	RSCGAPRSCMELACESARRTGAEPALSSAFSACCGVVTAAILTEHSAVSKASAAIAPRDVWRVAVRTGCVLAHRGARATLAT : 742
Imaj :	RSCGAPRSCMELACESARRTGAEPALSSAFSACCGVVTAAILTEHSAVSKASAAIAPRDVWRVAVRTGCVLAHRGARATLAT : 740
Ldon :	RSCGAPRSCMELACESARRTGAEPALSSAFSACCGVVTAAILTEHSAVSKASAAIAPRDVWRVAVRTGCVLAHRGARATLAT : 740
Domínio RT	
Lpan :	HSGGASCLYAGVRFPRDRKLRPIAVVRSASLRSLKEMARGSPSPYSYSHSTVRLLRRLGCSIDGQLPAATATLLRRVQARS : 823
Lama :	RGGGASCLYAGVRFPRDRKLRPIAVVRSASLRSLKEMARGSPSPYSHASATVRLLRRLGCSIDGQLPAATATLLRCVQARS : 825
Imaj :	RGGGASCLYAGVRFPRDRKLRPIAVVRSASLRSLKEMARGSPSPYSHASATVRLLRRLGCSIDGQLPAATATLLRRVQARS : 823
Ldon :	RGGGASCLYASVRFPRDRKLRPIAVVRSASLRSLKEMARGSPSPYSHASATVRLLRRLGCSIDGQLPAATATLLRRVQARS : 823

Domínio RT

Lpan : RRRNRTGFRFRHPLPHTLPHRAALCNALRCLVSGVEEQRVRDGLPRLSNLHQDEYAEIERSFCCEVRECHALSCEGLACTHS : 906
Lama : RHNRRIGVHRSAPLPPHLPKHAALCDALRCLVSGVEEQRVRDGLPRLSNLHQDEYAEIERSFCCEVRCRHAIECEGPAKTTF : 908
Imaj : RHNRRAGVHREHPLPPLPHKAAIRDALRCLVSGVEEQRVRDGLPRLSNLHQDEYAEIERSFCCEVRCRHALLCEGPACTAF : 906
Ldon : RHNHHTGVHREHPLPPLPHKAAIRDALRCLVSGLEEQRVVDGLPRLSNLHQDEYAEIERSFCCEVREHFAIECEGPACTAF : 906

Domínio RT

Lpan : AASPNEPCGAAA-CETFCAMQTVTLVRSASRCYDNLPCQERVLAAVSLVKHDAYRVLHFTVIHALNGEAPHAGGCLLRRFT : 988
Lama : AVSSASPPGATA-CFAPY----VTLRSDASRCYDNLPCQGRVLAAVSLVKHDAYRVLHFTVIHAVISETCKGGCLLRRFT : 986
Imaj : AVSSAEPFGSAACCFAPY----VTLVRSASRCYDNLPCQERVLAAVSLVKHDAYRVLHFTAIHALDGEACRGGCLLRRFT : 985
Ldon : VVSSASPESAACCFAPY----VTLVRSASRCYDNLPCQERVLAAVSLVKHDAYRVLHFTAIHALDGEAACKGGCLLRRFT : 985

Domínio RT

Lpan : TRTIPYTEMCEGVLARIPRGHIYW-EEEEGTQGGPHATAAASCAGIPSRRCGENFISGAAVRALLSEHVRHHLVVVSGGSLF : 1070
Lama : TRTIPCAEIECGVLARIPRGHIYW-EEEEGTPTGHTSAAVSRITDASSRCGANLISGAAVRALLSEHIRHHLVVVSGGSLF : 1067
Imaj : TRTIPCAEIECGVLARIPRGHIYW-EEEEGTQGGPHATAAASRTTANRRRCGANLISGAAVRALLSEHIRHHLVVVSGGSLF : 1068
Ldon : TRTIACAEIECGVLARIPRGHIYW-EEEEGTLSGHTTAAVSRITNASSRCGANLISGAAVRALLSEHIRHHLVVVSGGNLL : 1068

Domínio RT

Lpan : EQRVGIQSGSPVAMLLCDRLFADVVDTALSSILSEHAERSLLRRVDDVLVATSPAAAERCLAMCGWPSVGYLSNPSKLT : 1153
Lama : EQRVGIQSGSPVAMLLCDRLFSDVVDTALSSILSEHAERSLLRRVDDVLVATSPAAAERCLAMCGWPSVGYLSNPSKLT : 1150
Imaj : EQRVGIQSGSPVAMLLCDRLFSDVVDTALSSILSEHAERSLLRRVDDVLVATSPAAAERCLAMCGWPSVGYVSNPSKLT : 1151
Ldon : EQRVGIQSGSPVAMLLCDRLFSDVVDTALSSILSEHAERSLLRRVDDVLVATSPAAAERCLAMCGWPSVGYLSNPSKLT : 1151

Domínio RT

Domínio CTE

Lpan : LSNACGSIVPWCGLLLHDTTLEVSVEWRRIGVLLSLRVGDFHYVHRGDHDEPLYLTQRFVAVLQLRVPTALCGRMNSKTRQL : 1236
Lama : LSNACGSIVPWCGLLLHDTTLEVSVEWRRIGVLLSLRVGDFHYVHRGDHDEPLYLTQRFVAVLQLRVPTALCGRMNSKTRQL : 1233
Imaj : LSNACGRIVPWCGLLLHDTTLEVSVEWRRIGVLLSLRVGDFHYVHRGDHDEPLYLTQRFVAVLQLRVPTALCGRMNSKTRQL : 1234
Ldon : LSNACGSIVPWCGLLLHDTTLEVSVEWRRIGVLLSLRVGDFHYVHRGDHDEPLYLTQRFVAVLQLRVPTALCGRMNSKTRQL : 1234

Domínio CTE

Lpan : QTFYEVGLLWSRVVLEKVQETLPVARNRQVAVLLRPLAVCVGRLLRLLHRHQRFLAARQSSACVSAAEVRACVLTALHRTVC : 1319
Lama : QTFYRVGLLWIRVVLEKVQEALPVARNCCVAVLLRPLAVCVGRLLRLLHRHQRFLAARQSSACVSAAEVRACVLTALHRTVC : 1316
Imaj : QTFYEVGLLWSRVVLEKVQEALPVARNRQVAVLLRPLAVCVGRLLRLLSGHQRFLAARQSSACVSAAEVRACVLTALHRTVC : 1317
Ldon : QTFYEVGLLWSRVVLEKVQEALPVARNRQVAVLLRPLAVCVGRLLRLLNRHQRFLAARQSSACVSAAEVRACVLTALHRTVC : 1317

Domínio CTE

Lpan : AKLRVLQARTIRVMCECRCLPCRGASINRCPKSLCAHKAESAGHCQCKRRRKGCRRRINERTLRFSFWWMAAAEVELQWRR : 1402
Lama : AKLRVLQARTVRFAMIVCSGRTRCG-FSLKGRNNFCSKASSVGRKQCECRRRGRNRR-NTRVCLRSFWWLTAAEVESQWRRS : 1397
Imaj : AKLRVLQARTVVCVTAQRGRTERG-TSLKGRNNHFCSEHATSAGRQCCERRRGRNRRNTRVCLRSFWWLAAAEVESQWRRS : 1399
Ldon : AKLRVLQARTVVRVMTACRGRTERG-TSLKGRQKPPCINKASSAGCKYCERRRGRNRRCS-TRVCLRSFWWLAAAEVELQWRRS : 1399

Domínio CTE

Lpan : LQALHRAASFA--LGGHSTAFSSATSASLLMEDGLGSMHARALSATRLSCA : 1452
Lama : LQALYRAAPRPGEGCGSTASPSFASASLLMEDGPI SMHARALSATRLSQI : 1449
Imaj : LQALHRAAPRLGEARGPTALSPSECSASLLMEDGPI SMHARALSATRLSQI : 1451
Ldon : LQALHRAAHPGEGARGSTAFSLSESASASLLMEDGPI SMHARALSATRLSQI : 1451
LgAlhRAApr gea G ta spSp aSASLLMeDGpLSMHARALSATRLSQI

Anexo 4: Alinhamento de sequência múltiplo entre diferentes Tripanossomatídeos

Domínio TEN	
Lama :	MSASFPSPGFRGPTSLKAFLEEDYFGLHITFAAETASPSRAAATAETPSAAGFRALRTVVLPNQSFLVVVYVALHASSEF : 82
Lmaj :	MSASFPSPGFRGPTSLKAFLEEDYFGLHITFAVETASPSERAATAETPSAAEFRALRTVVLPNQSFLVVVYVALHASSEF : 82
Tcru :	MSASFPSPGFRGPTSLKAFLEEDYFGLHITFAAETASPSRAAATAETPSAAGFRALRTVVLPNQSFLVVVYVALHASSEF : 67
Tbru :	MAQMPFSPGFRGPTSLKAFLEEDYFGLHITFAAETASPSRAAATAETPSAAGFRALRTVVLPNQSFLVVVYVALHASSEF : 68
Domínio TEN	
Lama :	PEPTAHASPTERTNTGSRVALRPPRGSAGFRQAAWPIRRRY--RQAHITCVTRC-NPS-RQNE-FCLSSSSTCKNGNARSFWCA : 161
Lmaj :	PEPTAHASPTERTPT-----ALGRAASATGFERLRQLPHTQTVASSAHTCMTRC-NPA-RONSFWSSSSFAKKGHSTRDPWRA : 159
Tcru :	PE-----PEPEPAP-----LLVTTGEGEFASKGHDLLE-RATS-LTPFG----- : 104
Tbru :	PE-----QRSVRSV-----AVLSSCSNM--DG-HPQS-VPHFFG----- : 102
Domínio TEN	
Lama :	SASWLLYNTSHRPTDALLRHPWASFSAGLGPAMGFLEMYPVIVLQLEAMAGGVQVIGPRLKHALETFSFSATAQLSEM : 243
Lmaj :	SAFRLLYNTSHRPTDALLRHPWASFAACLGPAMGFLEMYPVIVLQLEAMAGGVQVIGPRLKHTAFERSFSAAAPSSEM : 241
Tcru :	-----RPLTSSLLRFPFNGLFTCLGPAVSVFILAWCPVVVCEEPENGGGLQVLGELKN----- : 158
Tbru :	-----RELTSSVLKHPFDALFSCITGFAVSVIVLWAPIVVVCEEPASGGGLQVLGELK----- : 156
Linker	
Lama :	KGLPPRGSASVSSRGVKRAVDTEGCSAPLEPLCKQKQVEAVASLP--KAGRRLCR--EVAPSHRPCRDPFSFSPVNRAAMPATW : 322
Lmaj :	KGPLPQSGSGCISRRGVKRALDTRCSAPLEPLCKQKQVEA---LPKAKAGRRLCR--GVAEDSRPCRDPFSFSSVNRAAMAATW : 319
Tcru :	-----AVERAVNRCRG-----EGW-TGSRKMGVYERNACHRILOGS-SESLI----- : 199
Tbru :	-----SASCPVWRCPG-----AGWWEGLKIVSR-----TETLTQNDPHCFCTEGTSSP----- : 199
Linker	
Lama :	AAALTRLDDETRLYSRYVSDSDCTGGG--GALPLPAGSLEAWPPIPHPRSLHRVLOAALPKRAAYGASTRHYCAGTEERGT : 403
Lmaj :	AAALTRLDDETRLYSARISDGDGAGGGSGALPLPAGSLEACWIRREPHSLHRAVLOAALPKRAAYGASTRYSVYGASERGT : 401
Tcru :	--SILYRNVSRLEPL--IGEAPLCTFAEIDRRVNER---LSKTWQEEH---GGMGAP-----SSCRCSVTAGAVPM : 261
Tbru :	-VALYRNVSRLEPL--VECSPLRSTALAHAVTLL--LEQWIEPHKSSCSPPSSDRLL--SCLLCAVATSATN : 266
Domínio TRBD	
Lama :	GC---TSVKNSMWEVAHVERWVLOCSGSEVAFQTPSPKFDLPSYLRRELTSTVDDQSRLLDRGAALRHTGYLEDAFR-- : 480
Lmaj :	GC---TSVEDPMEVWTHVERWVLOCPQSADASPTTPPKFDLPSLRRLSTVDDQSRLLDRGAALRHTGYLEDAFR-- : 478
Tcru :	PCADACSVLAVSRELVHLVFEYTL---SINVDGENTVKMASYEREITHVHCVLSGVCCMNERGAAIKHTGFLEERLSLF : 338
Tbru :	CCGDSCVVLAQARNQVNAVEYTL--VHDDAVSNGTVEK-----YLHVLHCLGVSQRMNIRGAAIKHTGFLEEKRSV : 342
Domínio TRBD	
Lama :	-----RCQQGVEPMDVQRLSTFVDVVVSYLRTLLPTLWAPLK-----EANGLFWGRDAGSERVIDALMRAVR : 545
Lmaj :	-----RCQQGVEPMDVQRLSTFVDVVVSYLRTLLPTLWAPLK-----ETNGSFWGRDAGSERVIDALMRAVR : 543
Tcru :	NHVRPR---GEAGACAAEVSSELTVEETINTSYHSLGGTWRAPQ-----ATERVNFWG-EL---VLERLCAITL : 404
Tbru :	TSISQKAQTACQENKRAIFVGETVVEEVVASYLETFLNMMWRVPPRFVGGAVTARVKFWGEEGT---VLTRELVTV : 420
Domínio TRBD	
Lama :	CWLIAGRQAVVSVSRFLDGVVAVCPWLNIFYTTTSLPFSDATSSASRHARRERSQVQCRVWLCGALFLDCLLPFLLRAS : 627
Lmaj :	CWLIAGRAVAVVSVSRFLDGVVAVCPWLNIFYTTTSLPFPATSSASRHARRERSHVQCRVWLCGALFLDCLLPFLLRAS : 625
Tcru :	SWLQCGRENEFELSEFLHCIPVAKLPWLRGFYTRGS-----GRRRSVQCRVWLCGFFFLCHVVPFLVRS : 472
Tbru :	DWICCGREVEVILSEFLDGVVAVCPWLRGFYTKG-----SRKRSVQCRVWLCGFLFLCQVVPFLVRS : 488
Domínio TRBD	
Lama :	FTITWSSKNTHKLLFEPFVWRRLVREVRTRSCHAERSQMSLAERTMSCADHAATVSAENACCGAVAAAFPSPHSAAS : 709
Lmaj :	FTITWSSKNTHKLLFEPFVWRRLVREVRTRSCGAPRSQMLPAGESARRIGEPAAISSAPSACCGVWTAALLPHSAVS : 707
Tcru :	FHVWSSKKNPAFLFIPKSVWIRLVSRRLQVCLRRRRRKRERDGDGAPQMA--LPPMALERLTSNTVGLKLEPGK---- : 549
Tbru :	FEVACTSRSPYIFFEIERAVVVRITREMRVVCIRRVKRMRSDTLSDPQMRGHGLPPMALERVTSEGLNQLIGPSSQPVGD : 570
Domínio TRBD	
Lama :	NASAAISAPRDEWRAVRTGGALAHWSARATATRGGGSCLYAGVREPERRKLRFIAVRS--SLRSIKENARGSESYSHA : 791
Lmaj :	KASAAIAAPRDEWRAVRTGGALTRGARATATRGGGSCLYAGVREPERRKLRFIAVRS--SLRSIKENARGSESYSHA : 789
Tcru :	-----DAIPAGTTAPELLYSSIRELLECKLRFIAVRS--NARWLLKMAADCLCS----- : 599
Tbru :	DSG-----GVSDVCKLIDSSPELLYSDVRELVGSKLRFIARFREGWQSSLLKPADEVCSS----- : 626

Domínio RT

Lama : SAIVRLRLRLGCSADADGQLEAATATLLRQVQARSRHNRRTGVHRSAPHLPPHLPKKAALQALRCLVSGVEEQVRVDGLPRL : 873
Imaj : SAIA RLRLRLGCSDTDGQLEAATATLLRFRVQARSRHNRFRAGVHRFPHLPKKAALRCLVSGVEEQVRVDGLPRL : 871
Tcru : -----ATLSVCNFSCRAVASTHTQTKRSSAVS-----PNATVLRGALRCLLAGLAEQVRVGSLSVKF : 654
Tbru : -----TASLAKAGKYLATDGKAWPES-----ENSVLRDALRCLDAGVEERITTEARQR : 678

Domínio RT

Lama : SNLSHCDEYAEELRSFCEEVGRHAIPECEGPAKTTPAVSSASPPGATA-CFAPYVTLRLSDASRCYDNLPCGRVLAAVRSLVK : 954
Imaj : SNLSHCDEYAEELRSFCEEVGRHALLCEGPAKTAPAVSSASPPGSAACCFAPYVTLVRSASRCYDNLPCQERVLAARVSLVK : 953
Tcru : TNLSHCDEYAEVRSFVEGARN----FHHPLLASTPESAAEVSSA----VVTMVRGDARCYDLPQEAVMNTVRLLP : 727
Tbru : AIGSHNDEYAEVRSFICCARQ----FCITIEPLVQGETSALLADATA----VTMVRGDARCYDELPQEAVVNIMRGLVT : 750

Domínio RT

Lama : HDAYRVLRFVVIHAVDS-----EATCKGGCLIRRTFTT----RTIECAEAECGFLARIPRGHIWEEF : 1013
Imaj : HDAYRVLRFVTAIHADG-----EGACRGGCLIRRTFTT----RTIECAEAECGVLARIPRGHIWEEF : 1012
Tcru : HEFVHTLSFTAVAFIAGGRGTANVEEVGGQGWKRAEDAPSVVLLHRRVKL----CTVSGKVCVNGTIPGIRKGMWYEEF : 805
Tbru : HETVYSLRFTVV-TLTTVKNVSKKTRSSHAESQLSSMKETNGSTVKKRKLQLSKRVKLVGRYVCEVLYGLARGCIAYEEC : 831

Domínio RT

Lama : --ERTPTGPHTSAAVSRTTDASSRCGANLISCAAVRLLSEHIREHLVVVSGGSLFECRVGILCGSEVAMLLCDRLFSNVVD : 1093
Imaj : EEHTQGGHHTAAVSRRTDANRRCGANLISCAAVRLLSEHIREHLVVVSGGSLFECRVGILCGSEVAMLLCDRLFSNVVD : 1094
Tcru : TPP-----SESVMRGDEMRTVIGCHLQKHLVVE-GRLYSCGVGITCGSEVAMLLCDILL-ETVD : 863
Tbru : SVE-----SDSLVSCEEIRSVLQKHLQSHLVMIN-GKMYICRMGINCGSEVAMLLCDILL-ERVD : 889

Domínio RT

Lama : TALSSILSEHAERSLLRRVDDVLVPTTSPAAERCLRAMCGWESVGYLSNPSKILTSLKACGSIVPWCGLLHDDTLEVSFV : 1175
Imaj : TALSGILSEHAERSLLRRVDDVLVPTTSPAAERCLRAMCGWESVGYVSNPSKILTSLACGRIVPWCGLLHDDTLEVSFV : 1176
Tcru : HSLNLSLQHEEPPALLRRVDDVLVWILSEVAARCDVALRHGWEVGFQDDEKLRCTTVA--IIPWCGLLWNPCTLEESV : 943
Tbru : AALSSILSEHEEPPALLRRVDDVLVWILSRRAASRCEDAIRSGWSEIGFECQEBKLRVH-CGQFVRCGLLWDPVTLLEFAV : 970

Domínio CTE

Lama : EWRIRIGVLLGSLRVDGPHYVHGDYEPILYLTQRELAFLVCLRVAETALCGRMNSKTRCLQTEFRVGLLWIRVVLKVOBALPV : 1257
Imaj : EWRIRIGVLLASIRVDGPHYVHGDHPEILYLTQRELAFLVCLRVAETALCGRMNSKTRCLQTEFRVGLLWIRVVLKVOBALPV : 1258
Tcru : EWRRLSSLLPHIASRP----MTGNEPLHCSIRMSIILCLRTPTVLCRINTKTRVVCTLCRIGFLWSRFFLRKLLNAMEF : 1020
Tbru : DWARIAKMPEYLAIRP----RTGCEPILSSREIRIRLRTRPTALCRGINSKSRVQCTLYEIGLLWSRFFLRKLLKANAFAF : 1047

Domínio CTE

Lama : ARNCCVTVLLRPLAVCVGRICRLLTRHQRFIAARCSACDVSAPEVRACVLTALERTVOAKIRVILQARTVRAMTVQSGRTQR : 1339
Imaj : ARNRCVAVLLRPLAVCVGRICRLLSQRFLAARCSACDVSAPEVRACVLTALERTVCAKIRVILQARTVCVVTACRGRTPR : 1340
Tcru : IRPHVRV--LLRPLALATASIRRLIRKCKGIERLGSFCDVTDDEIRLCVHFTLEHTLCEWIPRMISQI----- : 1087
Tbru : FRPHVRT--ILCPLALATATIRRLVRKHSFDLQRLGSEFCDVTDVEVRLCISALLYETIHQRIFAEMIG----- : 1112

Domínio CTE

Lama : GPSTLKGRRNFCSTKASSVGRKGCEQRRRGRN--RNTRVCLRSFWWITPAEVESQWFRSI-----GAIYRAAPRPEAC : 1412
Imaj : GTSLKGRRNHFCSSHEALSAGRKCCERRRGRNRRNRTRVCLRSFWWLAPEVESQWFRSI-----SAIHRAAPRLCEAR : 1414
Tcru : -----RRREGIKCFELLISALVRRKMAAALMDGASCEGDDEHIAASRGGSH : 1135
Tbru : -----RMSVIGKKKELILMTAVLKRKM-----GEL-----GSRINRNN : 1145

Domínio CTE

Lama : GSTASSF--SPAASALLMETGPLSMHRRALSATRLSCL----- : 1449
Imaj : GPTALSF--SPGASALLMDIGPLSMHRRALSATRLSCL----- : 1451
Tcru : GVAFAAVNEDDDVGRKLLLABGDDIVIVRRLAAVRFPPSRAFTHGENLWRLSEDPTGGRLKNNKILLPTVLK : 1207
Tbru : QMESSLE-----SALILLVVEGDDGVSRGLAVRFRFVETVRFQGCATRSA----- : 1192

APÊNDICES

Apêndice 1: Capítulo de livro: “NUCLEAR AND KINETOPLAST DNA REPLICATION IN TRYPANOSOMATIDS”

Molecular and Cellular Biology of Pathogenic Trypanosomatids, 2016, 86-125

CHAPTER 4

NUCLEAR AND KINETOPLAST DNA REPLICATION IN TRYPANOSOMATIDS

Marcelo S. da Silva¹, Maria Alejandra Viviescas², Raphael Souza Pavani¹, Edna Gicela Ortiz², Camila B. Storti² and Maria Isabel N. Cano^{2*}

¹Special Laboratory of Cell Cycle, Butantan Institute, São Paulo, Brazil; Center of Toxins, Immune Response and Cell Signaling-CeTICS, Butantan Institute, São Paulo, Brazil. ²Departament of Genetics, Biosciences Institute, Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP), Botucatu, São Paulo, Brazil

Abstract: Trypanosomatids are among the most primitive eukaryotes and therefore exhibit both conserved and unique non-conserved features in the DNA replication machinery. In eukaryotes, nuclear DNA replication is preceded by the assembly of the pre-replication complex (pre-RC), which is coordinated by the six-subunit origin recognition complex (ORC), which together with the Cdc6 and Cdt1 proteins play a central role in the loading of the hetero-hexameric Mcm2-7. In the domain Archaea there are no Cdt1 protein homologs, Mcm is a homo-hexamer, which is recruited by a protein that shows homology with ORC, and Cdc6 (called Orc/Cdc6). Curiously, trypanosomatid pre-RC differs from others eukaryotes in this context, and it appears more similar to that of Archaea, presenting a homolog of protein Orc/Cdc6 and no homologs of Cdt1, in addition to present Mcm as a hetero-hexameric complex. The completion of DNA replication, at trypanosomatid telomeres, apparently is similar to other eukaryotes, although the processing of the leading and lagging telomeres required to generate the 3' overhangs, which serves as telomerase substrate, remains unknown. With the generation of overhangs at the ends of the chromosomes, telomeres are frequently extended by the action of telomerase, whose control also remains unknown. It is worth mentioning that DNA replication in trypanosomatids initiates almost simultaneously in the nucleus and the kinetoplast, suggesting that regulation of DNA synthesis in the two DNA-containing organelles may be coordinated. The kinetoplast DNA (kDNA) consists of mini- and maxicircles, which are replicated by many proteins whose mechanisms of action remain unclear. This chapter aims to review and discuss the complex DNA replication mechanisms that act independently in the kinetoplast and the nucleus, as well as some fascinating peculiarities exclusive to trypanosomatids protozoa group.

Keywords: Nuclear DNA replication, kinetoplast DNA replication, DNA polymerases, Origin Recognition Complex, Origin Licensing, Origin Firing, topoisomerases, replisome, replication forks, telomere replication.

***Corresponding Author:** Maria Isabel N. Cano, Genetics Department - IB, São Paulo State University “Júlio de Mesquita Filho” (UNESP), Botucatu, São Paulo, Brazil; e-mail: micano@ibb.unesp.br

**Marcelo S. da Silva and Maria Isabel N. Cano (Eds.)
All rights reserved - © 2016 Bentham Science Publishers**

Apêndice 2: Artigo publicado em BBA-General Subjects: “A calmodulin-like protein (LCALA) is a new *Leishmania amazonensis* candidate for telomere end-binding protein”

BBA - General Subjects 1861 (2017) 2583–2597



Contents lists available at ScienceDirect

BBA - General Subjects

journal homepage: www.elsevier.com/locate/bbagen



A calmodulin-like protein (LCALA) is a new *Leishmania amazonensis* candidate for telomere end-binding protein



Edna G.O. Morea^{a,1}, Maria Alejandra Viviecas^{a,1}, Carlos A.H. Fernandes^{b,1}, Fabio F. Matioli^b, Cristina B.B. Lira^a, Maribel F. Fernandez^c, Barbara S. Moraes^d, Marcelo S. da Silva^e, Camila B. Storti^a, Marcos R.M. Fontes^b, Maria Isabel N. Cano^{a,*}

^a Genetics Dept., Biosciences Institute, São Paulo State University (UNESP), Botucatu, SP, Brazil

^b Biophysics and Physics Dept., Biosciences Institute, São Paulo State University (UNESP), Botucatu, SP, Brazil

^c Instituto Tocantinense Presidente Antonio Carlos LTDA., ITPAC-Porto Nacional S.A., TO, Brazil

^d PROAHSA - Programa de Estudos Avançados em Administração Hospitalar e Sistemas de Saúde, São Paulo, Brazil

^e Laboratório Especial de Ciclo Celular, Center of Toxins, Immune Response and Cell Signaling (CeTICS), Butantan Institute, São Paulo, SP, Brazil

ARTICLE INFO

Keywords:

Leishmania amazonensis
Telomere end-binding protein
Telomerase
Calcium
Calmodulin

ABSTRACT

Background: *Leishmania* spp. telomeres are composed of 5'-TTAGGG-3' repeats associated with proteins. We have previously identified LaRbp38 and LaRPA-1 as proteins that bind the G-rich telomeric strand. At that time, we had also partially characterized a protein: DNA complex, named LaGT1, but we could not identify its protein component.

Methods and results: Using protein-DNA interaction and competition assays, we confirmed that LaGT1 is highly specific to the G-rich telomeric single-stranded DNA. Three protein bands, with LaGT1 activity, were isolated from affinity-purified protein extracts *in-gel* digested, and sequenced *de novo* using mass spectrometry analysis. *In silico* analysis of the digested peptide identified them as a putative calmodulin with sequences identical to the *T. cruzi* calmodulin. In the *Leishmania* genome, the calmodulin ortholog is present in three identical copies. We cloned and sequenced one of the gene copies, named it LCaLa, and obtained the recombinant protein. Multiple sequence alignment and molecular modeling showed that LCaLa shares homology to most eukaryotes calmodulin. In addition, we demonstrated that LCaLa is nuclear, partially co-localizes with telomeres and binds *in vivo* the G-rich telomeric strand. Recombinant LCaLa can bind specifically and with relative affinity to the G-rich telomeric single-strand and to a 3'G-overhang, and DNA binding is calcium dependent.

Conclusions: We have described a novel candidate component of *Leishmania* telomeres, LCaLa, a nuclear calmodulin that binds the G-rich telomeric strand with high specificity and relative affinity, in a calcium-dependent manner.

General significance: LCaLa is the first reported calmodulin that binds *in vivo* telomeric DNA.

1. Introduction

Telomeres are essential terminal structures at eukaryote chromosomes, mainly composed of repetitive DNA in a double- and single-strand form associated with proteins [1]. Telomeres are maintained by the action of telomerase, a specialized reverse transcriptase that replenishes telomeric repeats at the end of the chromosomes, using its intrinsic RNA component (TER, telomerase RNA) that bears the template sequence used by the TERT (telomerase reverse transcriptase) component for the telomere repeat addition [2]. In most organisms, telomeric proteins play crucial roles in telomere maintenance and

homeostasis and are found forming huge complexes. Some of these proteins bind specifically to the double-strand form of telomeres, while others, known as the telomere end-binding proteins, show high affinity for the G-rich single-stranded telomeric DNA. Only a few proteins in these complexes do not interact with telomeric DNA, but bridge double- and single-strand telomeric proteins with each other or with telomerase. This results in dynamic and regulated interactions to maintain telomere homeostasis and to protect chromosome ends from being recognized as damaged DNA [1,3–7].

In *Leishmania* spp., telomeres are composed of the conserved repetitive sequence 5'-TTAGGG-3' and also maintained by the action of

* Corresponding author at: Genetics Dept., Biosciences Institute, São Paulo State University (UNESP), SP, CEP 18618-689, Brazil.

E-mail address: micano@ibb.unesp.br (M.I.N. Cano).

¹ These authors contributed equally to the paper.

<http://dx.doi.org/10.1016/j.bbagen.2017.08.011>

Received 18 January 2017; Received in revised form 6 August 2017; Accepted 14 August 2017

Available online 24 August 2017

0304-4165/ © 2017 Elsevier B.V. All rights reserved.

Apêndice 3: Artigo publicado em Biochimie: “Structural studies with BnSP-7 reveal an atypical oligomeric conformation compared to phospholipases α -like toxins”

Biochimie 142 (2017) 11–21



Contents lists available at ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi



Research paper

Structural studies with BnSP-7 reveal an atypical oligomeric conformation compared to phospholipases A_2 -like toxins



Lino Fernando G. de Lima^a, Rafael J. Borges^a, Maria Alejandra Viviescas^{a, b}, Carlos A.H. Fernandes^{a, **}, Marcos R.M. Fontes^{a, *}

^a Departamento de Física e Biofísica, Instituto de Biociências, Universidade Estadual Paulista (UNESP), Botucatu, SP, Brazil

^b Departamento de Genética, Instituto de Biociências, Universidade Estadual Paulista (UNESP), Botucatu, SP, Brazil

ARTICLE INFO

Article history:
Received 18 April 2017
Accepted 21 July 2017
Available online 25 July 2017

Keywords:
Myotoxic mechanism
Snake venom
Phospholipases A_2 -like toxins
Oligomeric conformation
Cinnamic derivatives
X-ray crystallography

ABSTRACT

There are 2.5 million cases of snakebite per year and approximately 100,000 to 150,000 deaths. Thus, it is considered an important public health problem by the World Health Organization. Snakes from the *Bothrops* genus may cause severe local effects in the victims, so it is important to develop inhibitors to treat local effects in patients. In addition, approximately 30 different species of bothropic snakes have been described that may present differences in their venom composition. Small structural differences in the venom proteins may result in different ligands binding. Herein, BnSP-7, a PLA_2 -like protein that causes local myotoxic effects, was analyzed using different biophysical techniques. Crystal structures of BnSP-7 binding to three different cinnamic acid derivatives were solved showing that the ligands bind in the membrane-dockage region (MDoS) of the protein. Spectroscopy fluorescence and microscale thermophoresis (MST) assays showed that these ligands also bind to BnSP-7 in solution and provide comparative information about their affinity to BnSP-7. MST experiments also showed that hydroxyl radicals of the ligands, involved in their binding with the MDoS region of BnSP-7, are essential to increase their affinity with the protein. As this region has been indicated as essential for the myotoxic mechanism, the ligands could potentially be used as inhibitors for BnSP-7. These results provide relevant insights to understand the PLA_2 -like proteins myotoxic mechanism and may eventually lead to design of new inhibitors for these toxins. Furthermore, a comparative structural analysis of BnSP-7 with other PLA_2 -like proteins showed that BnSP-7 has an atypical quaternary conformation, suggesting an intermediate state that is unlike other PLA_2 -like proteins. This information, combined with the absence or partial occupancy of molecules in their hydrophobic channel and the misaligned membrane-disruption region, led us to hypothesize that the protein is not able to fully exert its myotoxic activity like other PLA_2 -like proteins.

© 2017 Elsevier B.V. and Société Française de Biochimie et Biologie Moléculaire (SFBMM). All rights reserved.

1. Introduction

Envenomation by snakebites represents an important public health problem due to its frequency and the associated mortality [1]. The worldwide estimate of snake bites is 2.5 million cases per year, with approximately 100,000 to 150,000 deaths [2,3], especially in rural areas of tropical and subtropical countries in Asia, Africa, Oceania, and Central and South America [4,5]. In Latin

America, snakes from the *Bothrops* genus (Viperidae family) cause approximately 130,000 accidents per year, leading to important local effects that may lead to permanent tissue loss, disability or limb amputation [6,7]. These effects are not efficiently neutralized by conventional serum therapy [4,8]. Similarly, snakes from the *Trimeresurus* genus (Viperidae family), found in Asia, are also responsible for a high number of snakebites, mainly in rural areas of Bangladesh, Bhutan, India, Nepal, Pakistan, and Sri Lanka [4,9]. These bites also produce prominent local effects in the victims, generating high morbidity [9,10].

Metalloproteinases and phospholipases A_2 (PLA_2 s) are the toxins responsible for the local myonecrosis in accidents from snakes from Viperidae family. PLA_2 s are small (approximate molecular mass of 14 kDa) and stable proteins, which often have seven

* Corresponding author.

** Corresponding author.

E-mail addresses: fernandes@ibb.unesp.br (C.A.H. Fernandes), fontes@ibb.unesp.br (M.R.M. Fontes).

<http://dx.doi.org/10.1016/j.biochi.2017.07.009>

0300-9084/© 2017 Elsevier B.V. and Société Française de Biochimie et Biologie Moléculaire (SFBMM). All rights reserved.

Apêndice 4: Artigo aceito para publicação em *Current Proteomics* “Chaperones and Their Role in Telomerase Ribonucleoprotein Biogenesis and Telomere Maintenance”

Chaperones And Their Role In Telomerase Ribonucleoprotein Biogenesis And Telomere Maintenance

Maria Alejandra Viviescas^a, Maria Isabel Nogueira Cano^{*a}, Marcela Segatto^{1b}

^aGenetics Dept., Biosciences Institute, São Paulo State University (UNESP), Botucatu, SP, Brazil;

^{1b}Faculdade Brasileira, Multivix Vitória, 29075-080, Vitoria, Espirito Santo, Brazil



Abstract

Telomere length maintenance is important for genome stability and cell division. In most eukaryotes, telomeres are maintained by the telomerase ribonucleoprotein (RNP) complex, minimally composed of the telomerase reverse transcriptase (TERT) and the telomerase RNA (TER) components. In addition to TERT and TER, other protein subunits are part of the complex and are involved in telomerase regulation, assembly, disassembly, and degradation. Among them are some molecular chaperones such as Hsp90 and its co-chaperone p23 which are found associated with the telomerase RNP complex in humans, yeast and probably in protozoa. Hsp90 and p23 are necessary for the telomerase RNP assembly and enzyme activity. In budding yeast, the Hsp90 homolog (Hsp82) is also responsible for the association and dissociation of telomerase from the telomeric DNA by its direct interaction with a telomere end-binding protein (Cdc13), responsible for regulating telomerase access to telomeres. In addition, AAA+ ATPases, such as Pontin and Reptin, which are also considered chaperone-like proteins, associate with the human telomerase complex by the direct interaction of Pontin with TERT and dyskerin. They are probably responsible for telomerase RNP assembly since their depletion impairs the accumulation of the complex. Moreover, various RNA chaperones, are also pivotal in the assembly and migration of the mature telomerase complex and complex intermediates. In this review, we will focus on the importance of molecular chaperones for telomerase RNP biogenesis and how they impact in telomere length maintenance and cellular homeostasis.

Keywords: chaperones, telomerase, telomeres, ribonucleoprotein complex, biogenesis

1. INTRODUCTION

1.1. Telomeres

The ends of eukaryotic chromosomes, known as telomeres, have a special structure and properties and were demonstrated in early observations by Muller and McClintock [1,2]. They are essential for chromosome integrity since

they protect chromosome end termini from degradation and fusions.

Decades later, with the advance of molecular methods, Blackburn revealed that telomeric DNA consisted of Guanine-rich short repetitive sequences clearly related in many species [3,4] since even phylogenetically distant organisms such as protozoa and humans share the same